# Encyclopedia of Psychology of Decision Making

# Denis Murphy Danielle Longo Editors

NOVA

## **ENCYCLOPEDIA OF PSYCHOLOGY OF DECISION MAKING**

No part of this digital document may be reproduced, stored in a retrieval system or transmitted in any form or by any means. The publisher has taken reasonable care in the preparation of this digital document, but makes no expressed or implied warranty of any kind and assumes no responsibility for any errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of information contained herein. This digital document is sold with the clear understanding that the publisher is not engaged in rendering legal, medical or any other professional services.

# **ENCYCLOPEDIA OF PSYCHOLOGY OF DECISION MAKING**

DENIS MURPHY AND DANIELLE LONGO EDITORS

Nova Science Publishers, Inc. New York Copyright © 2009 by Nova Science Publishers, Inc.

**All rights reserved.** No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means: electronic, electrostatic, magnetic, tape, mechanical photocopying, recording or otherwise without the written permission of the Publisher.

For permission to use material from this book please contact us: Telephone 631-231-7269; Fax 631-231-8175 Web Site: http://www.novapublishers.com

#### NOTICE TO THE READER

The Publisher has taken reasonable care in the preparation of this book, but makes no expressed or implied warranty of any kind and assumes no responsibility for any errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of information contained in this book. The Publisher shall not be liable for any special, consequential, or exemplary damages resulting, in whole or in part, from the readers' use of, or reliance upon, this material. Any parts of this book based on government reports are so indicated and copyright is claimed for those parts to the extent applicable to compilations of such works.

Independent verification should be sought for any data, advice or recommendations contained in this book. In addition, no responsibility is assumed by the publisher for any injury and/or damage to persons or property arising from any methods, products, instructions, ideas or otherwise contained in this publication.

This publication is designed to provide accurate and authoritative information with regard to the subject matter covered herein. It is sold with the clear understanding that the Publisher is not engaged in rendering legal or any other professional services. If legal or any other expert assistance is required, the services of a competent person should be sought. FROM A DECLARATION OF PARTICIPANTS JOINTLY ADOPTED BY A COMMITTEE OF THE AMERICAN BAR ASSOCIATION AND A COMMITTEE OF PUBLISHERS.

#### LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Encyclopedia of psychology of decision making / editors, Denis Murphy and Daniel Longo. p. cm.
ISBN 978-1-61728-398-7 (E-Book)
1. Decision making--Psychological aspects. I. Murphy, Denis, 1964- II. Longo, Daniel R. HD30.23.E49 2009
153.8'303--dc22 2009028413

Published by Nova Science Publishers, Inc. + New York

## **CONTENTS**

Preface		xi
Expert Com	mentaries	
Α	Neuroeconomics and Decision Making: A Common Task for Psychologists, Economists and Neuroscientists José M. Martínez-Selva and Juan P. Sánchez-Navarro	1
В	End of Life Decision Making: Autonomy and Capability <i>Guido Miccinesi and Claudia Borreani</i>	5
С	Psychological Aspects of Decision Making as New Research Topics in Natural Resources Management <i>T. Hujala, T. Kainulainen and P. Leskinen</i>	9
D	New Directions in the Psychology of Decision Making Liqing Zhang	15
Ε	Psychology of Decision-Making: 6Rs for Qualitative Research Methodological Development Zenobia C. Y. Chan	21
Research and	d Review Studies	
Chapter 1	The Bayesian Statistical Approach for Decision Making in Medical Fields Roberto Bergamaschi and Cristina Montomoli	25
Chapter 2	Nurse - Physician Communication: The Impact of Disruptive Behaviors on Factors Affecting Decision Making Affecting Patient Outcomes of Care <i>Alan H. Rosenstein and Michelle O'Daniel</i>	45
Chapter 3	Time and Risk in Health-Related Decisions Monica Ortendahl	63

Chapter 4	Data-Driven Decision-Making in Clinical Settings: Applications to Unique Populations Jill D. Stinson, James C. Coleman, Niels C. Beck and Anthony A. Menditto	75
Chapter 5	Clinical, Psychological and Ethical Dilemmas: The Case of Unrelated Bone Marrow Transplantation in Adult Thalassemic Patients <i>G. Caocci, S. Pisu, A. Vacca, E. Piras, A. Ledda, R. Littera, C.</i> <i>Carcassi and G. La Nasa</i>	93
Chapter 6	Personal Meanings and Health Care Decision Making – The Case of Total Knee Replacement <i>F. Toye, J. Barlow, C. Wright and S.E. Lamb</i>	113
Chapter 7	A Constructivist Framework for Exploring Influences on Decision- Making About Treatments for Chronic Pain <i>Cary A. Brown</i>	157
Chapter 8	The Psychology of Using and Creating Video Decision Aids for Advance Care Planning <i>Muriel R. Gillick and Angelo E. Volandes</i>	183
Chapter 9	Decision Making and the Emotional Brain José M. Martínez-Selva and Juan P. Sánchez-Navarro	197
Chapter 10	A Neuroscientific Approach to the Psychology of Decision Making <i>Michael X. Cohen</i>	237
Chapter 11	Risk Avoidance and Risk Taking in Mental Health Social Work Shulamit Ramon	261
Chapter 12	Decision Making at the End of Life Kathryn A McCall and Ann Marie Rice	279
Chapter 13	Decision-Making at the End of Life: Cross Cultural Perspectives <i>H. Russell Searight</i>	297
Chapter 14	Feeling Responsible: Decision-Making in the Context of Genetic Risk <i>Holly Etchegary</i>	311
Chapter 15	Making Reproductive Decisions in the Face of a Late-Onset Genetic Disorder, Huntington's Disease: An Evaluation of Naturalistic Decision-Making Initiatives <i>Claudia Downing</i>	337
Chapter 16	Decision Making Process and Health Seeking Patterns of Young Women with Unplanned Pregnancies, Bangkok, Thailand <i>Wanapa Naravage and Rungpetch C. Sakulbumrungsil</i>	391

	Contents	vii
Chapter 17	Project Towards No Drug Abuse (TND): A Needs Assessment of a Social Service Referral Telephone Program for High Risk Youth <i>Steve Sussman, Silvana Skara and Patchareeya Pumpuang</i>	457
Chapter 18	Causes, Impact, Prevention and Medicolegal Aspects of Child Abuse and Neglect <i>B. R. Sharma</i>	475
Chapter 19	Adolescent Sexual Risk Behavior: Associations with Family, Peers and School Environment Aleksandra Jovic-Vranes and Boris Vranes	503
Chapter 20	Sexual Aggressions amongst Underaged: Vulnerabilities, Risk Factors, Signs of Calling for Help <i>Philippe Bessoles in collaboration with Myriam Linnell-Boutaud</i>	517
Chapter 21	Predictors of Middle School Youth Educational Aspirations: Health Risk Attitudes, Parental Interactions, and Parental Disapproval of Risk Regina P Lederman, Wenyaw Chan and Cynthia Roberts-Gray	531
Chapter 22	Young Children's Decisions about the Power of Inference Norman H. Freeman, Angeliki Varouxaki and Katerina Maridaki- Kassotaki	541
Chapter 23	The Role of Filial Piety in the Career Decision Processes of Postgraduate Students in China <i>Leili Jin, Mantak Yuen, and David Watkins</i>	557
Chapter 24	Reproductive Agency among Low-Income Women in Brazil Gina Hunter de Bessa	571
Chapter 25	Comparing Risky and Inter-Temporal Decisions: Views from Psychology, Ecology and Microeconomics <i>Tobias Kalenscher and Philippe N. Tobler</i>	593
Chapter 26	From Abstinence to Safer Sex: Ambiguities and Dialectics Donna E. Howard, Robin Sawyer and Jessica Brewster-Jordan	619
Chapter 27	Comparing Risky and Inter-Temporal Decisions: Views from Psychology, Ecology and Microeconomics <i>Tobias Kalenscher and Philippe N. Tobler</i>	637
Chapter 28	Deciding with whom to Mate: Do Female Finches Follow Fashion? John P. Swaddle and Earl Clelland	663
Chapter 29	Relativity of Financial Preferences: How Choice Options Influence Investment Decision Making Ivo Vlaev and Nick Chater	677
Chapter 30	The Psychology of Decision-Making in Economics: A Review <i>Jeffrey A. Livingston and Michael A. Quinn</i>	707

Chapter 31	The Conflict between Money and Self-Esteem in Investment Liqing Zhang	725
Chapter 32	Risk Mitigation: Individual and Market Behavior Jamie Brown Kruse and Kevin M. Simmons	735
Chapter 33	Frequency-domain versus Time-domain Estimates of Risk Aversion from the C-CAPM: The Case of Latin American Emerging Markets <i>Ekaterini Panopoulou</i>	751
Chapter 34	Preparing Consumers for Shared Decisions: Analysing the Effectiveness of a Decision-Aid for Women Making Choices About Birth After Caesarean <i>Allison Shorten, Brett Shorten and Sandra West</i>	767
Chapter 35	Learning, Acceptance and the Perceptions of Risks and Benefits of New Technology Developments in Spain Montse Costa and Joan Costa-Font	799
Chapter 36	Hurricane Risk Perceptions and Preparedness John C. Whitehead	811
Chapter 37	Risk Perception and the Value of Safe Rooms as a Protective Measure from Tornadoes: A Survey Method <i>Ozlem Ozdemir</i>	825
Chapter 38	Driving Dangerously: Motives Underlying Young Men's Decision to Engage in Risky Driving Robert Ho and Rita Yong Gee	841
Chapter 39	Risk Management Decision-Making in a High Risk Recreational Activity: Lessons from Mountaineering <i>Robert C. Lee and Linda S. Cook</i>	867
Chapter 40	Deciding about Risky Prospects: A Psychological Descriptive Approach Petko Kusev and Paul van Schaik	889
Chapter 41	Instability and Relativity of Preferences: How Context Determines Utilities and Decisions <i>Ivo Vlaev</i>	911
Chapter 42	Why Do People Change Their Minds? Evidence from the Purchase of Long-Term Care Insurance <i>Anne Theisen Cramer and Gail A. Jensen</i>	935
Chapter 43	Toward a Comparison-Induced Distortion Theory of Judgment and Decision Making <i>Jessica M. Choplin</i>	951

	Contents	1X
Chapter 44	Analysis and Decision-Making Models in the Process of the Assimilation of Change and Innovation in Learning Systems <i>Baruch Offir</i>	981
Chapter 45	An Error-Minimizing Rule for Animal Decision-Making under Uncertainty <i>J.A. Devenport and L.D. Devenport</i>	1013
Chapter 46	Developmental Trends in Decision Making: The Case of the Monty Hall Dilemma <i>Wim De Neys</i>	1033
Chapter 47	Charting the Course of Self Evaluations and Social Comparisons over Time Janet A. Sniezek, Phanikiran Radhakrishnan and Tahira M. Probst	1045
Chapter 48	The Effects of Perceptions on Mode Choice Maria Vredin Johansson	1073
Chapter 49	Do Outsider-Dominated Boards and Large Board Size Curtail Firm Risk Taking? Michael Graham and Alfred Yawson	1099
Chapter 50	Exploring Gender Differences in Decision-Making Using the Iowa Gambling Task Ruud van den Bos, Esther den Heijer, Suzanne Vlaar and Bart Houx	1115
Index		1135

## PREFACE

Expert Commentary A - In recent years there has been a great amount of research on the study of the neural mechanisms of economic decisions, a scientific field currently known as Neuroeconomics (Camerer, 2003). This has been promoted by the development and use of imaging techniques, particularly functional magnetic resonance imaging (fMRI), for the exploration of brain activity. Neuroeconomics is an interdisciplinary field, as is decision making, in which the efforts of psychologists and economists join up with those of the neuroscientists, resulting in the participation of research specialists of different scientific fields in a given research.

Expert Commentary B - End-of-life decision making has various ethical and psychological implications. One of the main aspects is the assumption that it is the patient who must take all decisions regarding him/herself. Clinical practice does not always fit this ideal, as each case must be taken in its legal and cultural context, with much depending also on the request of the individual. In Europe a huge variability in types and proportion of end of life medical decisions has been shown, ranging from 23% of all deaths in Italy up to 51% in Switzerland. In that population-based study (Eureld) about one third of all deaths were sudden and totally unexpected in each country. Among the remaining dying patients some decisions, that could have hastened death in the attending doctor's opinion, were made in at least one out of two cases in all countries except Italy. The highest difference between countries concerned the non-treatment decisions. The lowest proportion of that type of decision was in Italy (4% of all deaths) where the percentage of individual DNR decisions was also the lowest (16% of all non-sudden deaths), and the attitude to communicate about incurability was 4 to 8 times lower than in the other participating European countries as well as in Australia [1-4].

Expert Commentary C - Sustainable management of natural resources necessitates conscious and organised activity. When planning the use of a certain spatially explicit management unit, competing goals and interests call for multi-objective decision support. The environmental decision making also involves uncertainty concerning available information and future predictions. Traditionally, a typical procedure has included the use of preference elicitation tools, statistical models, and optimisation methods, to find an acceptable solution for the decision problem at hand.

When improving the decision support portfolio to work more efficiently, there is a need to consider the overall structure of decision problems. To demonstrate that, we have conceptualised the variety of forestry decision problems in three-dimensional decisionproblem space, which comprises geographical, temporal, and social dimensions. The crucial issue in the problem structuring and, thus, in the selection of decision support method, is scope awareness: recognising the fundamentals of the decision problem at hand. With the aid of that, the unwanted decision-model-based or support-methodology-based psychological biases can be avoided or smoothened.

By the results of the research proposed in this chapter, both the efficiency of single phases of decision making and the quality of the whole decision process would be improved. The results would also enhance learning about the characteristics of decision problems and decision makers' preference structures.

Expert Commentary D - The psychology of decision making has been gaining more attention than before. It is growing rapidly. Three main perspectives of the psychology of decision making include the standpoint of cognitive psychology, the influence of social psychology and the viewpoint of neuropsychology. The present commentary reviews briefly these three major approaches to the psychology of decision making, points out their contributions and limitations, and suggests new directions.

Expert Commentary E - The phenomena and facts of decision-making are closely related to our daily life, and the what, how and why of our psychological responses during the course of this process have fascinated psychologists throughout academic history. This commentary suggests the potential of qualitative approaches to research into the psychology of decision-making, and proposes a 6Rs framework to ensure good practice in the preparation and implementation of these studies. Finally, existing research should be recognized, and future researchers should be encouraged to consider the ability of qualitative inquiry to enrich our understanding of decision-making and other areas of psychology. (97 Words)

Chapter 1 - Medical controversies, such as the evaluation of different risks of disease or the efficacy of different treatments, are often faced with frequentist statistical analyses.

However, conventional statistics are someway improper for such decisions, because they dichotomize results according to whether they are or are not significant and do not allow decision makers to take account of additional evidence, such as previous knowledge, biological plausibility, biases in the studies, etc. A Bayesian approach can overcome these problems.

The essential components of Bayesian analysis are based on existing evidence available from other studies or from expert opinion. This evidence is summarized quantitatively via a *prior distribution* for existing beliefs. The real data are used to build a *likelihood function*. The *prior distribution* and the *likelihood function* can then be formally combined to give a *posterior distribution* for the parameters of interest. Finally, this can be interpreted as the updated beliefs or evidence about the effects of interest. Data presented as a series of posterior probability distributions are a better guide to policy. For these reasons, Bayesian methodology has been promoted as an alternative to the frequentist one when dealing with medical forecasting.

Chapter 2 - Disruptive behaviors can have a profound effect on staff perceptions, attitudes and reactions that affect decision making and communication flow. Feelings of anger, hostility and frustration lead to impaired relationships, confused expectations, and unclear roles and responsibilities which can impede the transfer of vital information that can adversely affect patient outcomes of care. Health care organizations need to be aware of the significance of disruptive behaviors and develop appropriate policies, standards and procedures to effectively deal with this serious issue.

Chapter 3 - Classical decision theory, which deals with situations in which a decisionmaker can choose between different actions, is based on the effect of probability of outcome and value of outcome on judgments and decisions. The importance of the time factor has been ignored in decision theories, and in preventive procedures and clinical decision-making as well the impact of time has largely been neglected. Most work on inter-temporal choice is based on a trade-off between two outcomes of different values over different periods. Although risk is known generally to have an important impact on judgments and decisions, the effects of time have not been studied to the same extent. However, the concepts of time and risk are both important to health-related decisions. The concept of risk is analyzed and a distinction is made between personal and general risk. The intricate relationships between time and risk imply that subjects are risk aversive, with the individual's risk aversion intensifying the sooner the outcomes of choices will occur in the future. Time-related aspects of risk have a potentially large impact on health-related decisions and responses to forecasts of morbidity and mortality from disease in the future.

Chapter 4 - Traditional assessment instruments designed to predict risk, identify appropriate treatment interventions, and inform placement decisions typically rely on heterogeneous normative samples that generalize to a broad range of clinical settings. While this approach is useful in providing empirical data for clinical decision-making, it lacks the specificity that is sometimes necessary in order to address the needs of unique client populations. An additional problem relates to the methodological techniques used to derive prediction equations, which primarily involve linear statistical models and univariate comparisons. We advocate the use of multivariate models and innovative statistical methods to more fully address the complex factors which drive the behaviors of interest in a clinical population.

Here we discuss several examples which are relevant to this issue, including prediction of aggression within a forensic psychiatric institution, assessment and treatment of problematic sexual behavior in individuals with intellectual and developmental disabilities, and prediction of elopement from public psychiatric facilities. For each of these, we address the use of heterogeneous samples to inform decision-making and how this process may be enhanced through the use of creative statistical procedures (e.g., latent class analysis, logistic regression) and sampling techniques more specific to institutional needs.

Chapter 5 - Thalassemia mayor is a chronic, progressive and, in the long term, lifethreatening condition. Extensive ongoing medical care in this pathology (regular transfusions and iron-chelation) can have a major impact on the physical, psychosocial well-being and quality of life (QoL) of patients and their families. Allogeneic bone marrow transplantation (BMT) from an HLA identical sibling is currently the only treatment available to cure these patients. Unfortunately, the probability of finding an HLA-identical donor within the family is less than 30%. In all other cases, it is necessary to search for a compatible donor among volunteer donors worldwide. Unrelated BMT in adult thalassemic patients is burdened by a transplant-related mortality of about 30%. For this reason, patients without a sibling donor need to make a difficult and radical decision. They can either continue traditional transfusion and chelation therapy, with no chance of complete recovery, or, they can accept the high risk of BMT in the hope of obtaining a definitive cure for the disease. This circumstance gives rise to several social, psychological and ethical problems.

We report the results obtained in 31 adult thalassemic patients transplanted from unrelated donors. To better understand the decision making process, we investigated the motivation factors, the communication strategies, the post-transplantation quality of life (QoL) and discussed the ethical issues of this therapeutic approach. Twenty-four patients (77.4%) are alive and transfusion independent. Seven patients (22.6%) died from transplant-related causes. The surviving patients were given two questionnaires to evaluate pre-transplantation motivation, communication factors and psychological aspects of decision-making. The global QoL was also assessed. All patients were strongly motivated and satisfied with the communication modalities of physicians before transplantation. None of the patients regretted their choice and would make the same decision if reliving the experience. All patients had a good Qol.

The central ethical issue of BMT in a non malignant disease can be framed in terms of a conflict between two fundamental principles of medical ethics: non-maleficence and beneficence. Given the fact that many situations present in medicine where benefits cannot be achieved without at the same time inflicting harm, the question arises as to who will be responsible for the decision and according to what criteria can a fair balance of benefits and burden be struck. The other topic is based on respect for patient's autonomy.

The issue is not simply "to cure" but "to take care" of the patient as a whole person.

The multi-disciplinary approach used in this study seems to offers a valuable contribution to the decision-making that surrounds the choice of treatment with a high mortality risk in a chronic non malignant disease.

Chapter 6 - There seems to be a mismatch between the severity of knee symptoms and the decision to undergo (total knee replacement) TKR. Personal meanings are important because decisions regarding need for TKR do not seem to be explained by symptoms alone. This study explores patients' personal meanings of knee osteoarthritis (OA) and TKR. 18 semi-structured interviews were conducted with a purposive sample of respondents who were listed for TKR at one specialist orthopaedic hospital. Data were analysed using Interpretive Phenomenological Analysis. Results suggest that the decision to undergo TKR is not related to symptoms alone, but to personal meanings. Some of these personal meanings may not be useful in accurately assessing need for TKR, and may result in mis-targeting of treatment.

Chapter 7 - This chapter will first present the theoretical underpinnings of a constructivist analysis of chronic pain and then discuss more specifically the relevance of the derivative Consciousness Model. The current literature will be presented to illustrate the applicability of a constructivist analysis of people's chronic pain experience and the decisions they make about treatments for their pain. Particular attention will be placed on exploring in greater depth the four components of the Consciousness Model (coherence, sense of self, purposiveness and affect). The second part of the chapter will present the outcomes of a research study that used Delphi methods (iterative rounds of feedback and mixed methods questionnaires) to explore the utility of the Consciousness Model in understanding how service users and providers construct meaning for chronic pain and how, in turn, that meaning influences decision-making about treatment. The findings presented in this chapter are extracted from a larger multi-stage study carried out in the United Kingdom with the multidisciplinary membership of the British Pain Society and three chronic pain sufferers' support groups in the North-West of England. The aim of this study was to examine how service users and providers made decisions about which treatment components they believed were important for chronic pain. Stage one of the large study identified, via postal survey, what treatment components service providers and service users identified as 'important'. In Stage two a constructivist framework was applied to develop an iterative Delphi questionnaire exploring service users' and providers' rationale for their choices. Stage three also employed a Delphi process and asked participants to reflect on their previous comments and the researcher's analysis of these data. Stages two and three used postal questionnaires and summary reports from preceding rounds for background information.

The study findings revealed that there were statistically significant differences between what service users and service providers perceived as influencing their decision-making. While service user findings showed that the influences on decision-making were fairly evenly distributed across all four of Chapman's Consciousness Model domains this was not the case for service providers. Most service providers clearly felt that *affect* and *self-image* had little influence on their decision-making. Rather they more strongly endorsed the categories of *coherence* and *purposiveness*. An explanatory model has been derived from these findings and is presented in the latter section of the chapter. The model details how differential access to control and knowledge can play a strong mediating role on all aspects of consciousness and consequently on decision-making.

Lastly the chapter will discuss that although a constructivist paradigm advances our understanding of influences on decision-making in chronic pain it remains an insufficient explanatory tool. The chapter concludes by identifying how the emerging transtheoretical and complex adaptive systems paradigms hold promise for advancing our knowledge about decision-making yet further in relation to what is defined, by progressive thinkers grappling with the post-modern issues of health in the 21<sup>st</sup> century, as the 'fifth vital sign': pain.

Chapter 8 - Decision-making in the medical context has traditionally focused on individual patients choosing between various medical treatments that have been proposed for an acute medical problem. For example, a pregnant woman may need to decide whether or not to undergo amniocentesis to rule out Down's syndrome in her fetus, or a man with an elevated prostate specific antigen test may need to choose between "watchful waiting" or radical prostate surgery. Much of this research has focused on how patients process information and how physicians can objectively present such information to improve decision-making.

Advance care planning is the process in which patients plan for their future medical care. Patients asked to consider the benefits and burdens of various interventions that might be considered for some hypothetical disease state, such as dementia. Advance care planning is advocated largely in order ensure that the treatment the patient receives is consistent with his own preferences—even if he is incapable of making decisions at the time of illness.

While advocated on theoretical grounds, advance care planning has not been widely adopted, either by patients or physicians. A major barrier to engaging in advance care planning is that patients are often unfamiliar both with the hypothetical disease state and with the specific medical interventions that might be used to treat them. Video has previously been used as a decision aid for patients facing real time decisions about their medical care. The application of video to advance care planning discussions is a new approach that has the potential to help surmount the barrier to informed decision making.

Video is a powerful and widely accessible medium that engages patients in a way not easily captured with words. Video allows patients to imagine hypothetical health states and can provide an accurate portrayal of medical interventions. This chapter will review the emerging corpus of work that uses video as an adjunct to advance care planning, specifically in planning for the possible development of dementia. It will analyze the biases that can arise from the use of an aesthetic and visual medium and discuss ways those biases can be avoided. The chapter will make the case that video is useful for patients with low health literacy and patients from minority populations. Finally, other specific disease entities that lend themselves to video portraits will be proposed.

Chapter 9 - There is an increasing amount of empirical data which give support to the idea that affective processes play an important role when making decisions. The main purpose of this chapter is to describe the brain mechanisms that support and/or mediate the involvement of emotion in decision making processes. A brief overview is presented of the theoretical approaches explaining how emotion influences decision making, such as the prospect theory, the somatic marker hypothesis, and other recent theories. The brain mechanisms involved in decision making and the processing of emotional information are reviewed. Particular attention is given to the role of the ventromedial and dorsolateral prefrontal cortices, anterior cingulate cortex and amygdala in decision making, and to the brain defence system and the dopaminergic mesocortical-mesolimbic system, together with their projections and related structures, in emotion. Research results from clinical studies, e.g. brain injury and substance abuse, and from the field of Neuroeconomics which are relevant to the topic are also discussed.

Chapter 10 - Considerable research in cognitive psychology, economics, and marketing has focused on the psychological mechanisms of decision-making. Coming from a different perspective, the past few decades have seen considerable research on the neurobiological mechanisms of reward processing, learning, and addiction. Can these neuroscience findings shed light on the psychology of decision-making? Here I review the neuroscience literature, ranging from physiological studies of single neurons in rats to noninvasive neuroimaging in humans, as it relates to and informs human reward-guided decision-making. I focus specifically on the nucleus accumbens, orbitofrontal cortex, and hippocampus, structures with critical roles in the flexible adaptation of decision-making behavior based on rewards and novelty. The major neurochemical system involved in reward-guided decision-making is dopamine, which originates in the midbrain, is transmitted throughout the striatum and frontal cortex, and provides a global signal about rewards and errors in reward prediction. Dysfunctions of the dopamine system might underlie impairments in decision-making such as impulsivity and irrational decision-making, as seen in, for example, substance abuse. These structures do not operate in isolation, but rather form a tightly linked network, in which information is passed from structure to structure, and is dynamically gated in the nucleus accumbens, such that incoming information is permitted throughput only during specific neural events. Consideration of these systems as functional circuits has the potential to shed light on the neurobiological processes that underlie psychological decision making, as well as impairments in decision-making, such as those seen in addiction disorders.

Chapter 11 - This chapter will look at issues related to risk in mental health social work, especially the over-focus on risk avoidance and the need to create space for calculated risk taking. Examples from Britain and continental Europe will be utilised.

Chapter 12 - Palliative care is becoming an increasingly accepted approach to care when cure is no longer possible. Decision making during the palliative phase of any illness can be complex, involving integration of knowledge, legal & ethical considerations, & excellent communication skills for healthcare professionals, patients and, where appropriate, their family/carers. This decision making process may become even more challenging as the focus of care changes from uncertainty in terms of survival or time to progression of disease to the often highly charged emotional terminal stage of the illness.

Patient participation is increasingly purported as a goal in end of life care, and research suggests that most patients prefer a collaborative role in decision making. However there is a paucity of information about how those nearing the end of life can actively engage with healthcare professionals & family members to achieve this in such diverse areas as personal care; symptom management; hydration & nutrition; advance directives; compliance with medical advice; and place of end of life care.

The patient's right to self-determination is evidenced in the increasing use and possible questionable legality of advance directives, both written & documented oral statements, and in the ongoing current debate regarding physician-assisted suicide. Traditional decision making roles are changing and the paternalistic medical model, which has previously led much end of life care, is having to be modified, challenging professional caregivers to help the decision making process to be more collaborative.

This chapter aims to explore the psychology of decision making in end-of-life care where competency to make decisions may be threatened by changes in cognitive ability due to advancing disease, the unique psychosocial stressors related to terminal illness, or altered judgement related to the use of high dose drugs in symptom management.

Palliative care is becoming an increasingly accepted approach to care when cure is no longer possible. Decision making during the palliative phase of any illness can be complex, involving integration of knowledge, legal and ethical considerations, and excellent communication skills for healthcare professionals, patients and, where appropriate, their family/carers. This decision making process may become even more challenging as the focus of care changes from uncertainty in terms of survival or time to progression of disease to the often highly charged emotional terminal stage of the illness.

Patient participation is increasingly purported as a goal in end of life care, and research suggests that most patients prefer a collaborative role in decision making. However there is a paucity of information about how those nearing the end of life can actively engage with healthcare professionals & family members to achieve this in such diverse areas as personal care; symptom management; hydration & nutrition; advance directives; compliance with medical advice; and place of end of life care.

The patient's right to self-determination is evidenced in the increasing use and possible questionable legality of advance directives, both written and documented oral statements, and in the ongoing current debate regarding physician-assisted suicide. Traditional decision making roles are changing and the paternalistic medical model, which has previously led much end of life care, is having to be modified, challenging professional caregivers to help the decision making process to be more collaborative.

This chapter aims to explore the psychology of decision making in end-of-life care where competency to make decisions may be threatened by changes in cognitive ability due to advancing disease, the unique psychosocial stressors related to terminal illness, or altered judgement related to the use of high dose drugs in symptom management.

Chapter 13 - Increased cultural diversity in the U.S. has raised challenges for health care law, ethics, and clinical decision making. In the United States, medical ethics emphasizes individual autonomy by encouraging patients to obtain all relevant information about their medical condition from health care providers and to use that information to choose treatment options. The federal 1990 Patient Self-Determination Act (PSDA) mandated health care institutions to discuss advance directives (AD) and to encourage patients to establish an AD to take effect if the patient is no longer able to communicate their medical decisions. Some ethnic communities did not find either type of AD, the living will or durable power of attorney (DPOA), to be consistent with prevailing cultural norms. Asian- and Hispanic-Americans were less likely than White European- or African-Americans to believe that patients should be directly informed about a terminal illness such as metastatic cancer. In these ethnic groups, families often received information about the illness, its prognosis and treatment options rather than the patient, themselves. These ethnic groups often saw informing the patient of serious illness and requiring the patient to make independent decisions as burdensome and isolating. The increased discussion of control over end of life care also highlights a legacy of distrust of the healthcare system among many African-American patients. From the perspective of decision-making models, patients and families appear to rely on arguments influenced by personal narratives to make health care decisions, while physicians are likely to rely on probability and clinical experience as a guide for developing their own recommendations. A more culturally sensitive approach would be to encourage patients to choose if they want information about their condition, decisional control, and who they would prefer to have made their healthcare decisions.

Chapter 14 - Decision-making in the context of genetic risk is normally constructed as an opportunity for choice. For example, a person at risk for a genetic disorder could choose to accept or decline a genetic test. Since genetic risk has implications for other family members, however, perceptions of responsibility to important others could constrain some of the choices to be made around genetic risk. Drawing upon semi-structured interviews with at risk people and their family members, this chapter investigates decisions about genetic risk for Huntington disease (HD) - a fatal genetic disorder. In particular, it explores perceptions of genetic responsibility since qualitative data analysis suggested that decisions about genetic risk were often influenced by obligations to other family members. For example, some at risk people felt responsible to determine their genetic risks through testing, particularly for their at risk offspring. Responsibility to current and future partners, to plan for a future that might include HD and to communicate genetic risk to other family members also emerged as important dimensions of genetic responsibility. It is argued that perceptions of genetic responsibility constrain some of the choices of those who live with genetic risk having implications for test decisions, post-test adjustment and family relationships. Theoretically, the findings suggest that decision-making models might include measures of perceptions of responsibility to better account for their influence on decisions. At least in the context of genetic risk, a consideration of perceptions of responsibility to others could provide a richer picture of decision-making.

Chapter 15 - This chapter makes two contributions to the psychology of decision making. It draws on empirical work to document how family members make and live with reproductive decisions as they become aware of their risk for a serious late-onset genetic disorder, Huntington's disease (HD). Decision-making involves negotiating two dimensions of reproductive risk – that for any child which might be born and the uncertainty that arises about the at-risk parent's ability to sustain a parenting role should he or she become symptomatic. A detailed account is given of how the model of responsibility was generated from their narratives. This model encapsulates what families find important when making reproductive decisions. It demonstrates that how people make decisions can become as important as what they decide. Examples are given to show how the model provides a framework for comparing how people deal with each dimension of risk, to compare different people's decision-making and to illuminate decision-making in the face of change. This

includes how they address new options generated by recent developments in molecular genetics which can resolve uncertainty about, or avoid these risks. A detailed account is also given of the process of their decision-making. These findings are used to evaluate claims being made by naturalistic decision-making (NDM) initiatives to account for decision-making in the real world. An outline of this initiative prefaces the research findings. Beach's (1990) image theory, used to illustrate the NDM approach, emphasises the role that values play in forming, negotiating, implementing and living with decisions that arise in everyday lives. Participants' accounts support this claim - revealing how values such as responsibility become established and contribute to decision-making. This and other findings lend support to the NDM claims. The chapter concludes with suggestions about how the model might be used more generally to further our understanding of the psychology of decision-making in the face of risk.

Chapter 16 - *Objectives*: This study examined the decision making process and health seeking patterns of low-income young women with unplanned pregnancies who opted for abortion, putting baby up for adoption, or keeping the baby.

*Methods*: Research methodology in this study focused on formative research utilizing qualitative data. The study was conducted covering all five shelters and low-income communities in Bangkok, Thailand. The data collection was through Focus Group Discussions, and in-depth interview. Forty-five cases were purposively recruited into the study. The data were analyzed using content analysis.

*Results*: From the qualitative research, the results indicated that the majority of the young women tended to select abortion as their first choice, while a few cases continued their pregnancies to term without any attempt to terminate the pregnancy. These women tended to delay seeking assistance once they found out their pregnancy situation. Majority of the low-income young women were more likely to resort self-medication as their first alternative, which was sometimes hazardous to their health. Decision on self-medication or abortions with unskilled personnel sometimes resulted in serious and life-threatening complications. Many pregnancy-termination situations resulted from self-medication. While most of these women made the important decision themselves, they still sought advice, guidance, and support from their partners, peers and parents. When the first attempt failed, they would seek a second or third attempt until they felt it was not possible to achieve what they had planned.

Most women with unplanned pregnancies knew that they had options regarding termination of their pregnancies but their main concern was confidentiality. Consequently, the women who wanted to terminate their pregnancies adopted three patterns of action, i.e., 1) visiting drugstores / grocery stores, 2) visiting private clinics or hospitals, and 3) using physical pressure or vigorous actions. Most of the women realized that private clinics provided effective methods for terminating pregnancy, but due to its high cost, they would first resort self-medication or self-management. If they were not successful, they would then visit a private clinic and asked others to support the cost.

Understanding the women's decisions making process and their health seeking patterns utilization allows us to understand their decision and course of actions. The program managers, implementers, providers, partners, parents/relatives, and friends should do as much as possible to support the decision of the women in order to provide better information and services to reduce the impact, both physical and mental, of their selected choice.

Chapter 17 - This chapter describes a quantitative and qualitative needs assessment of a potential social service resource telephone program component among high risk youth who

received the Project Towards No Drug Abuse (TND) classroom-based program (approximately 1-year earlier). Information was obtained to determine whether the targeted youth would be interested and receptive, or even need the information available from such a program. Results supported youths' overwhelming receptiveness of a social service referral program. The vast majority of respondents indicated a strong desire for resource and referral information on vocational, educational, recreational, transportation, and mental health and drug counseling. Participants' responses will be used to better structure and tailor our booster program. Further research is needed to investigate the effectiveness of the provision of social service resource information on drug use among emerging adults.

Chapter 18 - Abuse and neglect of children and adolescents continue to be major public health problems all over the world. It has been found to be correlated with emotional and behavioral problems, suicidal behavior, substance abuse, and many psychiatric disorders including post-traumatic stress disorder. However, despite the high prevalence and the mental health consequences of childhood abuse for the victims, insufficient efforts are being made to recognize these consequences and there are too few resources available to provide the needed mental health services to these victims and their families. On the other hand, exciting advances have taken place regarding care of sexual abuse victims. Clinical research has demonstrated the effectiveness of psychotherapies and suggested the efficacy of medications in ameliorating the behavioral and emotional consequences of childhood abuse. Specific clinician interviewing strategies have been developed to avoid suggestion and interference with sexual abuse-associated litigation. However, psychiatry and allied mental health professionals have much to contribute in terms of advocacy for abuse victims and an increased understanding of the need for provision of rehabilitation opportunities to these children. Medical professionals can also contribute to lower the high acquittal rate in cases of child sexual abuse by proper documentation of the history and examination findings.

Chapter 19 – Objectives: During the dynamic period of adolescence when the passage from childhood to maturity takes place, sexuality takes on new dimensions; feelings become more intense, relationships become more complex, and the consequences of sexual behavior are radically altered. In general, earlier puberty, later marriage, a decline in the family leading to less control and more autonomy, and intense exposure to sexual stimuli via the mass media and travel across cultural boundaries have made pre-marital adolescent sexual activity more common. Adolescent sexual risk behavior places them at risk of unwanted pregnancy and childbirth, induced abortion in hazardous circumstances, HIV infection, and other sexual transmitted diseases. Objective of the study was to achieve a picture of adolescent sexual risk behavior. It also provides better material for the planning of specific preventive activities.

Methods: Random cluster sample consisted of 1540 15-year old adolescents, 822 female and 718 male. The multiple choice questionnaire included 119 items, covering topics such as: demographic characteristics, psychosocial determinants of health, leisure time behavior, family and peer context, risk behavior (tobacco use, alcohol and drugs consumption, sexual activity), and perceptions of school and the school's influence. Descriptive statistics, Chisquare test and logistic regression were applied for statistical analyzes.

Results: Thirteen percent of adolescent reported that they had sexual intercourse before age 15 years. Average age of first sexual intercourse was 13.88. Average number of sexual partners they had was 3.10. One out of three never used condom. Eight percent reported multiple sex partners, and fourteen percent alcohol use before intercourse. Factors associated

with adolescent sexual risk behavior include other form of risk behaviors, family and peer context, and school environment.

Conclusion: Adolescent relationships are complex and future research should consider not only causality of adolescent sexual risk behavior, but also the etiology of the satellite behaviors.

Chapter 20 - Post-trauma notion of vulnerability, following a sexual aggression, appears too general in infantile clinical psychopathology to be sufficiently operating. Differences regarding criteria for symptomatological assessments, differential diagnosis with post-trauma disorders, personality changes inherent to child development, specificity (or not) of a posttraumatic semiology inherent to under fifteen, etc. add to the complexity of the question. According to research, vulnerability can be, at once, synonym of psycho-social factors of risk, signs of victimarius call for help or badly circumscribed psychological vulnerabilities. International studies often diverge as investigation methodologies are so contradictory and epidemiological data not consensual enough.

Chapter 21 - School-wide surveys in five middle schools were used to measure educational aspirations, attitudes toward sexual health risk behaviors and drug use, and perceptions of parental interactions and disapproval of risk behavior at baseline and one year later. Participants were male and female students of Black (n = 222), Hispanic (n = 317), White (n = 216), and Asian or other heritage (n = 85), ages 11 to 14. Analyses were performed for three factors with Cronbach's alpha coefficients  $\geq 0.65$  (youth's attitudes, discourse with parents, and parents' disapproval of risk behavior), and three single items inquiring about use of alcohol, use of marijuana, and sexual behavior. Generalized Linear Model (GLM) with logit link was used to evaluate the contribution of these measures at baseline as predictors of educational aspirations at the one-year follow-up. Results showed race/heritage (p < .001), attitudes toward health risk behaviors (p < .01), extent to which youth talked with parents about use of drugs and other health risk behaviors (p < .05), and perceptions of their parents' disapproval of risk behavior (p < .05) each made significant contributions in predicting educational aspirations. Gender did not contribute to the prediction of educational aspiration nor did self-report of actual risk behavior. These results indicate that youth interactions with parents regarding health risk behaviors is worthy of further exploration to develop interventions to reduce adolescent health risks and increase educational aspirations.

Chapter 22 - Seeing in a box is a direct way of knowing what is in the box. There is evidence that 4-year-olds engage with the idea that equally reliable knowledge can also be gained indirectly, via inference. For example, if you see that there is one cup to each saucer, the cups can be put away out of sight, and just by counting the saucers you can infer the number of the uncounted cups. But sometimes it is possible to make an inferential mistake. We review recent evidence of young children's decisions on whether people know things via inference, and add new evidence to test a claim about a false inference test. That test involved asking children to judge another's knowledge through inference when the other was misled to input the wrong premises into her calculation. The finding that children performed better in the false inference task compared to the true inference task, together with their explicit verbal justifications, attests to children's theoretical understudying of inference. We suggest that that is the growth point for the next round of research.

Chapter 23 - The research attention paid to career decision making among Mainland Chinese adolescents and college students is of recent origin. Such attention is mainly due to

major changes in two salient policy areas by the government of the People's Republic of China (PRC). Firstly, the labor distribution system has been changed from a planned to a market economic system in which employers and graduates are free to choose each other. Before the middle of the 20th century in the PRC, college graduates had to accept the jobs assigned by the government. The freedom of selection by both employers and fresh employees has stimulated Chinese graduates to consider how to choose a suitable career for their own sake. Secondly, higher education in the PRC is undergoing a radical transformation from an elite to a mass system. The increasing intake of the youth into college had led to the sharp competition in the labor market, which in turn has produced the employment pressure among college graduates. The number of college graduates pouring into the labor market each year has increased from 1.15 million in 2001 to 4.2 million in 2006 (China Ministry of Education, 2005). Meanwhile, graduate school admission has become a top priority for the majority of undergraduates: 0.714 million senior undergraduates sat for the national entrance examination for a Master's Degree in January, 2007. Although the need for career and other counseling services in the Mainland is continually increasing, it is obvious that the antecedents of adaptive career decision making among Chinese postgraduates still remain unclear.

Chapter 24 - In the United States, discussions about reproduction, contraception, and abortion often are framed within the language of "choice." Feminists, however, have long argued that the dominant ideology of choice in family planning obfuscates the larger social and institutional forces which structure and constrain the choices people can make. Many scholarly debates in social sciences, too, have shifted between theories of active and passive agency, that is, between hypotheses that posit a rational actor making calculated fertility choices (often in "modern" societies) versus those that see people passively following norms and conventions (usually attributed to so-called "traditional" societies).

In this article, I use data from ethnographic fieldwork in one low-income, urban neighborhood in Brazil to examine poor women's reproductive decision-making. Through consideration of how these women understand their reproductive options, I illustrate the social and cultural forces in the home and community that structure reproductive "choices." In particular, I highlight women's decisions to permanently end their fertility through surgical sterilization. I argue that the decision to undergo sterilization is conflicted and constrained, but is an expression of these women's active agency.

The Brazilian context is especially interesting for examining issues of reproductive choice in relation to female surgical sterilization. Brazil registers some of the highest rates of female sterilization in the world. The reasons for this are complex, but cannot be attributed directly to family planning programs or policies, as is the case in some other developing countries. While most instances of sterilization in Brazil are voluntary, Brazilian women's reproductive choices are clearly subject to myriad structural and cultural constraints. An examination of the factors women take into account when making reproductive decisions, such as the decision to permanently end child-bearing, point to the centrality of social context and social networks in their lives.

Chapter 25 - When making decisions between different options, we often consider two basic properties of these options, how risky they are and when they will occur. For example, we may choose to gamble or to wait for a larger reward. Decisions under risk refer to decisions among known probabilistic options, inter-temporal decisions refer to choices between options that will be realized at known future timepoints.

xxiii

Risky and inter-temporal decisions have been captured theoretically primarily by Ecology and Microeconomics but findings from Behavioral Economics, Psychology and Neuroscience often contradicted theoretical predictions. As a consequence, a wealth of more descriptive models has emerged to explain the findings. A subset of these models has stressed the similarities between risky and inter-temporal decisions. In this chapter we review both core theoretical approaches and empirical findings. We discuss possible explanations for discrepancies and identify key behavioral experiments.

Chapter 26 - Despite a steady decrease in teen pregnancy, birth rates and abortion since 1991, adolescents in the United States continue to experience unintended pregnancy at higher rates than their peers in nearly all other industrialized nations and are at high risk for Sexually Transmitted Infections (STIs). At the same time, debate rages over whether comprehensive sexuality education or abstinence-only curricula should be taught in schools. Additionally, there is growing concern, among adolescent sexual behavior researchers, that our limited ability to affect changes in sexually risky behavior may be related to discordance between the survey questions pertaining to adolescent sexual behavior and actual youth practices. That is, research suggests that a significant proportion of data purporting to accurately gauge adolescent participation in specific acts of sexual behavior may be not only incomplete but flawed. This chapter will review the debate over sexuality education, examine the language used on current national health surveys that assess the sexual behavior of adolescents and young adults and provide recent findings from an exploratory study which examined the denotative meaning of sexual terms among a convenience sample of late adolescent university students. In this study, it was found that there still exists little consensus among adolescents about the behavioral referents for the terms safe and safer sex. Information in this chapter will help inform the development of rigorous investigations to further examine ambiguities surrounding the language and practice of sexual behavior among adolescents and consider how this imprecision affects policy and programmatic decision making.

Chapter 27 - When making decisions between different options, we often consider two basic properties of these options, how risky they are and when they will occur. For example, we may choose to gamble or to wait for a larger reward. Decisions under risk refer to decisions among known probabilistic options, inter-temporal decisions refer to choices between options that will be realized at known future timepoints. Risky and inter-temporal decisions have been captured theoretically primarily by ecology and microeconomics but findings from behavioral economics, psychology and neuroscience often contradicted theoretical predictions. As a consequence, a wealth of more descriptive Models has emerged to explain the findings. A subset of these models has stressed the Similarities between risky and inter-temporal decisions. In this chapter we review both core theoretical approaches and empirical findings. We discuss possible explanations for discrepancies and identify key behavioral experiments.

Chapter 28 - Choosing an appropriate mate is one of the most important decisions that any animal has to make. The traditional view in non-human systems is that animals are largely slaves to their genes and an individual's mate choice is handed down from their parents. However, in recent years it has become clear that many animals show active decision making in who to mate with and that females may copy mate preferences from other females in the population. In other words, females' mating decisions are affected by the current fashion in their population. Here, we explore whether "mate choice copying" occurs in a model monogamous mating system—the zebra finch. Females were given the opportunity to observe another female courting a particular type of male (we manipulated male appearance by placing small colored leg bands on each bird). In preference tests, our focal females significantly shifted their mate preferences towards the type of male that they had observed as being courted by other females. Therefore, female finches do seem to copy mate preferences, implying that there is social inheritance of information that fundamentally affects mating decisions. This is one of the first demonstrations of mate choice copying in any monogamous system and implies that many other birds may also use social information to affect their mating decisions. We need to rethink evolutionary models of mate choice and sexual selection incorporating this form of social decision making process.

Chapter 29 - We report a study in which methodologies from psychophysics are adapted to investigate context effects on financial decision making related to retirement savings and risky investment. The aim was to determine how the range of the options offered as possible saving rates and levels of investment risk influences decisions about these variables. The respondents were presented with either a full range of choice options or a limited subset of the feasible options. The study was conducted on a sample of working people, and we controlled whether the participants can financially afford in their real life the decisions taken in the test. The results showed that choices of saving and risk are affected by the position of each option in the range of presented options. Various measures of risk aversion did not account for the risk taken in each condition. Only the simplest and most direct risk preference measure was a significant predictor of the responses within all contexts (conditions), although the actual choices were still very much influenced by each context. Thus, the results reported here suggest that judgments and choices are relative, rather than absolute, which corroborates, in a more applied and realistic setting, previous related work with abstract gambles and hypothetical risky investments.

Chapter 30 - Economics has always focused on how individuals make decisions. Traditionally, the discipline has viewed individuals as rational agents maximizing their own utility. However, economists have recently begun to incorporate research from the field of psychology in creating a richer view of decision making. This push is the result of challenges to the neoclassical model made by theoretical advances such as Kahneman's Nobel-winning prospect theory model and from the field of experimental economics. This growing field has revealed many aspects of human behavior that cannot be explained by traditional economic models. Some of these aspects of behavior include loss aversion (as explored by Kahneman); relative deprivation (the theory that individuals consider their relative position as compared to others when making decisions), motivations of altruism, fairness, and reciprocity; and the endowment effect (individuals tend to value goods more highly if they are already in possession of them). These innovations have impacted economists' views of issues such as consumption, worker-firm relations, labor supply, equities and real estate.

This paper reviews the impact of psychology on economic models of decision making. The major trends will be discussed, along with implications that these changes have for both economics and public policy.

Chapter 31 - The present theory proposes that investors not only think of future monetary benefits, but also value the choices' implications regarding their self-esteem in decision making. Self-esteem is one's subjective evaluation of the self. Most people want to maintain a positive self-image. When they decide to invest in a project, people expect to receive financial rewards, and they also hope to enhance their self-esteem through the success of the project. Thus, when their initial investment produces negative economic return, they not only suffer

financial loss, but also encounter challenges to their self-esteem. They can withdraw from the project to minimize their monetary loss, or they may keep throwing additional money into the project to demonstrate that their initial decision was correct. It is painful to admit a mistake because it poses negatively to the investors' self-concept. As a result, investors may be entrapped within a losing project and suffer accumulated financial loss. The present theory suggests that when investors encounter conflicts between money and self-esteem in decision making, they may choose to give up money in order to defend their self-esteem.

Chapter 32 - Wind research has focused mainly on the meterological aspects of storms and the engineering challenges to construct resistant structures. Over the last several years, however, an effort has been launched to understand the behavior of people living in areas threatened by violent windstorms. The aftermath of a significant windstorm event brings instant media attention to the affected area and government aid is usually forthcoming to lessen the financial blow to residents and businesses. It is rare, but some storms produce damage so severe that it alters the economic conditions, not only of those living and working in the path of the storm but third party insurers as well. Hurricane Andrew is a relatively recent example. As losses from these events mount, several questions continue to arise. What can be done to lessen the effect of a tornado or hurricane? And second, will people adopt mitigation measures that may be available to them?

Chapter 33 - Campbell (2003) confirms the equity premium puzzle in an international context based on the Consumption-CAPM and cross-country evidence on implausibly large coefficients of relative risk aversion. In this paper we adopt a spectral approach to re-estimate the values of risk aversion over the frequency domain for six Latin American emerging markets. We complement our analysis with the traditional time series approach and confirm the results of existing literature of large coefficients of relative risk aversion. Our frequency domain findings, however, indicate that at lower frequencies risk aversion falls substantially across countries, thus yielding in many cases reasonable values of the implied coefficient of risk aversion.

Chapter 34 - *Background*: The *Making Choices for Childbirth* study was undertaken to examine the effectiveness of a decision-aid to support women making choices for birth after caesarean section.

*Methods*: A prospective randomized controlled trial (RCT), involving 227 women recruited at 18 weeks gestation, measured levels of decisional conflict, knowledge and preferences for mode of birth at 28 weeks gestation and 36 weeks gestation. The decision-aid was administered to women in the intervention group after completion of the 28 week survey. A follow-up survey at 6-8 weeks postpartum focused on birth outcomes, health state and factors relating to preparation for decision-making, satisfaction with choice, decision support and overall birth experience.

*Results*: Of the 227 women, 193 (85%) completed surveys at both 28 and 36 weeks, with 169 (74%) returned at 6-8 weeks postpartum. Results for the prenatal components of the RCT are reported in Shorten et al (2005). The chapter will explore the relationship between various decision factors such as knowledge, decisional conflict and perceived preparation for decision-making. Analysis of postpartum data will include elements related to women's views on the decision support they received during pregnancy. The extent to which the decision-aid improved the likelihood of women feeling informed and being prepared to participate in decision-making will be discussed.

*Conclusions*: Evidence from the RCT suggests that the decision-aid is potentially valuable in assisting women who have experienced previous caesarean section to consider the risks and benefits of their options for birth after caesarean. However, this level of preparation for 'informed' decision making did not appear to translate into helping women discuss their options with their healthcare provider and others. In order to facilitate a process of shared decision making, stand alone products may need to be enhanced by strategies to better prepare practitioners to share decisions with consumers.

Chapter 35 - The acceptance of risks associated with new technologies is a key issue that is likely to limit the extent of innovation in a 'risk society'. However, given the limited comprehensible information available to the public of new technologies, it is likely that risk information provision will have a heterogeneous effect on public perceptions. In order to examine this issue, we empirically examine the determinants of risk perceptions, benefit perceptions and risks acceptance of new technology developments in Spain. Our findings indicate that risk and benefits perceptions are not independent but affected by common information sources. Furthermore, by taking into account this effect individual's knowledge of science heterogeneously increases both risks and benefits perceptions.

Chapter 36 - Preparedness is one component of effective hurricane risk mitigation. Emergency managers attempt to persuade residents of coastal zones to watch weather reports closely, develop and rehearse evacuation plans, store necessary supplies, and prepare property when storms threaten. Many coastal zone residents do not follow these recommended guidelines, even when recent experience suggests that the benefits of preparation exceed the costs. Understanding the factors that encourage these behaviors is an important step toward motivating coastal zone residents to become better prepared and minimize the losses from hurricanes. Toward this end we explore the relationship between wind and flood risk Preparedness is one component of effective hurricane risk mitigation. Emergency managers attempt to persuade residents of coastal zones to watch weather reports closely, develop and rehearse evacuation plans, store necessary supplies, and prepare property when storms threaten. Many coastal zone residents do not follow these recommended guidelines, even when recent experience suggests that the benefits of preparation exceed the costs. Understanding the factors that encourage these behaviors is an important step toward motivating coastal zone residents to become better prepared and minimize the losses from hurricanes. Toward this end we explore the relationship between wind and flood risk.

Chapter 37 - Over the past thirty years, researchers have developed a theoretical framework concerning risk and the protective mechanisms chosen by individuals (e.g. insurance, seat belts, storm shelters) against disasters. In most of these models, buying insurance is studied as the protective mechanism and monetary loss is taken as the damage from the hazard.

Chapter 38 - Two studies were conducted to: (1) identify the motives underlying dangerous driving among young males, and (2) evaluate the hypothesized structural relations (both direct and indirect) between the personality construct of sensation seeking, perception of danger, and the identified motives in representing the way risky driving decisions are made. In study 1, exploratory factor analysis (N = 200) yielded a three-factor structure representing three major motives for risky driving – *driving fast/risk taking, confidence in one's driving skills, disrespect for traffic laws.* Confirmatory factor analysis (N = 264) confirmed and further clarified this factor structure in representing the motives underlying young males' driving behavior. In study 2, path analysis (N = 384) provided overall support

for the 'fit' of the hypothesized model for dangerous driving. The implications of the findings with regard to the development of effective intervention strategies for dangerous driving among young males are discussed.

Chapter 39 - Mountaineers can be viewed as "edgeworkers" who carefully manage risks in a voluntary activity that has the potential for serious injury or death. We maintain that the act of managing risks contributes to a sense of "flow" or transcendence that is a major psychological motivation for mountaineers. We also maintain that decisions associated with engaging in mountaineering as an activity, choices of particular types of mountaineering, locations, specific mountains, etc., and decisions associated with the act of mountaineering itself are conducted, implicitly or explicitly, within a rational multi-objective risk management framework. There are numerous risk-risk, risk-benefit, and benefit-cost tradeoffs throughout this risk management hierarchy. Furthermore, the risk management process is dynamic on a number of different levels. This rational process must be conducted within a context of uncertainty and fear, balanced with the sense of flow, and survival is sometimes at stake. Based on a review of the peer-reviewed, mountaineering association, and popular mountaineering literature, interviews with mountaineers and professional guides, and personal experience, we were unable to find any explicit exploration of dynamic hierarchical multi-objective risk management in voluntary, recreational risky activities. The psychology of decision-making in this context is also largely unexplored. Thus, we have developed a qualitative framework, based on Hammond, Raiffa, and Keeney's PrOACT framework, that helps elucidate and inform the decision-making processes associated with mountaineering. This approach may help mountaineers and society understand the psychology and tradeoffs associated with this activity, and may be useful for risk management of other types of risky recreational activities.

Chapter 40 - Results from many studies suggest that people violate the principles of rational choice in both the domain of gain and that of loss. People usually treat probabilities non-linearly by overweighting low and underweighting moderate and large probabilities. The violations of rational choice in human decision-making were disregarded by the normative point of view until Allais and Kahneman and developed a descriptive theoretical approach. In this chapter we investigate what might affect decision-makers' preferences with respect to described real-world protective prospects. People's precautionary ('protective') decision-making in the face of risk implies that they may judge and weight the probability of risky events in characteristic ways that deviate from both normative EUT and psychological descriptive theory of decision-making tested with abstract gambles. The following theoretical frameworks contribute to an explanation of protective decision-making: (a) experience-based decision-making models - past and immediate experience affect decision-makers' preferences - and (b) accessibility of information - not all available observations of risks are equally accessible in memory.

Chapter 41 - The aim of this paper is to draw attention to what is arguably a very general and pervasive feature of human cognition that may have important implications for our understanding of human decision making and also for some aspects of economics. The major claim, defended here, is that when people judge the attributes of choice options (like utilities, payoffs, and probabilities), they are not able to represent the absolute magnitudes of these attributes; instead, they represent magnitudes ordinally---in relation to other magnitudes that they can sample from memory or from the current environment. Also, when people represent a magnitude, they can only do so on the basis of whether it is larger or smaller than other sampled magnitudes. Such sampling of knowledge from memory and transferring it to the current situation produces certain biases in judgment because stimuli are judged only relative to each other and therefore utility of an option is dependent on the other options that can be retrieved from memory. As a consequence, there may be no ability to represent cardinal scales, for any magnitude and judgments involving such magnitudes are determined by the context. The core evidence for this claim comes from recent research in psychophysics on the perception of the intensity of basic psychophysical magnitudes such as the brightness of a light or the loudness of a sound, and also from research on the effects of context on decision making under risk and uncertainty.

Chapter 42 - Today more than ever, people try to anticipate financial needs and to plan wisely for a lifetime of financial security. Information about financial options is plentiful, and financing for health and long-term care (LTC) is no exception. With all of the information and advice that is available, under what circumstances would a person decide that his/her decision was no longer the best option?

We address this question by looking at the market for LTC insurance, and estimate logistic regressions to model consumer decisions to drop or renew an existing LTC insurance policy. We explore events that occurred after the policy was last purchased and before the current policy was dropped or renewed. The price and benefit design of each policy is not directly observable so several proxy measures of the price of a policy are explored.

Data is obtained from the publicly available Health and Retirement Survey (HRS). Data from 2002 is used to identify those who have a LTC policy and to establish baseline financial circumstances. Data from 2004 is used to determine whether the policy was renewed, and to identify potentially influential events that occurred since 2002.

The study sample includes 1,375 individuals who reported an existing, private LTC insurance policy in 2002, and were therefore eligible to renew the existing policy before 2004. Proxy prices were calculated and assigned using publicly available price schedules.

Preliminary findings suggest that price was an influential factor in the decision to drop an existing policy, even though the price of the policy did not increase as a result of age. Those with newer policies were less likely to allow a policy to lapse. Those with low levels of assets (less than \$200,000) were more likely to allow a policy to lapse, as were those with more than \$1.5 million in assets.

Our results suggest that financial considerations are important, and a thorough review of an individual's financial circumstances may be effective in enabling people to make a lasting choice when they decide how to plan for LTC.

Chapter 43 - This chapter demonstrates how biases produced by verbal comparisons (Choplin & Hummel, 2002) might produce a variety of phenomena in the psychology of judgment and decision making. The biases produced by verbal comparisons cause people to overestimate the importance of small differences and underestimate the importance of large differences. Simulations will demonstrate that overestimating the importance of small differences from a reference point (default value, status quo, etc.) would produce s-shaped evaluation functions. Simulations will also demonstrate that overestimating the importance of small differences and underestimating the importance of large differences from a status quo, etc.) would produce s-shaped evaluation functions. Simulations will also demonstrate that overestimating the importance of small differences and underestimating the importance of large differences from other contextual stimuli would produce distribution-density effects (often called frequency effects). Because large differences are underestimated, when the variance in people's unbiased estimates is large as in anchoring effects, these biases will often look like assimilation effects. Biases produced by

verbalized social comparisons would also overestimate the importance of small differences and underestimate the importance of large differences so that moderate downward comparisons will produce higher evaluations of the self than will extreme downward comparisons and moderate upward comparisons will produce lower evaluations of the self than will extreme upward comparisons. Comparison strategies might also help explain decoy effects such as asymmetric dominance effects and phantom decoy effects. Testable and falsifiable assumptions of this model are described thereby laying a foundation for future empirical research.

Chapter 44 - Research has shown that a relation exists between the level of learning in the schools and universities and a country's strength. A relation also exists between education and the level and quality of life. Education today is a significant factor for ensuring society's normal existence, development and prosperity. However, major cities can afford the student the opportunity to acquire knowledge more than cities found in the periphery. A gap therefore exists between the level of learning in the major cities and the level of learning in the peripheral settlements. Students with high learning abilities who live in the cities can participate in university courses and other learning centers, whereas students with learning abilities who live in the periphery do not have a framework which can afford them knowledge in accordance with their talents and abilities.

Chapter 45 - Decision-making under uncertainty is to be expected in natural environments. The greatest source of uncertainty comes from the passage of time, because time—and the environmental variability it allows to proceed—discounts the reliability of information on which decisions are based. Information is most reliable if it can be acted on immediately, but as time passes, an average of past values of the alternatives is the best estimate of current value, since this subjective process matches the objective tendency of biological variables to regress to their means. The most optimal strategy, therefore, would be to flexibly shift from tracking the most recent outcomes to averaging across them. A model, the temporal weighting rule (TWR), accomplishes this transition. The output of TWR is a dynamic average whose rate matches the rate of environmental change. We review empirical studies showing the wide range of species that make dynamic foraging decisions consistent with TWR, the special predictions the model makes and their accuracy, its ecological relevance, and the memory mechanisms it appears to rely on. We conclude that this quantitative model and its accompanying decision rule, or something very similar to it, solve the one of the commonest problems animals face in their variable environments. TWR minimizes decision error made under uncertainty.

Chapter 46 - The Monty Hall Dilemma (MHD) is a notorious brain-teaser where people have to decide whether switching to another option in a game is advantageous. Most adults erroneously believe that chances of winning remain equal or that they should stick to their original choice. The present study tested the impact of cognitive development on MHD reasoning to examine possible differences in the nature of the erroneous intuitions. Twelve to seventeen year old high school students were presented the MHD and selected one of three responses (switch, stick, or chances equal). Results showed that whereas maturation decreased adherence to the erroneous "stick with your first pick" belief, the "chances are equal" belief became more dominant with increasing age. Consistent with predictions, children who selected the latter response also scored better on a syllogistic reasoning task. Results further showed that twelve year old eighth graders selected the correct switching response more frequently than senior high school students. Implications for popular reasoning and decision making theories are discussed.

Chapter 47 - Self-evaluations of performance are important in theory and practice. In contexts with multiple persons performing the same task, the evaluation of one's own performance is expected to be a process involving judgments about the performance of others, and comparisons between one's own and others' performance. We conducted a longitudinal study tracking 79 participants' evaluations of their own and others' performance on five repetitions of a task over a four-month period. Three temporal factors that Radhakrishnan, Arrow, and Sniezek (1996) identified as influences on self evaluations of performance were examined: Temporal Perspective, Time Horizon, and Experience. In this chapter, the authors investigated in more detail, the role of these factors, on judgments evaluations at multiple time-points before and after each task performance event. Results show that in general, evaluations of own and others' performance as well as on social comparisons. Participants made evaluations at multiple time-points before and after each task performance event. Results show that in general, evaluations of own and others' performance and social comparisons both had a positively leniency bias. This bias in self evaluations and social comparisons decreased when estimates were made (a) after performance than before; (b) closer to the performance event than farther away from it; and (c) with increasing experience. However, evaluations of only one's own performance were more variable with changes in the temporal factors. Further, the increase in bias with longer time horizons was reduced considerably with increasing experience. Changes in inter- and intra-individual validity followed those for bias. Interestingly, changes in solo evaluations over time were similar to those for social comparisons.

Chapter 48 - While remaining within the traditional micro-economic framework of rational utility maximization, the authors enrich the standard and random parameters logit choice models with perceptions data. From the estimated models they derive a value of time and also make a tentative attempt to derive a value of safety. Because they estimate the values simultaneously, they are able to explore whether values estimated in conjunction differ from values estimated in isolation. Survey data is used to measure the individual's perceptions of five modal attributes (time, cost, safety/risk, environmental friendliness and flexibility) and show how these perceptions affect the modal choice for work trips. The respondents' perceptions are elicited by a novel approach in which the names of two modes (car and bus) are used as attribute levels instead of objective levels. A difference between the atuthors' survey and traditional ones is that they do not attempt to educate the respondents about, for example, the risks of travelling. Instead, they *record* the respondent's perceptions about the risk and the other the modal attributes.

Chapter 49 - This chapter examines the impact of board composition and board size on three firm risk taking variables; strategy risk, stock returns risk, and income instability risk. There is recognition in the literature that corporate governance processes need to encompass mechanisms for motivating board and managerial behavior towards enhancing firm risk taking. Agency theory and regulatory recommendations advocate for an increasingly greater roles for outsiders on the board of directors. The evidence documented here indicates a positive relationship between majority independent board composition and firm risk taking. The literature also suggests board size affects firm activities independent of other board attributes and that there are biases against risk taking as board size grows. We do not, however, find any evidence to support the proposition that large board size influences firm risk taking. Thus, whereas the recommendation of increasing independent members on board of directors by regulatory bodies as well as the Cadbury and Hampel facilitate wealth creation by firms, the total representation on the board of directors does not exert any measurable influence on firms' wealth creation.

Chapter 50 - Human and non-human animals essentially face the same problem: how to find the best long-term option within an environment that contains uncertainty for relevant items. Against this background the Iowa Gambling Task is a biologically relevant task to study such decision-making processes. The authors have recently developed an animal analogue of this task in rodents. An interesting cross-species finding in this task is that performance differences exist between males and females: while males tend to focus exclusively on long-term goals, females tend to balance short- and long-term interests, in other words males shift from exploration to exploitation, while females remain exploratory. In this chapter the authors try to answer the question what may underlie these differences between males and females, focussing thereby on humans. First, they discuss a neurobehavioural model for the Iowa Gambling Task. Subsequently, they look at the contribution of the menstrual cycle using both data from the literature and an experiment that they conducted. The authors conclude that the menstrual cycle is not a decisive factor for these differences to occur. Finally, they discuss the possibility that differences in choice behaviour may be due to differences in the general dynamics of neurotransmitter systems. Based on recent experiments they conclude that differences in brain serotonergic and dopaminergic activity may contribute to the observed behavioural differences.

**Expert** Commentary A

## NEUROECONOMICS AND DECISION MAKING: A COMMON TASK FOR PSYCHOLOGISTS, ECONOMISTS AND NEUROSCIENTISTS\*

## José M. Martínez-Selva and Juan P. Sánchez-Navarro School of Psychology, University of Murcia, Spain

In recent years there has been a great amount of research on the study of the neural mechanisms of economic decisions, a scientific field currently known as Neuroeconomics (Camerer, 2003). This has been promoted by the development and use of imaging techniques, particularly functional magnetic resonance imaging (fMRI), for the exploration of brain activity. Neuroeconomics is an interdisciplinary field, as is decision making, in which the efforts of psychologists and economists join up with those of the neuroscientists, resulting in the participation of research specialists of different scientific fields in a given research.

An initial hurdle arises from the use of different concepts and approaches by psychologists, neuroscientists and economists, which complicates the reciprocal understanding of data, reviews, synthesis and theoretical proposals. A common theoretical background could facilitate the progress and knowledge of the state of the art in the different problems under investigation, such as the role of the emotion in decision making, the preference for immediate or delayed rewards, or the problem of cooperation in iterated or single trial games. Concerning the differences in language and approaches between psychologists, economists and neuroscientists, there have been important efforts to unify concepts such as "utility" and the different types of time discounting, among others. An example of this is the volume edited by Brocas and Carrillo (2003). Another important aspect would be to incorporate the study of individual differences, i.e. in age, cultural background or personality, in a field of research which is currently characterised by a generalistic approach to functional processes.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Medicine and Health Care*, edited by Thomas E. Lynch, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

We seek to draw the attention here to two relevant methodological issues that arise in Neuroeconomics and which, unless properly tackled, could hamper progress in this sphere and undermine some of the currently accepted findings. We refer specifically to the diversity of the experimental tasks and to the assessment of the results coming from functional neuroimaging techniques.

### **DIVERSITY OF TASKS**

Research on economic decision making employs a wide variety of tasks, from simple gambling tasks to serial and complex ones, such as the Iowa Gambling Task (Bechara et al., 1994), tasks coming from Game Theory, such as the Prisoner's Dilemma, or the Trust or the Ultimatum Games. Many of these tasks have nothing in common, and it is frequent that the same group of researchers switch from one task to another in each new experiment. Variety occurs also in the type of rewards, from taste stimuli, such as fruit juice, to cash, or symbolic rewards. It is also frequent to employ naturalistic situations, such as the simulation of stock market investments, which have more ecologic validity but whose results are difficult to generalize. Features, such as the greater or lesser uncertainty generated by the task, or the presence or absence of feedback coming from the outcomes of a single choice or the whole set of choices, likewise contribute to this situation.

The study of the relationships between brain and behaviour requires the design of standardised situations whose key variables and characteristics are perfectly known and replicable, situations usually known as "model-systems" (Thompson & Spencer, 1966). These model systems should meet certain features: they have to be well known, and the principles that govern them can be appropriately measured, manipulated precisely and with detail, and they should be closely related to the process under study and sensitive to the variables addressed. Meeting these requirements will accelerate progress in this sphere (Martínez-Selva et al., 2006).

In our opinion, the use of the parsimony principle is linked to the development and use of these standardised tasks. For that purpose it would be necessary for the ecological validity to give way to a more rigorous methodology that ensures an adequate control of the factors involved in the tasks (Martínez-Selva & Sánchez-Navarro, 2007).

#### **FUNCTIONAL NEUROIMAGING TECHNIQUES**

In recent years a large amount of data have been made available from the use of fMRI, which shows us the increased o decreased metabolic activity in different brain regions while the subject is performing decision making tasks. These data allow us to identify the structures and systems that may be participating in a significant way in the different processes under study.

Kringelbach and Rolls (2004) have appropriately underlined the limitation of fMRI and the need for the combined use of different methodological approaches in this type of studies. Functional neuroimaging provides important spatial information about the brain activity by showing the average metabolic activity in a given brain region. However fMRI is less sensitive in detecting the role of different neuronal groups and their interactions. This technique shows us structures that are related, for example, to a given cognitive process. However, due to the very nature of the signal obtained by fMRI, it can also show activation in brain regions that are not necessary for the task, i.e. structures that are active due to other physiological processes unrelated to the cognitive or emotional processing required by the task under study (Rorden & Karnath, 2004). Its limits may also lie in the relevance that small populations of neurons have to a differential response to different situations, processes or stimuli that could be overlooked with this technique. Although the technique has provided many useful data, the results should be considered together with those coming from other methodologies, both from human (electrophysiology, magnetoencephalography, patients with brain injury) and from nonhuman subjects (neurochemistry, electrophysiological single cell recording).

The correlational nature of functional neuroimaging also imposes limits on the interpretation of its results. Ideally, the data should be combined with those coming from purely experimental approaches that ensure a due manipulation of the variables under study, and which allow us to show as clearly as possible the causal role of the different brain structures in the process under study.

As happens in other fields of research, the consideration of methodological questions goes beyond applied, specific issues, and leads us to ask what we understand by a valid scientific explanation. In our case we refer to the explanation of the role of the biological mechanisms (brain structures or systems, neurons, neurotransmitters, drugs) in specific psychological processes (choices, anticipation of reward, prediction errors, preference for short or long term rewards, intensity and valence of the emotions), and what the nature, quantitative or qualitative (correlation, cause, interaction between variables, simple effects) of this relationship is.

We think that the methodological and interdisciplinary approach put forward here will contribute to improve the quality of the data and, in general, of the research in this field.

# REFERENCES

- Bechara, A., Damasio, A. R., Damasio, H. & Anderson, S. W. (1994) Insensitivity to future consequences following damage to the human prefrontal cortex. *Cognition*, 50, 7-15.
- Brocas, I & Carrillo, J. D. (Eds.) (2003) *The psychology of economic decisions. Vol. I, Rationality and well-being.* Oxford: Oxford University Press.
- Camerer, C.F. (2003) Strategizing in the brain. Science, 300, 1673-1675.
- Kringelbach, M. & Rolls, E.T. (2004) The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, 72, 341-372.
- Martínez Selva, J. M. & Sánchez Navarro, J. P. (2007) Decision making and the emotional brain. En Thomas E. Lynch (Ed.) *Psychology of Decision Making in Health*. Hauppauge, NY: Nova Science Publishers. (In press).
- Martínez Selva, J. M., Sánchez Navarro, J. P., Bechara, A. & Román, F. (2006) Mecanismos cerebrales de la toma de decisiones. Réplica. [The brain mechanisms involved in making decisions. Reply] *Revista de Neurología*, 43, 190.

- Rorden, C. & Karnath, H.-O. (2004) Using human brain lesions to infer function: a relic from a past era in the fMRI age? *Nature Reviews Neuroscience*, 5, 813-819.
- Thompson, R. F. & Spencer, W. A. (1966) Habituation: A model phenomenon for the study of neural substrates of behavior. *Psychological Review*, 73, 16-43.

Expert Commentary B

# **END OF LIFE DECISION MAKING: AUTONOMY AND CAPABILITY<sup>\*</sup>**

# Guido Miccinesi<sup>1</sup> and Claudia Borreani<sup>2</sup>

<sup>1</sup>Epidemiology Unit, Centre for the Study and Prevention of Cancer, Florence-Italy <sup>2</sup>Psychology Unit, National Cancer Institute, Milan-Italy

End-of-life decision making has various ethical and psychological implications. One of the main aspects is the assumption that it is the patient who must take all decisions regarding him/herself. Clinical practice does not always fit this ideal, as each case must be taken in its legal and cultural context, with much depending also on the request of the individual. In Europe a huge variability in types and proportion of end of life medical decisions has been shown, ranging from 23% of all deaths in Italy up to 51% in Switzerland. In that population-based study (Eureld) about one third of all deaths were sudden and totally unexpected in each country. Among the remaining dying patients some decisions, that could have hastened death in the attending doctor's opinion, were made in at least one out of two cases in all countries except Italy. The highest difference between countries concerned the non-treatment decisions. The lowest proportion of that type of decision was in Italy (4% of all deaths) where the percentage of individual DNR decisions was also the lowest (16% of all non-sudden deaths), and the attitude to communicate about incurability was 4 to 8 times lower than in the other participating European countries as well as in Australia [1-4].

Therefore end-of-life decision making is surely different depending on the specific legal and cultural context. Other conditions can make decision making vary from case to case and far from modern western society's ideal of the patient who decides by him/herself at the end of life. Even today patients' awareness of their poor prognosis can be lower than expected: in a recent national study among dying cancer patients in Italy only 13% explicitly knew about their terminal condition [5]. The patient's process of becoming aware of his/her incurability is mirrored by the doctors' task of providing information regarding the development of the

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

illness, and neither can be easily reduced to a simple problem of breaking bad news. The time frame in which these processes usually take place is generally not compatible with the necessary time required to follow ordinary protocol of communication, which would allow the patient to come to terms with his/her own situation. Furthermore, it has been shown how much a patient's power for decision making can be unstable at the end of life, being connected to an instability of mood and more generally to the difficult task of dealing with desperation and search for meaning at the end of life [6,7]. The duration of staying in palliative care is becoming shorter, therefore it becomes more difficult to discuss openly endof-life issues with the patient in that setting, especially if he/she has not been able to deal with these issues before. To make best use of a patient's remaining life, it is important that the end-of-life decision making process is based on a patient's history and available resources, among which his/her family should be considered first. Only rarely does a dramatic change in the patient's will, attitude towards death and awareness occur; more often the clinicians or the psychologists working with a fragile person choose a conservative approach towards the psychological equilibrium already reached by the patient in the previous phases of the illness. The aim of psychological support in this phase is to facilitate the development of an emotional context able to follow the non linear evolution of the dying patient and the progressive separations involved.

It has been shown that only 2% of patients were reportedly distressed at being interviewed about their own death, whereas 20% of them considered the experience to be useful [8]. We could say that a major obstacle towards full participation of patients in the end-of life decision making is the behaviour and attitude of the majority of medical doctors, who are not emotionally ready and not sufficiently trained to deal with the relational and emotional aspects of the therapeutic relationship, not only in the case of dying patients. Some authors, studying the components of the communication between patient and physician at the end of life, highlighted also the importance of 'willingness to communicate' with the patient about dying and death [9]. The availability of the medical doctors to listen to the patients does not resolve all the communication problems at the end of life. It is also important that they develop the appropriate cognitive and emotional capacities to accompany patients on this path where often the common understanding of relations and roles is inadequate.

If this conceptual and practical framework is correct, there needs to be a shift away from a model of good death which includes complete autonomy of the patient towards the idea of enhancing his/her capability. Following A. Sen, capability means to be able to perform specific functions, to make certain patterns of being and doing real; it does not mean being normatively obliged towards those patterns of being and doing. Therefore tools to elicit patients end of life preferences are needed, to allow the patient to take the specific route necessary towards awareness of his/her situation. It is important to make people able to express some preferences about the dying phase, even to prefer to trust medical doctors or informal caregivers for decision-making at the end of life, or to leave some decisions to doctors or caregivers while keeping some other decisions for themselves. How do we reach such a deep understanding of each dving patient? Some authors put in evidence the importance to encourage questions from dying patients and being sensitive to when patients are ready to talk about death [9]. Some stated that time of the clinical encounter with a terminally ill patient might also be used more efficiently if providers have a prior list of themes to touch upon [10]. Others made us aware that such preferences are not something that exists as a pure type, a clearly definable entity, that can be straightforwardly categorised and

counted. Neither are they necessarily active choices, meaning that they can be the refusal of interventions. Rather, they take the form of stronger or weaker leaning in one direction, qualified by speculations about how things might change with events [11].

We recently participated in the developing of a relational instrument (End of Life Preferences Interview) that facilitates and stimulates communication on these aspects. Selected topics for the structured interview were organised in six macro-areas: information/communication, symptom control and treatment, relational needs, preparation for death, spiritual needs, religious practices. The instrument was divided in two parts, each one introduced by a key question, which was derived according to the clinical experience of the authors and aimed at allowing the patient to decide whether and when to stop exploring such sensitive matters. Discussing preferences means to begin to create the psychological conditions in the patient in order to, if he wishes, exert his self-determination in a delicate phase of his own life. The physician should be ready to relate to these aspects with respect to the times and defences of the patient. The instrument is still in a testing phase, but it is interesting as a proof of the arguments sustained throughout this commentary, to note how difficult it was for the physicians to apply it and to integrate it in their usual way of dealing with these issues. Although they considered the instrument highly useful, comprehensible and exhaustive, in the majority of eligible cases they did not find the conditions and fundamental requisites so that physicians and patients could confront the contents of the ELPI. The criticisms that emerged mainly concerned the fear of triggering intense emotional reactions in the patients but also the difficulty in finding the right moment and time necessary to confront the issue. A possible role in the difficulty in application could also be related to our decision to ask that administration of the instrument be carried out by the physicians. Nurses and psychologists may have had better compliance since they are more used to communicating with patients on these issues. Our decision was motivated by the fact that talking about these issues with the physician assumes for the patient a stronger significance but also by the fact that physicians, generally little inclined and little prepared to confront emotional and relational aspects of the therapy [12], cannot delegate completely the administration of these aspects to other professions. Finally, it is noteworthy that our research was conducted in palliative settings, the most sensitive context in the theme of quality of dving. This indicates the need for a change of attitude in the medical field with regard to the end of life. Debates, education, training in end-of-life care can contribute to this change, as well as the clear indication of how to look at these issues. Our proposal is to choose a capability approach over an autonomist one. This could guarantee that the passage to a greater consideration of the self-determination of the patient at the end of life can come with respect for the individuals' cultural and psychological needs, leaving the patient with the possibility to decide whether, when and which among these themes to confront with his own therapy team.

### REFERENCES

 Cartwright C, Onwuteaka-Philipsen B, Williams G, Faisst K, Mortier F, Nilstun T, Norup M, van der Heide A, Miccinesi G on behalf of the EURELD consortium. Physician discussion with terminally ill patients: a cross-national comparison. *Palliat Med*, 2007, in press.

- [2] Bosshard G., Nilstun T, Bilsen J, Norup M, Miccinesi G, van Delden J, Faisst K, van der Heide A. Forgoing Treatment at the End of Life in Six European Countries. *Arch Intern Med*, 2005; 165 : 401-407.
- [3] Van Delden JJ, Lofmark R, Deliens L, Bosshard G, Norup M, Cecioni R, van der Heide A; on behalf of the EURELD consortium. Do-not-resuscitate decisions in six European countries. *Crit Care Med*, 2006;34(6):1686-9.
- [4] Van der Heide A, Deliens L, Faisst K, Nilstun T, Norup M, Paci E, van der Wal G, van der Maas PJ; EURELD consortium. End of life decision-making in six European countries: descriptive study. *Lancet*, 2003, 362: 345-350.
- [5] Costantini M, Morasso G, Montella M, Borgia P, Cecioni R, Beccaro M, Sguazzotti E, Bruzzi P, on behalf of the 'ISDOC Study Group'. Diagnosis and prognosis disclosure among cancer patients. Results from an Italian mortality follow-back survey. *Ann Oncol*, 2006, 17:853-859.
- [6] Chochinov HM, Hack T, Hassard T, Kristjanson LJ, McClement S, Harlos M. Dignity in the terminally ill: a cross-sectional, cohort study. *Lancet*, 2002;360(9350):2026-30.
- [7] Breitbart W, Gibson C, Poppito SR, Berg A. Psychotherapeutic interventions at the end of life: a focus on meaning and spirituality. *Can J Psychiatry*, 2004; 49(6):366-72.
- [8] Emanuel EJ, Fairclough DL, Wolfe P, Emanuel LL. Talking with terminally ill patients and their caregivers about death, dying, and bereavement: is it stressful? Is it helpful? *Arch Intern Med*, 2004;164(18):1999-2004.
- [9] Wenrich MD, Curtis JR, Shannon SE, Carline JD, Ambrozy DM, Ramsey PG. Communicating with dying patients within the spectrum of medical care from terminal diagnosis to death. *Arch Intern Med*, 2001;161(6):868-74.
- [10]Steinhauser KE, Christakis NA, Clipp EC et al. Preparing for the end of life: preferences of patients, families, physicians, and other care providers. *J Pain Symptom Manage*, 2001; 22:727-37.
- [11] Thomas C, Morris SM, Clark D. Place of death: preferences among cancer patients and their carers. Soc Sci Med, 2004; 58:2431-44.
- [12]Baider L, Wein S. Reality and fugues in physicians facing death: confrontation, coping, and adaptation at the bedside. *Crit Rev Oncol Hematol*, 2001; 40:97-103.

Expert Commentary C

# PSYCHOLOGICAL ASPECTS OF DECISION MAKING AS NEW RESEARCH TOPICS IN NATURAL RESOURCES MANAGEMENT<sup>\*</sup>

# T. Hujala<sup>1</sup>, T. Kainulainen<sup>2</sup> and P. Leskinen<sup>2</sup>

<sup>1</sup>Department of Forest Resource Management, P.O.Box 27, FI-00014 University of Helsinki, Finland <sup>2</sup>Finnish Forest Research Institute Metla, Joensuu Research Unit, P.O. Box 68, FI-80101 Joensuu, Finland

## ABSTRACT

Sustainable management of natural resources necessitates conscious and organised activity. When planning the use of a certain spatially explicit management unit, competing goals and interests call for multi-objective decision support. The environmental decision making also involves uncertainty concerning available information and future predictions. Traditionally, a typical procedure has included the use of preference elicitation tools, statistical models, and optimisation methods, to find an acceptable solution for the decision problem at hand.

When improving the decision support portfolio to work more efficiently, there is a need to consider the overall structure of decision problems. To demonstrate that, we have conceptualised the variety of forestry decision problems in three-dimensional decision-problem space, which comprises geographical, temporal, and social dimensions. The crucial issue in the problem structuring and, thus, in the selection of decision support method, is scope awareness: recognising the fundamentals of the decision problem at hand. With the aid of that, the unwanted decision-model-based or support-methodology-based psychological biases can be avoided or smoothened.

By the results of the research proposed in this chapter, both the efficiency of single phases of decision making and the quality of the whole decision process would be

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Economics, Business and Finance*, edited by Klaus P. Hofmann published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

improved. The results would also enhance learning about the characteristics of decision problems and decision makers' preference structures.

**Keywords:** adaptive decision analysis, decision support systems, measurement scales, mental patterns, preference elicitation, psychological biases, uncertainty.

### BACKGROUND

In natural resources management, the intertwining of global and local, as well as public and private perspectives evokes the need for more sophisticated decision support methods. These methods would contribute e.g. to collaborative planning (e.g. Leskinen 2006), conflict management (e.g. Niemelä et al. 2005, Losa and Belton 2006), biodiversity maintenance (e.g. Kurttila et al. 2006), and legitimacy of environmental and forest policies (e.g. Rantala 2004). All of these aspects imply that a decision-makers' subjective views should be taken into account thoroughly, with the aid of cognitive and social psychology, as well as game theoretical aspects (e.g. Bender and Martin 2003). By the term decision-maker we do not refer merely to national, regional or local politicians or business managers but also to landowners, representatives of interest groups, and ordinary citizens, who participate planning or comparable decision making processes in some way or another.

The properties of decision problems as well as the understanding, motives, and commitment of decision-makers vary largely in different cases. Therefore, a traditional decision support system with fixed parameters and properties is susceptible to fail in fulfilling its expectations. In order to respond to this challenge we have enhanced the grounds for *adaptive decision analysis*, which adapts flexibly to the requirements of the output data and availability of the input data, taking into account the social system of decision considerations. With the adaptive approach, it is possible to select and apply appropriate decision models for different types of decision-problems and decision-makers.

A useful decision support system (Sprague 1993) helps the decision-maker to make better decisions or to find acceptable solutions more comfortably. If the system is transparent and easy to understand, the usage of such system results in learning about the problem (e.g. Buchanan 1994) and in improved commitment to the decisions. When the decision problem is straightforward and has a simple structure, our previous research (e.g. Hujala et al. 2007) shows that merely an illustration and discussion about the decision alternatives may be sufficient decision support to the decision-maker. Otherwise, e.g. when the decision-problem is about a large geographical area with a long planning period, cardinal, i.e. interval (e.g. Keeney and Raiffa 1993, Kainulainen et al. 2007) or ratio scale preference, elicitation techniques (e.g. Saaty 1980) are advisable.

The advantage of the cardinal techniques is that they offer a more sophisticated view on the problem, including the analysis of trade-offs between the conflicting decision criteria (c.f. Leskinen 2001). In addition, cardinal methods enable versatile possibilities for uncertainty analysis (e.g. Alho et al. 2001), which is an essential feature in natural resource management with complex structures. However, when reaching efficiency and acceptability of the solutions, profound computational methods will require also qualitative conversational support, especially in structuring the decision problems and assuring the decision-makers' understanding of the results.

## **DISCRETION REQUIRED IN USING ELICITED PREFERENCES**

The mental aspects of preference elicitation can have significant impacts in decision making. Psychological biases in decision making are well-documented in the behavioral decision making literature (c.f. Hammond et al. 1998, Pöyhönen et al. 2001). Also, the performance of multi-criteria decision making (MCDM) methodology has been examined from the behavioral perspective. As for the results, *anchoring effects* have turned out to be a common type of psychological bias in interactive MCDM methods. For example, Buchanan and Corner (1997) identified that the starting point of decision process affects the weight assessment in the interactive Zionts and Wallenius (1983) method. In cardinal MCDM methods, a specific alternative or a certain interval is used as a reference for the preference elicitation process. One interesting issue for future research would be to examine whether the changing of the reference alternative impacts the outcome of the analysis. If such effect can be observed and verified in experiments and sensitivity analyses, some anchoring-based scaling recommendations for preference elicitation could be derived.

Another important topic for future research would be to conduct a systematic preference elicitation test, in which both interval and ratio scale models would be tested through direct weighting and pairwise comparisons techniques. It would then be interesting to examine whether the different methods lead to systematic differences in the estimated preferences (c.f. Belton 1986, Hujala and Leskinen 2006) and uncertainty measures. In addition, the rationale of individual decision-makers' thinking could be explored through qualitative interviews and direct observations. Implications of the empirical findings could then be drawn back to the design of preference enquiries.

# **INTRODUCING DECISION-PROBLEM SPACE**

In order to respond to the requirements of taking into account decision-makers' perspectives in natural resource management, we hereby introduce a concept of *decision-problem space* (Figure 1). The space illustrates, in the context of forestry decision making, the potential diversity of decision-problems. In geographical dimension, one may focus on a single tree, forest compartment, an estate, region etc. The time horizon may be only one season, a couple of years, a decade or two, sometimes a century, or even more. The people who participate in the process form the social dimension of decision making. We measure the social dimension by the depth of participation and by adapting Arnstein (1969).

The scales in different dimensions are dependent, i.e. on the applicability of a decision support method in one certain dimension that is simultaneously affected by the case-specific scales in the other two dimensions. Thus, we can not study decision aid needs in one dimension at a time, independently of the other two dimensions. On the other hand, because decision support systems are tools to be used by humans, we consider that it is especially critical to study the methodological problems related to the social dimension. In addition to comparing single points in the decision space, it might be beneficial to study decision scales hierarchically so that decisions made at one level are dependent upon decisions or information at other levels.

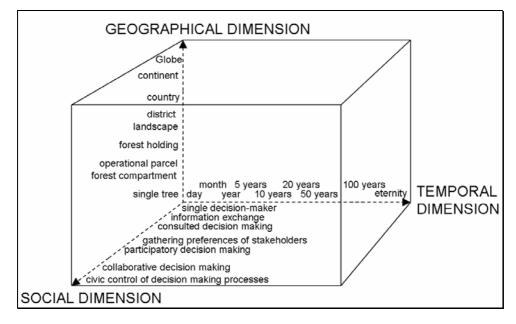


Figure 1. Illustration of three-dimensional decision-problem space (case forestry decisions).

The aim in hierarchical planning is to find formulations that preserve consistency between the decision levels. Production of information that supports finding solutions for problems arising from violated constraints, inefficient solutions, or infeasible problem formulations between or inside the levels of hierarchy are important (e.g. Davis and Liu 1991, Weintraub and Cholaky 1991, Davis and Martell 1993). The hierarchical analysis is important e.g. when projecting operational actions from strategic goals, implementing national policy in regional level, or compiling aggregate goals from the preferences of several individuals. The decision hierarchy may be both top-down- or bottom-up-oriented. It will be crucial that the available decision support methods both enable and support the hierarchical decision analysis.

# CONCLUSION

We emphasise that it is important to attain and distribute new knowledge of decisionmakers' perceptions of various decision support techniques and processes, especially the psychological biases that emerge within decision-making processes that need anticipation and smoothening. The research should open-mindedly take advantage of mixing quantitative and qualitative approaches. Developers and facilitators of decision support systems as well as decision-makers themselves need state-of-the-art recommendations for more convenient and transparent decision making. These guidelines would help decision-makers in natural resource management to meet the contemporary challenges such as multi-objectivity, interactivity, collaboration, and risk management. They would also improve learning possibilities in the context of decision making about different scales and facilitate the legitimacy of policies related to all levels and disciplines of natural resources management.

## REFERENCES

- Alho, J.M., Kolehmainen, O., Leskinen, P. (2001). Regression methods for pairwise comparisons data. In: Schmoldt, D.L., Kangas, J., Mendoza, G.A., Pesonen, M. (eds.). *The Analytic Hierarchy Process in natural resource and environmental decision making*. Kluwer Academic Publishers, Dortrecht. p. 235–251.
- Arnstein, S., (1969). A ladder of citizen participation. Journal of American Institute of Planners 35: 216–224.
- Belton, V. (1986). A comparison of the Analytic Hierarchy Process and a Simple Multiattribute Value Function. *European Journal of Operational Research* 26: 7–21.
- Bender, H.W., Martin, W.E. (2003). Modelling conflict using spatial voting games: an application to USDA Forest Service lands. *International J. of Global Environmental Issues* 3:149–167.
- Buchanan, J.T. (1994). An experimental evaluation of interactive MCDM methods and the decision making process. Journal of the Operational Research Society 45: 1050–1059.
- Buchanan, J.T. and Corner, J.L. (1997). The effects of anchoring in interactive MCDM solution methods. *Computers and Operational Research*, 24: 907–918.
- Davis, L.S., Liu, G., (1991). Integrated forest planning across multiple ownerships and decision makers. *Forest Science* 37: 200–226.
- Davis, R.G., Martell, D.L. (1993). A decision support system that links short-term silvicultural operating plans with long-term forest-level strategic plans. *Canadian Journal* of Forest Research 23: 1078–1095.
- Hammond, J. S., Keeney, R. L. and Raiffa, R. (1998). The hidden traps in decision making. *Harward Business Review*. Sept.-Oct.: 47–58.
- Hujala, T. and Leskinen, P. (2006). Preferences in interval and ratio scales An experiment with non-industrial private forest owners. Oral presentation at EUROXII European Conference on Operational Research 2-5 July, 2006. Reykjavik, Iceland. Abstract published in: Plà, L.M. (ed.) (2006). Proceedings of the Second Meeting of the EURO Working group on Operational Research (OR) in Agriculture and Forest Management. *Journal of Agricultural Science*, 144, 467–472.
- Hujala, T., Pykäläinen, J., Tikkanen, J. (2007). Decision-making among Finnish nonindustrial private forest owners – the role of professional opinion and desire to learn. To appear in Scandinavian Journal of Forest Research.
- Kainulainen, T., Leskinen P., Korhonen, P., Haara, A. (2007). A Statistical approach to assessing interval scale preferences in discrete choice problems (submitted).
- Keeney, R. L., Raiffa, H. (1993). Decisions with Multiple Objectives. *Preferences and Value Tradeoffs*. Cambridge University Press, Cambridge.
- Kurttila, M., Pykäläinen, J., Leskinen, P. (2006). Defining the forest landowner's utility loss compensative subsidy level for a biodiversity object. *European Journal of Forest Research* 125: 67–78.

- Leskinen, L.A. (2006). Purposes and challenges of collaborative forest planning. *Working Papers of the Finnish Forest Research Institute* 38: 66–69.
- Leskinen, P. (2001). Statistical methods for measuring preferences. University of Joensuu. *Publications in Social Sciences* 48. D.Sc. thesis. 111p.
- Losa, F.B., Belton, V. (2006). Combining MCDA and conflict analysis: an exploratory application of an integrated approach. *Journal of the Operational Research Society* 57: 510–525.
- Niemelä, J., Young, J., Alard, D., Askasibar, M., Henle, K., Johnson, R., Kurttila, M., Larsson, T.-B., Matouch, S., Nowicki, P., Paiva, R., Portoghesi, L., Smulders, R., Stevenson, A., Tartes, U., Watt, A. (2005). Identifying, managing and monitoring conflicts between forest biodiversity conservation and other human interests in Europe. *Forest Policy and Economics* 7: 877–890.
- Pöyhönen, M., Vrolijk, H. and Hämäläinen, R.P. (2001). Behavioral and procedural consequences of structural variation in value trees. *European Journal of Operational Research* 134: 216–227.
- Rantala, T. (2004). Conceptions of democracy of key informal interest groups in Finnish forest policy. In: Andersson, F., Birot, Y., Päivinen, R. (eds.). (2004). Towards the Sustainable Use of Europe's Forests – Forest Ecosystem and Landscape Research: *Scientific Challenges and Opportunities. EFI Proceedings* No. 49.
- Saaty, T.L. (1980). The Analytic Hierarchy Process, McGraw-Hill, New York, NY.
- Sprague, R.H. (1993). A framework for the development of decision support systems. In: R.H. Sprague, Jr. and H.J. Watson (eds.). (1993). *Decision Support Systems* (3<sup>rd</sup> ed.): putting theory into practice; pp.3–28. Prentice-Hall, NJ, USA.
- Weintraub, A., Cholaky A. (1991). A hierarchical approach to forest planning. *Forest Science* 37: 439–460.
- Zionts, S. and Wallenius, J. (1983). An interactive multiple objective linear programming method for a class of underlying nonlinear utility functions. *Management Science* 29 (5): 519–529.

Expert Commentary D

# NEW DIRECTIONS IN THE PSYCHOLOGY OF DECISION MAKING\*

# Liqing Zhang<sup> $\dagger$ </sup>

Peking University, Beijing, China

The psychology of decision making has been gaining more attention than before. It is growing rapidly. Three main perspectives of the psychology of decision making include the standpoint of cognitive psychology, the influence of social psychology and the viewpoint of neuropsychology. The present commentary reviews briefly these three major approaches to the psychology of decision making, points out their contributions and limitations, and suggests new directions.

The approach of cognitive psychology has produced a flourish of research and gained much attention from other fields, such as economics, political science, and management. From the viewpoint of cognitive psychology, decision making is considered an outcome of information processing. Early work addresses how people compare and weigh different dimensions of a choice set and then select the optimal (or best) outcome. Later research demonstrates that people may not be able to process all the relevant information related to a choice due to their limited cognitive resources. They may also be unwilling to process all the information related to a decision. For instance, people do not think and compare an extensive list of foods for breakfast, they may just pick a satisfactory option. Thus, people may stop their search for information when they encounter a satisfying option (Simon, 1957). Additionally, people may rely on heuristics and other rules of thumb in decision making. The application of heuristics reserves cognitive resources and adapts to its social environment in many situations (Gigerenzer & Goldstein, 1996). However, the application of heuristics also generates significant costs and produces decision errors (Tversky & Kahneman, 1974).

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: Liqing Zhang, Department of Psychology, Peking University, 5 Yi-He-Yuan Avenue, Beijing, P. R. China, 100871. Email: liqingzhang@pku.edu.cn.

Another important contribution of the cognitive approach is the prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 2000). Prospect theory includes three major propositions. First, people perceive loss and gain differently. The pleasure people derive from a certain amount of gain is much less than the pain people derive from the same amount of loss. Second, whether an outcome is perceived as a gain or a loss depends on how the outcome is framed based on a reference point. The reference point may be determined by one's wealth, income, allowance, past experiences, future expectation or social comparison. Third, people are usually risk averse in the domain of gain and are risk seeking in the domain of loss. The prospect theory has been applied to different fields of decision making, such as marketing and management. A recent development of this standpoint is to apply prospect theory into new real life decisions through field experiments. Another development is to address limitations of the prospect theory. Researchers identify situations in which people are risk seeking in the domains of gain. For instance, if they had to choose between a sure gain of a penny and a 0.01% probability of attaining \$100, most people may be willing to take the risk.

The major limitation of the perspective of cognitive psychology is that it neglects some important human motivations. Thus, practitioners often argue that these findings are different from the decision process in everyday life. Research on social psychology and decision making may have the potential to fill in the gap on motivations and choices.

The contribution of the approach of social psychology includes research on how fairness, emotion and self-control influence choices. First, people take fairness into consideration in their choices. Fairness puts a crucial constraint on profit maximization. For instance, restaurants usually do not raise their prices during popular hours, such as evenings on weekends because they worry that consumer may view such an opportunistic behavior to be unfair. Second, emotions influence choices. Research suggests that not only do emotions elicited by the choice situation impact judgment and decision making, but incidental emotions generated by sources irrelevant to the choice situation also influence the decision. Research has contributed to the field by examining the diverse effects of positive and negative emotions on choice. Recent research suggests that some emotions may have their specific effects. For instance, even though both anger and fear are negative emotions, anger leads to risk seeking, while fear produces risk aversion (Lerner & Keltner, 2001). A next step of the research on emotions and choices may address whether these specific emotions differ on systematic dimensions, such as arousal-calm, approach-avoidance, and examine how these different dimensions of emotions influences choices. Third, self-control influences an important aspect of decision making, intertemporal choices, which involve tradeoffs between costs and benefits at different times (Loewenstein, Read, & Baumeister, 2003). People often need to exert their self-control in stifling their immediate desire for short-term gain so as to achieve long-term benefits. For instance, dieters have to resist the temptation of tasty desserts in order to keep their body fit. People may override their desire for luxuries and save money for their children's education.

One major question that remains in the research on motivation and decision making is whether motivation influences decisions directly or motivation shapes behaviors through biased information processing. Encouraging progress has been made on how emotion shapes choices. Some research suggests that people ignore the probability of winning when they are under intense emotions (Loewenstein, Weber, Hsee, & Welch, 2001). The other research has focused on how emotion influences the perception of information (Schwarz, 1990). The theoretical model on how emotion or motivation shapes behavior deserves more attention (Baumeister, Vohs, DeWall, & Zhang 2007; Zajonc, 1980).

The study on neuropsychology and decision making is an emerging topic (Eshel, Nelson, Blair, Pine, & Ernst, 2007). The advantage of meuropsychology is that it provides an objective measure of mental processes. However, this approach is constrained by the fact that the structure and function of the human brain is still a mystery. This approach has contributed more to confirming available theories using a new method than to construct new theories.

The psychology of decision making includes three major approaches –cognitive, social, and neurological perspectives, each has contributed to the field and gains much attention in the scientific disciplines beyond psychology. The crucial question is what the next step of the psychology of decision making is. The progress of a general theory may be needed. In the past decades, research on psychology and decision making has benefited from inspiring theories, such as prospect theory. In the recent thirty years, research has progressed more in generating middle-level theories than constructing integrative theories. The impressive middle-level theories, elaborated experimental techniques, and abundant empirical findings have accumulated building blocks for new theories. The emergence of new integrative theories may contribute greatly to the further development of this field.

What perspectives have the potential to integrate the previous findings and provide general theories? First, a better understanding of the attention process may offer an integrative theory on the psychology of decision making. Attention is an operational phase of information processing. Recent research suggests that attention is related to the platform of all sorts of cognitive activities—working memory (Engle, Tuholski, Laughlin, & Conway, 1999). Since the resource of attention is limited, attention may explain application of heuristics in decision making. The limited resource of attention may also explain bounded rationality.

Second, an understanding of the impact of self-esteem on choices may facilitate the formation of an integrative theory on motivation and decision making. People are motivated to defend, maintain and enhance their self-esteem (Baumeister, 1988; Greenwald, 1980; Steele, 1988; Taylor & Brown, 1988). In social psychology, a large amount of research has documented the importance of self-esteem in shaping goals, emotion and cognition. Self-esteem may affect why people care about fairness, when and why people experience intense emotions, and how they choose options that are important to their identity. Recent research has started to explore how self-esteem influences individual choices and interpersonal negotiations (Larrick, 1992; Zhang & Baumeister, 2006; Zhang, 2004; Zhang, 2007). Future research should construct a theory on the role of self-esteem in decision making.

Researchers may find other viewpoints in constructing integrative theories. The science, psychology of decision making, may have a bright future if a few new breakthrough theories are proposed.

## **AUTHOR NOTE**

I want to thank Huiling Yu for her comments on an earlier draft of this paper.

### REFERENCES

- Baumeister, R. F. (1998). The self. In. G. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology(Vol.1, 4th Ed., pp.680-740)*; New York, NY: Oxford University Press.
- Baumeister, R. F., Vohs, K. D., DeWall, C. N., & Zhang, L. (2007). How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation? *Personality* and Social Psychology Review, 11, 167-203.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E. & Conway, A. R. A (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. Journal of Experimental Psychology: General, 128(3), 309-331
- Eshel, N., Nelson, E. E., Blair, R. J., Pine, D. S., & Ernst, M. (2007). Neural substrates of choice selection in adults and adolescents: Development of the ventrolateral prefrontal and anterior cingulate cortices. Neuropsychologia, 45, 1270-1279.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.
- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, *35*, 603-618.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, 47, 263-291.
- Kahneman, D., & Tversky, A. (2000). Choices, values, and frames. New York, NY, US Cambridge University Press.
- Larrick, R. P. (1993). Motivational factors in decision theories: The role of selfprotection. *Psychological Bulletin*, 113, 440-450.
- Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. Journal of Personality and Social Psychology, 81, 146-159.
- Loewenstein, G., Read, D., & Baumeister, R. (2003).*Time and decision: Economic and psychological perspectives on intertemporal choice*. New York, NY, US : Russell Sage Foundation.
- Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127, 267-286.
- Schwarz, N., (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins, & R. M. Sorrentino (Eds.) *Handbook of motivation and cognition: Foundations of social behavior*, (Vol. 2., pp. 527-561) New York, NY, US : Guilford Press.
- Simon, H. A. (1957). Models of man; social and rational. Oxford, England Wiley.
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self.(In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261—302). New York: Academic Press.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193-210.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inference. *American Psychologist, 35,* 151-175.

- Zhang, L., & Baumeister, R. F. (2006). Your money or your self-esteem: Threatened egotism promotes costly entrapment in losing endeavors. *Personality and Social Psychology Bulletin*, 32, 881-893.
- Zhang, L. (2004). Self-esteem boost reduces responders' rejections in the ultimatum bargaining game. *Dissertation*. Case Western Reserve University.
- Zhang, L. (2007). An exchange theory on money and self-esteem. *Working paper*. Peking University.

Expert Commentary E

# PSYCHOLOGY OF DECISION-MAKING: 6Rs for Qualitative Research Methodological Development\*

# Zenobia C. Y. Chan<sup>1,†</sup>

Centre for Health Education and Health Promotion School of Public Health,Faculty of Medicine <sup>1</sup>The Chinese University of Hong Kong, Hong Kong

## ABSTRACT

The phenomena and facts of decision-making are closely related to our daily life, and the what, how and why of our psychological responses during the course of this process have fascinated psychologists throughout academic history. This commentary suggests the potential of qualitative approaches to research into the psychology of decisionmaking, and proposes a 6Rs framework to ensure good practice in the preparation and implementation of these studies. Finally, existing research should be recognized, and future researchers should be encouraged to consider the ability of qualitative inquiry to enrich our understanding of decision-making and other areas of psychology. (97 Words)

# COMMENTARY

Decision-making occurs in, and affects, our daily life in many ways: it allows us to choose the most appropriate actions or strategies for a particular event or task in order to attain the best outcome; it allows us to be flexible in an ever-changing world, reacting quickly to both routine and specific life matters in a timely manner; and it allows us to enhance our chances of success and minimize our chances of failure by doing things in a smart and correct

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: <sup>1</sup>The Chinese University of Hong Kong, Hong KongE-mail: zehippo@yahoo.com

manner as much as possible. In light of the importance of this process, many disciplines have put great effort into contributing to our knowledge of it and making it more transparent. Psychology is one of the most active and committed of these disciplines, making a continuous effort to explore this significant research area: humans' psychological responses toward decision-making in various cultures, when faced with various life events.

This commentary has two main objectives: to state the importance of adopting qualitative research approaches when studying the psychology of decision-making related issues; and to suggest a "6Rs framework" for quality psychological research which can reveal the what, why and how of participants' experience with decision-making in their individual contextual circumstances. Naturalistic and interpretive approaches can offer new insight into this process.

Stealing is prohibited and protecting everyone's property is stressed in this civilized world. The same is true in academic realms, where acknowledgement of any ideas, theories, research contributions and research methods suggested by any inventors, authors, scholars and researchers is a pivotal means for us to express respect and recognition of original claims and discoveries in advancing knowledge to benefit human welfare. However, in qualitative research in psychology, it seems that the importance of accreditation of our colleagues' ideas on the development and modification of an approach to exploring social facts, human behaviors or experiences in relation to health issues has not been well supported, and it may be time for an in-depth and constructive discussion. It is understandable why qualitative inquiry and its methods are perceived as secondary to quantitative research, due to the politics of methodological development. Applying standards and rigor to all aspects of research, including collection and analysis of data and publication of results, in the realm of the psychology of decision-making, are the issues I wish to draw to our readers' attention, by using some thought-provoking ideas from Dr. Morse's editorial in Qualitative Health Research (Vol. 16, 1) titled "The Politics of Developing Research Methods."

When qualitative researchers set out to develop research methods, political considerations should be discussed, including:

- (A) Who is the inventor or moderator of this research method?
- (B) Where has this research method been mentioned in an extensive and relevant literature review?
- (C) Where was this research method adopted?
- (D) In what context and setting was this research method conducted?
- (E) What advantages and disadvantages of this research method have been documented?
- (F) What of interest about the process and outcome of this research method has been noted?

As a concrete action plan to answering these questions, I suggest the "6Rs:"

• The first R - Reference others' work. This is the primary step to crediting and honoring our colleagues' contributions. Originality has been stressed as a top priority of scholarship. This is not only a matter of hard work; it is a matter of wisdom, intuition, creativity, passion for knowledge, and sensitivity in human encounters. Therefore, when someone is going to suggest, adopt or modify a

research method, they should review the related literature thoroughly and properly reference others' ideas that have affected their conceptualization and implementation of such methods. I think only when their research methods are so innovative and have truly never before been mentioned or employed, can the creator be rightfully proud of developing such methods. Otherwise, it is shameful to call the method their own creation.

- The second R Reflexivity is a crucial mental process for every researcher throughout the entire research process. They should reflect on their own assumptions, biases, knowledge base and interpretations which might affect the development of research methods. We must check in with our hearts and minds to determine whether the research method really has been developed by ourselves without any inspiration or input from other people or not.
- The third R Replication is another issue for handling the politics of development of research methods. The research results might not be the same when conducted by other researchers due to the subjectivity of analyzing and interpreting data, but the clear and systematic documentation of each research method and process can ensure replication when the same research method is employed by other researchers in other settings. So, justification for using one or more research methods should be provided. A detailed description of how to implement the research process should be offered as well.
- The fourth R Remarks. This is a significant role for an editor or reviewer when a colleague submits a paper for peer review. Critical and frank reading can help spot any research methods that have not been referenced, if these methods are known to the editor or reviewer from their own previous experience but perhaps not to the writer of the paper. Some researchers do not have certain knowledge about their suggested research methods, and may end up claiming the methods for their own unintentionally.
- The fifth R Reproach must be applied to anyone who does not reference other researchers' methods properly. Whether someone's use of another's method without giving due credit was intentional or not, there is no excuse for tolerating any attempt to claim a method as one's own when it is not. The worst case -- consciously using another's ideas without quoting the author is a grave offence.
- The sixth R Registration of one's proposed research methods. Registration systems or organizations could be developed to record qualitative research methods in an international data base. Registering research methods can provide comprehensive information about who, when, where, what and how such methods were developed. It also admits the importance of research ethics not for the participants' benefit only, but for all qualitative researchers. It allows us to learn about who they are, what their research methods are, where they adopted their research methods, who their participants were and how the research methods have been implemented. As such, I sincerely call for an international inquiry to consider how we can register our methods and categorize them both historically and by health topic.

The 6Rs is a framework proposed for researchers or scholars who are interested in using a qualitative approach to research into the psychology of decision-making. The key issue of the politics of research method development should be revisited. I deeply believe that the simple pursuit of being published is inadequate. It is time for us to remember the importance of integrity as academics, researchers and scholars and that we should be true to ourselves about the origins of the research methods stated in our studies – just as we teach our children that honesty is one of the primary responsibilities of every human being. Integrity and the 6Rs offer a guide that can help us remain aware of the politics of the development of research methods, and one that will hopefully safeguard both originality and ethics.

### REFERENCES

Morse, J. M. (2006). Editorial –. The Politics of Developing Research Methods. *Qualitative Health Research*, 16(1), 3-4.

Chapter 1

# THE BAYESIAN STATISTICAL APPROACH FOR DECISION MAKING IN MEDICAL FIELDS\*

# Roberto Bergamaschi<sup>1</sup> and Cristina Montomoli<sup>2</sup>

<sup>1</sup>Department of Clinical Neurology, Neurological Institute C.Mondino of Pavia, Italy <sup>2</sup>Department of Health Sciences, Section of Medical Statistics and Epidemiology, University of Pavia, Italy

# ABSTRACT

Medical controversies, such as the evaluation of different risks of disease or the efficacy of different treatments, are often faced with frequentist statistical analyses.

However, conventional statistics are someway improper for such decisions, because they dichotomize results according to whether they are or are not significant and do not allow decision makers to take account of additional evidence, such as previous knowledge, biological plausibility, biases in the studies, etc. A Bayesian approach can overcome these problems.

The essential components of Bayesian analysis are based on existing evidence available from other studies or from expert opinion. This evidence is summarized quantitatively via a *prior distribution* for existing beliefs. The real data are used to build a *likelihood function*. The *prior distribution* and the *likelihood function* can then be formally combined to give a *posterior distribution* for the parameters of interest. Finally, this can be interpreted as the updated beliefs or evidence about the effects of interest. Data presented as a series of posterior probability distributions are a better guide to policy. For these reasons, Bayesian methodology has been promoted as an alternative to the frequentist one when dealing with medical forecasting.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

## **1. INTRODUCTION**

Controversies over medical problems, such as the geographical distribution of a disease, the evaluation of different risks of disease, the efficacy of different treatments, are often deal with classical statistical approaches.

However, conventional statistics are someway improper for such determinations because they dichotomize results according to whether they are or are not significant and do not allow decision makers to take account of supplementary evidence, such as biological plausibility, previous knowledge, biases in the studies, etc. In addition, frequentist analyses have troubles to fit complex data sets affected by measurements errors, censored or missing observations, multilevel correlations structures and multiple endpoints (Dunson, 2001).

A Bayesian approach can be a different way to face numerous medical problems.

In the present chapter we will discuss the main characteristics of the Bayesian methodology, its advantages respect to the frequentist methodology, the utility of the Bayesian statistical approach to support decision making in several medical fields, as well as some practical applications.

# 2. THE BAYESIAN STATISTICS

Bayesian method, so called after the British scientist Thomas Bayes, who firstly developed it in the XVIII century, is a largely exploited analytical tool in many branches of applied sciences.

Bayesian analysis is based on the Bayes's theorem, that describes, given prior knowledge or opinion, an updating process based on real observations.

Formally, given an H hypothesis, the Bayes's theorem calculates the posterior probability

of the hypothesis (given data) in the following way:  $P(H \mid data) = \frac{P(data \mid H)P(H)}{P(data)}$ 

If we consider the posterior probability of a parameter  $\theta$  (e.g. the different risk of developing a disease when being exposed to two different risk factors), the theorem says that the posterior probability (or conditional density) of the parameter will be:  $p(\theta \mid data) = \frac{P(data \mid \theta) p(\theta)}{P(data)}$ 

The factors P(data | H) or  $P(data | \theta)$  are the likelihood functions estimated on the H hypothesis or on the  $\theta$  parameter.

By this way, Bayes's theorem relates the conditional density  $p(\theta | data)$  of a parameter  $\theta$  with the non conditional density  $p(\theta)$ . As  $p(\theta)$  depends on knowledge antecedent the study, it is a prior probability.

The previous knowledge is summarized quantitatively via a *prior distribution* for the value of interest, based on existing knowledge or beliefs or evidence about the parameters in the model. The actual data (new evidence) are used to build the *likelihood function*. The *prior distribution* and the *likelihood function* can then be formally combined to give a *posterior distribution* for the parameter of interest (Lilford and Braunholtz, 1996).

A possible way to describe the Bayes's theorem will be: posterior probability  $\propto$  likelihood function X prior probability where  $\propto$  stands for "proportional to" and X for "multiply by"

To summarize, the *prior probability* represents our hypothesis (possibly supported by previously acquired data such as the known incidence of a disease) before the observation of the actual data, the *likelihood function* represents the actual data, and the *posterior probability* is the final opinion derived from prior knowledge applied to the actual data. Thus, one of the essential components of the Bayesian approach is the capability to take advantages of the existing evidence (handled openly and explicitly) available from other studies or from expert opinion. Data presented as a series of posterior probability distributions would be a much better guide to policy, because they better represent the fact that real evidence is often continuous, not dichotomous (as descends by frequentist methodology).

Thanks to its characteristics, Bayesian methodology has been promoted as an alternative to the frequentist one when dealing with medical forecasting.

## **3.** THE BAYESIAN APPROACH VS THE FREQUENTIST APPROACH

The Bayesian and the frequentist approaches are two competing philosophies of statistical analysis. The frequentist analyses are presently the more commonly employed and reported in scientific literature. The Bayesian studies are fewer, also due to the computational difficulties in facing very complex data. Nevertheless, the remarkable improvement of speed and power of computers and the development of computing techniques, with particular reference to the Markov chain Monte Carlo (MCMC) sampling methods (Gilks et al, 1996), are increasing and facilitating the use of Bayesian methods. MCMC algorithms iteratively generate samples of the parameters in a statistical model. After convergence, these samples represent serially correlated draws from the joint posterior distribution of the model parameters. Based on a large number of iteratively generated samples, estimates of the posterior distribution of any parameter (or function of parameters in a model) can be easily obtainable (Dunson, 2001).

The main difference between the two theories is the way they deal with probability. Both approaches are based on the calculation of the probability and on empirical evidence; however, the Bayesian approach gives a particular relevance to the existing knowledge and to the predictive processes. In fact, as already said, Bayesian methods are based on the idea that unknown quantities follow a probability distributions. The probability distribution for a population proportion expresses our prior knowledge or belief about it, before we add the knowledge which comes from our real data (Bland and Altman, 1998).

For example, we could aim to know if a given population has a lower or a higher risk to develop a disease, compared with a control population. The frequentist statistic will answer to this question in terms of "p-value", namely showing the probability that the results should be more extreme respect to those observed under the null hypothesis H0 (that is: the risk of the analyzed population is equal to the risk of the whole population, with a probability of 1% or 5% to be wrong). This can be synthesized as: P (DATA/H0), where DATA refers to observed data. Of course, the data set considered as "the more extreme" depend to the collected data and to the characteristics of the given population. As a population analyzed in a study is

generally a selected population, thus not fully representing the whole population, different independent studies focusing on the same outcome, but analyzing different populations, could give different results. On the contrary, the Bayesian statistic will answer to the question showing the probability that the risk of the analyzed population is equal (or very similar) to the risk of the whole population, given the results collected in the study. This can be synthesized as: P (H0/ DATA). (Freedman, 1996)

In summary, the frequentist statistics make use of the probability of the results, starting from hypotheses, while the Bayesian statistics make use of the probability of the hypotheses and of the probability of future results, starting from prior knowledge (prior distribution). As a matter of fact, one of the most controversial aspects of the Bayesian approach is the choice of the prior distribution. In fact, since the prior distribution is a subjective synthesis of the available knowledge, the conclusions of the study could be influenced by the different sources and lead to findings derived from possibly overenthusiastic or overskeptical beliefs. Frequentists affirm that this specification introduces an element of arbitrariness and subjectivity into an otherwise objective procedure. However, also priors can be objective, when chosen on the basis of results obtained by previous studies. On the other hand, frequentist statistics themselves have elements of subjectivity, most notably in the choice of p<0.05 as a criterion for statistical significance (Freedman, 1996).

# 4. BAYESIAN STATISTICS FOR DECISION MAKING IN MEDICAL FIELDS

The Bayesian method is advocated as one of the most powerful tools in the investigation of complex biological processes (Berzuini and Larizza, 1996; Gilks et al, 1996). It has been successfully applied in a number of areas of medical forecasting. In the following paragraphs we will indicate some fields of the Bayesian method's application.

#### 4.1. Medical Diagnosis

The most intuitive application of Bayesian methodology is to medical diagnostic processes.

As a matter of fact, the Bayesian idea is the statistical formalization of the natural clinical process of evaluating evidence. For example, a neurologist visits a young patient suffering from headache with a recent onset. The patient is afraid because he thinks to have a brain tumour, so he estimates his risk at 20 percent. The neurologist has a much lower expectation of the risk (about 1%), because he bases his forecast on his direct experience (he has seen a number of patients suffering from headache due to causes different from brain tumours) and on the knowledge of medical literature. However, he prefers to submit the patient to a brain Computed Tomography (CT) and therefore to a brain Magnetic Resonance Imaging (MRI), before expliciting a negative diagnosis. The process of updating the patient's and the neurologist's beliefs about whether the patient has a brain cancer by using the real data (CT and MRI information) is effectively a Bayesian process. If D indicates true disease status (D = 1 if disease, D = 0 if no disease), T indicates the result of the diagnostic test (T = 1 if

29

positive, T = 0 if negative), P(D = d) is the prior probability that an individual has disease status d, P(T = t | D = d) is the likelihood of the test result t conditional on disease status, then the posterior probability of disease conditional on the test result is P(D = 1 | T = t). From the neurologist's perspective the patients's prior probability of brain cancer P(D = 1) is low. CT and MRI are used to update his prior. The resulting posterior probability forms the basis for the physician's diagnosis. Typically, physicians will order more tests (i.e., collect more data) until their posterior probability P(D = 1 | X) is close to 0 or 1, where X denotes all data collected for a patient (Dunson, 2001).

Bayesian diagnostic processes can be attained through computer-aided systems, that can offer measurable advantages over more conventional approaches. For example, in a study comparing the accuracy of the diagnosis of acute abdominal pain, a Bayesian computing system demonstrated to be much more accurate and reliable than a "human" clinician team in detect the true diagnosis (91.8% vs 79.6%) (de Dombal et al, 1972).

#### 4.2. Design of Clinical Trials

In the planning stages of clinical trials one must make important decisions which may strongly influence the conduction of the research. One of these choices is that of the minimum difference between the effects of different therapies to be considered clinically relevant. This is a very important decision, because it conditions the calculation of sample size, whose dimension strongly influence the feasibility (also economical) of the study, and can lead to results with questionable practical impact. Anyway, the choice of the minimum difference retained significant is always determined by arbitrary rules-of-thumb. A Bayesian approach, incorporating medical previous beliefs and knowledge as prior distribution, may better sustain this choice. In addition, a Bayesian model that integrates economic principles, may also be helpful to funding agencies in the organizational phase of the clinical trial (Detzky, 1985).

#### 4.3. Interim Analysis in Clinical Trials

Bayesian methods can be applied also in monitoring data accumulated during the clinical trials. In the course of a clinical trial, an interim analysis is frequently used for recommending early termination of the trial when clear benefit or harm from one of the treatments is demonstrated. However, there is no agreed method of calculating a P value or confidence interval for the treatment effect after the use of frequentist group sequential methods, nor are these methods flexible to the emergence of new external data that might influence early termination. Bayesian methods that express prior skepticism about the existence of benefit from a new treatment take account of new external data in making the final inference. For these reasons, Bayesian methodology has been proposed for the design and the monitoring of several cancer trials sponsored by Britain's Medical Research Council (Parmar et al, 1994).

#### 4.4. Analysis of the Results in Clinical Trials

The Bayesian idea can be conveniently adapted to the analysis of the results of a clinical trial. Consider a study comparing the efficacy of two different treatments (X and Y). On the basis of a frequentist analysis we can affirm that treatment X is superior to treatment Y when there is a low probability (p < 0.05) that an extreme effect difference would have been observed when the treatments were in fact equivalent. Conversely, Bayesian analysis begins with the observed difference and then estimates the probability that treatment X is really superior to Y. Bayesians induce the probability of the existence of the true but as yet unknown underlying state (in this case, X is superior to Y) given the data (Freedman, 1996).

A trouble related to the results' analyses is the interpretation of multiple hypothesis testing. Clinical trials often address the effect of a treatment by testing hypotheses about a range of putative risk factors for a given disease in different subgroups of patients. Frequentist methods aim to control the probability of finding false subgroup effects or risk factors using more stringent significance levels, such as Bonferroni procedures, where the degree of conservatism in the conclusions increases with the number of subgroup effects or risk factors tested. Bayesian methods of dealing with this multiple testing problem depend not on the number of subgroup effects or risk factors but on the prior information regarding the possibility of these effects (Freedman, 1996).

#### 4.5. Adaptive Studies

In the last years, development of new pharmacological therapies is becoming more and more demanding and expensive; in addition, some projects fail during research process. It should be advantageous to early dispose (during phase I-II studies), firstly, of data enabling the better decision making as to whether or not to continue with high-priced drug development programs and, secondly, of reliable preliminary information on which to found the design of efficient phase III trials. This is of interest of investigators and of pharmaceutical companies. As a consequence, particular attention has been recently pointed on adaptive studies, which could improve the effectiveness of clinical trials.

To better comprehend the following considerations, we will remind the main features that characterize the phases of drug development studies.

Phase I trials aim to find the maximum tolerated dose (MTD) for a new drug, while efficacy is not considered in the design (Storer, 1989). The very limited sample size in phase I trials can lead to identify unreliable MTDs, that, in turn, can influence the subsequent phase II and III trials.

Traditional Phase II trials are randomized, parallel group studies, comparing placebo and several active doses (generally three or four) of a new drug, designed to determine the optimal dosage which has the highest effectiveness as well as tolerability (Lesko et al, 2000). Phase II study should be an exploratory, learning mode of investigation, rather than a confirmatory, testing mode (Sheiner, 1997). Since the objective of this trial is to estimate a dose-response function for a fixed total number of patients, it should be better to test as many doses as possible. Anyway, a study of this type designed in a standard mode can be slightly informative, because the number of tested doses is often too small and consequently the interval between successive doses is too wide. Therefore, a major problem in clinical Phase II

studies is the incomplete understanding of the dose-response, that potentially leads to inappropriate choice of the dose to be utilized into Phase III studies (Grieve and Krams, 2005).

To identify the better dose-response function, it should be better to test more doses with fewer patients per dose, than fewer doses with more patients per dose. Anyway, increasing the number of doses is not a good option if we adopt a confirmatory stance (via frequentist testing), since whether we use 3 or 10 doses has little impact on the number of subjects needed in a single group to detect a difference of dose (Grieve and Krams, 2005). In order to face this problem, it is better to adopt a learning attitude (via Bayesian estimation). In fact, while at the beginning of a trial we know little about the dose/response curve within a dose interval, as the trial progresses, information grows as to the response of patients to differing doses. A Bayesian model can learn and use this information to adapt patients' allocation. For example, if we learn from a Bayesian design (that continuously captures outcome data) that the optimal dose-response curve is around medium doses, we could proceed with an adaptive allocation, reducing the chance of allocating patients to very low doses or to very high doses. On the contrary, with a inflexible classical study design, patients allocated to low doses should be wasted, since their response will essentially be the same as the response of patients to placebo. Similarly, patients allocated to the top doses will respond similarly to patients receiving medium doses, but they will be exposed to a major risk of side effects.

An outcome-adaptive Bayesian procedure has been proposed for assigning doses of an experimental treatment to successive cohorts of patients in breast medical oncology study. The method partitioned the two dimensional toxicity–efficacy probability domain by introducing a trade-off contour that was used to quantify the desirability of each dose. In an adaptive fashion, the Bayesian design updated the dose information coherently as more data were observed in the trial. This provided a basis for determining a best dose for each cohort. The method combined the goals of conventional Phase I and Phase II trials, and thus may be called a "Phase I-II" design (Thall and Cook, 2006).

The learning process about the dose/response curve is a useful support to decisions making on future conduction of Phase III trails or stopping early. If there is no sufficient evidence of drug efficacy, than we should stop, as continuing would be futile. Conversely, if there is enough evidence to identify a dose with an adequate risk/benefit profile, we could go into a Phase III trial. A Bayesian design based on dynamic termination rule for early discontinuation (either for efficacy or futility) was deployed in ASTIN, a phase II proof-of-concept trial of the neuroprotectant, neutrophil inhibitory factor in acute stroke (Grieve and Krams, 2005).

Adaptive Bayesian approach can improve the advancement of comparative Phase III clinical trials, where the randomization probabilities may be unbalanced adaptively by utilizing the baseline prognostic covariates and interim data available at each patient's entry time, to favor the treatment having comparatively superior outcomes. This is ethically appealing because more patients are assigned to the most successful treatments. Consequently, physicians are more likely to enroll patients into trials where the randomization is outcome-adaptive rather than balanced in the conventional manner (Cheung et al, 2006).

Adaptive studies based on Bayesian methods can accurately estimate pharmacokinetic, that has a wide inter-patients variability. A maximum a posteriori Bayesian estimator predicting individual pharmacokinetic parameters and exposure indices, has demonstrated to

be useful in the estimation of anticancer chemotherapy to prevent toxicity and to facilitate studies of the relationships between efficacy and exposure (Rousseau et al, 2000; Rousseau et al, 2002).

To summarize, adaptive studies employing Bayesian methods may increase the efficiency of decision making in phase I studies, whereas, in phases II and III, they may help to chose the right dose regimen and to perform the trials more efficiently and in a reduced time. This is fully in line with the present paradigm of "learning and confirming" in regulated clinical development (Maurer, 2005).

## **5. PRACTICAL APPLICATION OF BAYESIAN METHODS**

We have above illustrated some applications of Bayesian statistics in the management of clinical trials. Furthermore, Bayesian methods can deal with many other medical problems.

A Bayesian approach is particularly indicated to study of some aspects of chronic diseases, such as the geographical distribution, the natural history, the early prediction of the long-term evolution. In the following paragraphs, to further support the theoretical statements in favor of the Bayesian approach, we will describe practical applications of Bayesian methodology that we carried out in the field of a specific chronic neurological disease, that is Multiple Sclerosis (MS).

#### 5.1. Analysis of the Geographical Distribution of the Disease

The analysis of the geographical distribution of diseases is appealing, as it can give clues to etiological explanations, can test specific hypotheses about suspected ecological risk factors on which to decide future epidemiological studies, can promote adequate decisions about the distribution of resources.

Unfortunately, this kind of research is particularly difficult to carry out employing a classical frequentist statistical approach, especially when the aim is to study an uncommon disease, and when the disease is distributed in small areas with relatively scarce population.

A Bayesian approach, which is particularly recommended in epidemiology where knowledge increases with the accumulation of data, could be the better key to modeling the geographical distribution of a disease. It can overcome the above described problems with spatial smoothing of the rates, filtering out the random variations from the estimated prevalence.

Suppose we observe  $d_i$  events in a population of  $n_i$  individuals in area *i*. We assume:

 $d_i \sim \text{Binomial}(n_i, p_i),$ 

where '~' means 'is distributed as' and  $p_i$  is the area-specific prevalence of disease. To investigate whether prevalence shows a particular geographical distribution through the geographical areas, we would ideally like to map the true prevalence  $p_i$ . Since it is unknown,

the most obvious strategy is to calculate its maximum likelihood estimate,  $\hat{p}_i = \frac{d_i}{n_i}$ .

Mapping the  $\{\hat{p}_i\}$ , however, can be misleading because sampling variability can dominate the map and disguise genuine trends. In particular, areas that have exceptionally high or low  $\hat{p}_i$  tend to be those where sampling variability is most pronounced. Furthermore, if the disease risk is not constant within areas, the observed data may display extra-Binomial variation (overdispersion), that is, data are more variable than the Binomial model assumes.

Several strategies for dealing with sampling variability in maps have been proposed. The current state-of-the-art is to adopt a fully-Bayesian hierarchical-spatial model (Bernardinelli and Montomoli, 1992).

The model permits the decomposition of the extra-Binomial variation into two components. The first component of variation is simply a spatially unstructured extra-Binomial variation: we call it heterogeneity. The second component of variation varies smoothly across areas: we call it clustering. The model is:

 $logit(p_i) = e_i^{[1]} + e_i^{[2]},$ 

where the heterogeneity and the clustering components of variation are represented by  $e_i^{[1]}$ and  $e_i^{[2]}$ .

The  $e_i^{[1]}$  are random components which are assumed to be independently normally distributed with the mean given by the overall mean, while  $e_i^{[2]}$  are assumed to be distributed normally with mean given by the mean of neighbouring  $e_i^{[2]}$ . We consider the neighbours of an area to be those geographically adjacent to it. The sizes of  $e_i^{[1]}$  and  $e_i^{[2]}$  are controlled by variance parameters. The variance in  $e_i^{[1]}$  and  $e_i^{[2]}$  is assumed to be distributed as a X<sup>2</sup> variable.

The model for  $p_i$  is a Bayesian hierarchical model, since we specified a prior distribution at two levels of the model ( $e_i^{[1]}$  and  $e_i^{[2]}$  and their variance parameters).

To estimate the unknown parameters  $p_i$ , we have to combine our prior belief about the risks, which is embodied in the prior distribution of *heterogeneity* and *clustering* with the information contained in the data (likelihood). This is a familiar operation within a Bayesian statistical framework. Deriving the analytical form of the desired posterior distribution and the likelihood is often difficult, since it involves analytically intractable integrations. The solution is supplied by a Markov chain Monte Carlo technique which can be used to generate samples from the posterior distribution (Gilks et al., 1993) and hence to obtain the entire posterior distribution of the parameters of interest,  $p_i$  in our model.

Summaries of the posterior distribution of  $p_i$  can be computed; for example, the point estimate of prevalence in area I can be computed through the mean of the posterior distribution of  $p_i$ . Moreover, for each area, we can compute the proportion of the posterior samples which have a mean greater than the global posterior mean, that is, the overall mean prevalence. This quantity is the posterior probability (PP).

Let us now discuss the effect of specifying the heterogeneity or the clustering prior to the estimate of  $p_i$ . The effect of an a priori distribution for the area-specific prevalence is to yield,

for each area, an estimate that represents a compromise between the area-specific prevalence and a reference value. Under the *heterogeneity* component of variation, the reference is the mean of all the prevalence values in the map. Under the *clustering* model, the reference for any area-specific prevalence is the mean of the prevalence values of the neighbouring areas. Extreme prevalence estimates are thus pulled towards the reference, the pulling being more substantial when the estimates are unstable, that is when the corresponding areas do not contain a high number of cases, and thus do not provide substantial evidence in favour of the extreme value. The result of this mechanism is that the estimated maps tend to be smoother and more interpretable from an epidemiological point of view.

The choice between the *clustering* and the *heterogeneity* model depends upon our prior belief about the size of high/low risk clusters. A cluster size bigger than the area size leads to a *clustering* model, while a cluster size smaller than the area size leads to a *heterogeneity* model. Although it is possible to include both terms in the model, this may not be necessary. Indeed, for high resolution maps the *heterogeneity* term will often be unnecessary, and, if the number of events per area exceeds 1, is not even identifiable. In our application we chose the *clustering* model for the analysis at a fine geographical scale (Bernardinelli and Montomoli, 1992; Gilks et al, 1993).

We used a Bayesian approach to study the prevalence of multiple sclerosis (MS) in a specific geographical area.

MS differentially affects women, people 30-60 years old, and Caucasians. Evidence indicates that it is a complex disease determined by both environmental factors and genetic susceptibility. Prevalence studies finding clusters of MS in specific small areas could underscore the role of hypothetical environmental exposures in the development of the disease. As a consequence, they could support the decision making for additional epidemiological studies and engagement of resources.

We described the application of the Bayesian modeling compared to the frequentist methodology in the study of the geographical distribution of MS across the province of Pavia (northern Italy, 2965 km2, 493,753 inhabitants) subdivided into 190 very small sub-areas (municipalities) (Bergamaschi et al, 2006).

The frequentist prevalence rates were calculated for each municipality simply by dividing the number of MS cases by the resident population.

The main idea underlying the hierarchical Bayesian model we used was that in each areaspecific prevalence was based on pooling information from neighboring small areas (prior distribution). The Bayesian maps of prevalence were accompanied by a map of the posterior probability (PP) that indicated if the prevalence rate for each area was significantly lower or greater than a given reference value. The PP is the Bayesian equivalent of the p-value (Meng et al, 1987) and enables the testing of specific hypotheses. In our analysis, for example, we choose the median age-standardized MS prevalence rate of the whole province of Pavia as a reference value in order to test whether each area-specific prevalence was different from the median provincial values. We subdivided the range of PP (0-1) into five intervals (<0.10, 0.10-0.25, 0.25-0.75, 0.75-0.90, >0.90). A PP value higher than or equal to 0.90 strongly indicated that the area-specific risk was higher than the reference value, while a PP value smaller than or equal to 0.10 strongly indicated that the prevalence was lower than the reference value. In those areas where the PP value fell within the fourth interval (0.75-0.90), there was only an indication that the risk was higher than the reference value. Similarly, in those areas where the PP value belonged to the second interval (0.10-0.25), there was an indication that the prevalence was lower than the reference value. When the PP value fell within the central interval (0.25–0.75), there was not enough evidence to judge. In conclusion, the map of PP can be used to evaluate how confident we should be when we analyze a prevalence value that is much higher or lower than a reference value. We used the WINBUGS software package for the Bayesian analysis (Spiegelhalter, 1999).

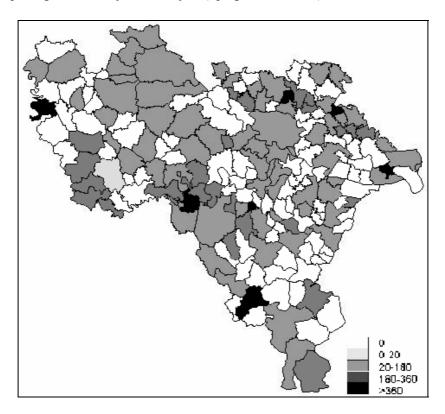


Figure 1. Map of the crude MS prevalence obtained with a frequentist approach (from Bergamaschi et al., J Neurol Sci, 2006, 244, 127–131).

The Bayesian area-specific prevalence ranged from 84 to 126/100,000 inhabitants.

We built the maps of the geographical variations of MS prevalence across the 190 communes of the province both with a classical frequentist approach and a Bayesian approach. The maps had to be interpreted by considering that different shades of grey were proportional to the prevalence value. So the darker the area was, the higher the prevalence of disease was.

The frequentist approach produced an extremely dishomogeneous map (Figure 1), strongly influenced by random variability. This map was dominated by extreme prevalence values found in the areas with scarce population.

On the contrary, the Bayesian map (Figure 2) was much smoother.

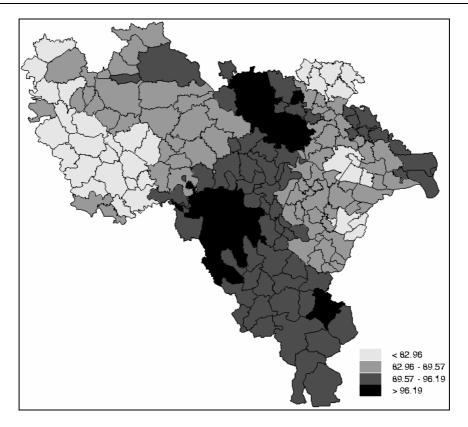


Figure 2. Map of the Bayesian estimates of MS prevalence rates (from Bergamaschi et al., J NeuroL Sci, 2006, 244, 127–131).

The PP Bayesian map (Figure 3) revealed two small clusters where the PP value fell into the last category (PP>0.90), indicating that for these areas there was evidence that the prevalence was significantly greater than the median provincial rate.

To summarize, the use of a conventional statistical approach seems not to be suitable for this kind of analyses. In fact, the frequentist map of the disease resembles a patchwork, being strongly subject to random error. As a consequence, the map is not interpretable from an epidemiological point of view. On the contrary, the Bayesian approach overcomes these problems with spatial smoothing of the rates, both allowing us to estimate area-specific prevalence rates and filtering out the random variations from the estimated prevalence due to the small number of cases in each commune, thus showing the true underlying variations of MS prevalence.

We think that our study is a good example of the usefulness of Bayesian methods for the epidemiological analysis in small areas with small populations, providing the advantages to obtain reliable maps of disease prevalence, to identify interesting clusters with a high prevalence of the disease where deserve further epidemiological investigations and engage resources.

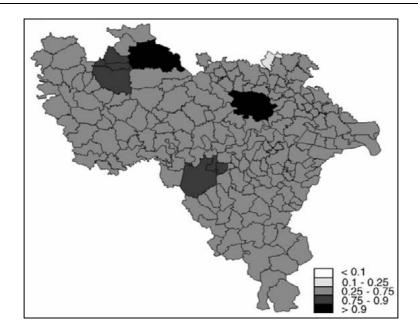


Figure 3. Map of the Bayesian posterior probability that each area-specific MS prevalence rate is higher than the median MS provincial rate (*from Bergamaschi et al, J Neurol Sci, 2006, 244, 127–131*).

#### 5.2. Disease Natural History and Prognosis

The majority of the works in this area are based on statistical models which relate the early manifestations of a disease, as observed in the patient before a specified baseline time, to the time at which a specified endpoint event, or failure, occurs. However, in those studies, the statistical analysis often ignores the uncertainty affecting prebaseline information, and it also ignores the rich wealth of information available from postbaseline observations, which may be used as a surrogate for the censored failure times.

A Bayesian model overcomes these limitations. The model specifies the full joint probability distribution for a set of random variables which characterize the entire course of the disease, including early indicators observed before baseline, intermediate indicators observed between baseline and failure, and time-to-failure. In the Bayesian model, the intermediate indicators are treated as a surrogate response event, so that in rightcensored patients they provide supplementary information pointing towards the unobserved failure times. Moreover, a full probability modeling approach allows us to incorporate in the analysis the considerable uncertainty which affects certain early indicators. By using a model with the aforementioned features, the ability of early indicators to predict failure can be assessed more accurately and reliably and, moreover, explained in terms of the relationship between early and intermediate indicators. The essence of this statistical approach is to regard disease course as a random realization of a random multidimensional stochastic process (Berzuini and Larizza, 1996).

We applied a Bayesian model aiming to study the natural history and the prognosis of Multiple Sclerosis (MS).

MS is a chronic disimmune disease of young adults. The natural course of MS is characterized by several acute episodes of neurological impairment, followed by a partial or complete remission with clinical stability between attacks. The latter is defined as relapsing-remitting phase. After a mean time of 10 years, the relapsing-remitting phase is followed by the secondary-progressive phase, which is characterized by progressive clinical impairment, with or without superimposed relapses and remissions, that eventually leads to irreversible disability. Only 10 to 15% of MS patients accumulate a progressive disability from the beginning, referred to primary progressive course (Bergamaschi, 2006).

MS course is classically described as unpredictable, because it is spotted by several events characterized by different neurological patterns, each spaced by different time intervals.

Since 1993, the use in MS treatment of Disease Modifying immunomodulating Therapies (DMTs) has partially changed the natural evolution of the disease. In particular, DMTs have demonstrated to be able to reduce the risk of relapses and of new brain lesions. Moreover, some clinical trials have demonstrated the efficacy of early treatment in patients with a first clinically isolated syndrome (CIS) suggestive for MS onset. On the basis of this findings, the American Academy of Neurology Subcommittee affirmed that it is appropriate to consider DMTs for treatment in any patient who is at high risk for developing clinically definite MS. Hence, the broad and early use of DMTs in MS patients is strongly encouraged, although it is still uncertain if DMTs are really effective in slowing the disease's progression and some criticisms should limit the particularly early use of DMTs. The probability of a patient with an attack suspected for MS of having a second clinical relapse and then becoming a clinically definite MS patient is directly related to the duration of the disease. Thus, even if the conversion from CIS to the clinically definite diagnosis of MS is quite frequent, it might happen a long time after onset, or not at all. In addition, studies proved that 15-25% of MS patients show a "benign" course, which is characterized by a mild or absent disability despite a long duration of the disease. Finally, we must take into account the cost-benefit ratio that the chronic use of DMTs entails, since these therapies are onerous both for the health service, not only in economic terms, and, above all, for the patients, who are required to do frequent injections and to be exposed to possible adverse events. As the balance between the pros and cons of using DMTs is still uncertain, it is desirable to assess reliable predictors of a poor prognosis, useful to select the patients at higher risk.

We exploited a Bayesian approach with a Markov chain Monte Carlo simulation to model the natural history of 186 MS patients with initial relapsing-remitting (RR) course, in order to foresee the risk of reaching an unfavourable end-point, that is the secondary progression (SP) (Bergamaschi et al, 2001). The selected patients were never treated with immune "preventive" therapies, thus fairly represented the natural history of the disease.

The model was based on the analyzes of a large number of clinical variables collected since the onset of the disease for a total period of pre-endpoint follow up of 1112 person years, with a mean follow-up of 7 years. The Bayesian model allowed us to build the graph in Figure 4 as a faithful representation of the main qualitative relationships among the prognostic variables of MS. This graph represented a selected subset of the early predictors (nodes near the bottom edge of the figure), as influencing patient proneness,  $\Phi$ , and it represented  $\Phi$  as influencing a selected subset of intermediate predictors (nodes near the left edge). The thickness of the arrows in Figure 4 reflected their corresponding estimated strengths. For example, the thick arrow leading from  $\Phi$  to "disability score (named EDSS) outside relapse, years 2–3" reflects strong data evidence that patients with higher proneness to failure tend to exhibit a higher rate of EDSS growth during years 2 and 3. According to the graph, the disease evolution pattern in patients highly prone to failure is best characterized during years 2 and 3 in terms of a higher rate of sphincter–pyramidal involvement, and of a faster rate of deterioration of the ability to walk, as expressed by the disability scale EDSS.

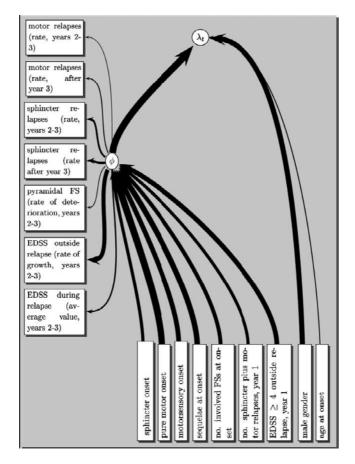


Figure 4. Graphical model representing MS natural history. Nodes near the left edge: intermediate predictors; nodes near the lower edge: early predictors;  $\Phi$ : patient proneness to failure;  $\lambda$  (t): patient's hazard of failure (*From: Bergamaschi et al, J Neurol Sci, 2001,189, 13–21*).

The strength of the arrows from the early manifestations to  $\Phi$  was quantified by the corresponding estimated local relative risks (LRR). Therefore, each prognostic factor was associated to a specific statistical "weight", the Bayesian local relative risk (LRR), interpretable in the same way of the relative-risk of the frequentist statistic.

The results from this Bayesian analysis turned out to be largely consistent with those of the previous studies regarding onset variables (the female gender had a protective effect, while a late age at onset, a polysimptomatic onset, a motor and sphincter onset, and an sequele outside relapse, were unfavorable prognostic factors). Anyway, the main particularity of the Bayesian model was the identification of post-onset adverse indicators, which only few studies described. As already said, a high number of relapses in the first years was previously found to have an unfavorable prognostic value, but the Bayesian model added a qualitative feature to these findings: the importance of the type of relapses. Specifically, the motor/sphincter relapses have a strong impact on the prognosis, whereas certain types of relapses appear to have no impact at all. This suggests that when we study a long-term disease, we should not merely look at the number of events, but that we should rather consider the relative importance of their different characteristics. Currently, some clinical therapeutic trials in MS use quantitative endpoints such as the "total count of relapses". The choice of these earthy endpoint is partially due to the difficulty to manage complex qualitative data by frequentist statistics. In the light of Bayesian results, such an endpoint appears to be inappropriate, since it undervalues the fact that certain types of relapse are more important than others.

Thanks to the Bayesian model of the natural history of MS, we were able to calculate, for any given patient, a risk score based on information collected within the first year of the disease. This score was obtained by the sum of the statistical weights over the appropriate log LRRs, according to a simple algorithm. Afterwards, the risk score was tested on a new and larger sample of individuals (1245 MS patients, mean follow-up 17.1 years), in order to verify its trustworthiness and its capability to really predict the disease evolution (Bergamaschi et al, 2007). The score was significantly related (p<0.000001) to the unfavourable end-point (to reach secondary progression) (Figure 5).

As a diagnostic test, the Bayesian score was extremely specific. In fact, patients with high ( $\geq$  95th percentile) values of risk score had a very high probability of reaching SP within 10 years (specificity 0.99), while all the patients with low values ( $\leq$  5th percentile) remained progression-free. Thanks to its specificity, the Bayesian score can be used as a useful tool to support the treatment decision making. In the case of MS, although immune treatment in all patients with a diagnosis of relapsing-remitting MS is considered appropriate, we should recognise that patients destined to have a favorable course would benefit from not receiving treatments. With a Bayesian score available for each single patient, we could wait and see in the case of low risk patients, reserve early treatment or very efficient but relatively unsafe therapies (such as hard immunosuppressors) for high risk patients, and consider the use of soft immunomodulators in the remainder.

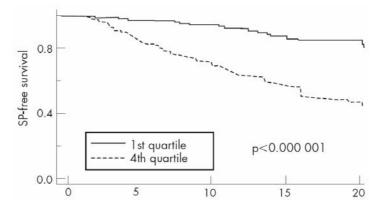


Figure 5. Kaplan–Meier curves for estimates of the time needed to reach the secondary progressive (SP) phase, starting from the 1 year duration of disease mark, among patients with multiple sclerosis who were never treated with immune therapies, with low or high Bayesian Risk Estimate for Multiple Sclerosis values ( $\leq$ 1st quartile or  $\geq$  4th quartile) *(from Bergamaschi et al, J Neurol Neurosurg Psychiatry 2007).* 

We think that our study is a good example of the usefulness of Bayesian methods for the analysis of the natural history of chronic diseases. A Bayesian model allows us to characterize the pattern of disease course in high-risk patients, and to identify short-term manifestations which are strongly related to long-term evolution of disease, as potential surrogate markers. It is also able to predict the long-term evolution of individual patients on the basis of observations taken during an early stage of the disease. Such forecasts are obviously important to the patients who want to be informed about their prospects, and to the clinician to identify at an early stage of disease those patients who urgently need to be treated, before

the therapy becomes ineffective. Finally, Bayesian predictive studies may provide a useful contribution to rationalizing the design of clinical therapeutic trials and the choice of therapy in clinical practice.

## **6.** CONCLUSION

Statistical frequentist approaches are not always sufficiently adequate to cope with medical problems, whose solution is often related to decision making.

During the last years, the Bayesian approach has been strongly encouraged as support to medical decision making, because Bayesian analyses take into account the prior knowledge, fully exploiting previous evidence available from other studies or from expert opinion, and use them to build a function of the real data, producing a series of posterior probability distributions useful to guide policy. The recent improvement of computing techniques have facilitated the employment of Bayesian methods.

For these reasons, Bayesian methodology has been promoted as an alternative to the frequentist one when dealing with medical forecasting.

The possible applications of Bayesian methodology as support for decision making are numerous.

It can be useful in diagnostic processes and in the different phases of clinical trials, from the design to the final analysis. Particularly, it get better the choice of the minimum difference in treatment efficacy to be considered significant and the computation of the number of patients to be included in a trial. They also contribute to the interim analysis (and eventually to the decision of early discontinuation of the study), as well as to the final analysis.

Bayesian philosophy is the basis of adaptive studies, addressed to increase the efficiency of clinical trails.

Bayesian analysis is also exploitable to the study of the geographical distribution of specific diseases, whose awareness can give clues to etiological explanations, can test specific hypotheses about suspected ecological risk factors, can promote new epidemiological studies and adequate decisions about the distribution of resources.

Finally, an appealing application of the Bayesian approach is the modeling of the natural history of chronic diseases and, as a consequence, the accurate and reliable identification of early predictors of their long-term evolution. This can be useful to: identify the high-risk patients who could require early or more aggressive therapies; identify the low-risk patients who could avoid lifelong, expensive and potentially troublesome treatments; favor a more homogeneous selection of patients for clinical therapeutic trials; evaluate the effect of

therapies in the field of observational studies, facilitating an a posteriori subdivision of nonrandomized patients on the basis of their different "propensity" to reach a poor end-point.

## REFERENCES

- Bergamaschi R, Berzuini C, Romani A, Cosi V. Predicting secondary progression in relapsing-remitting multiple sclerosis: a Bayesian analysis. J Neurol Sci, 2001,189, 13– 21.
- Bergamaschi R. Prognosis of multiple sclerosis: clinical factors predicting the late evolution for an early treatment decision. *Expert Rev Neurotherapeutics*, 2006, 6, 357-364.
- Bergamaschi R, Montomoli C, Candeloro E, Monti MC, Cioccale R, Bernardinelli L, Fratino P, Cosi V. Bayesian mapping of multiple sclerosis prevalence in the province of Pavia, northern Italy. *J Neurol Sci*, 2006, 244, 127–131.
- Bergamaschi R, Quaglini S, Trojano M, et al. Early prediction of the long-term evolution of multiple sclerosis: the BREMS score. J. Neurol. Neurosurg. Psychiatry, in press (published online 12 Jan 2007)
- Bernardinelli L, Montomoli C. Empirical Bayes vs fully Bayesian analysis of geographical variation in disease risk. *Stat Med*, 1992, 11, 983–1007.
- Berzuini C, Larizza C. A unified approach for modelling longitudinal and failure time data, with application in medical monitoring. *IEEE Trans Patt Anal Mach Intell*, 1996, 2, 109– 23.
- Bland JM, Altman DG. Bayesian and frequentists. BMJ, 1998, 317, 1151.
- de Dombal FT, Leaper DJ, Staniland JR, McCann AP, Horrocks JC. Computer-aided diagnosis of acute abdominal pain. *BMJ*, 1972, 2, 9-13.
- Cheung YK, Inoue LY, Wathen JK, Thall PF. Continuous Bayesian adaptive randomization based on event times with covariates. *Stat Med*, 2006, 25, 55-70.
- Clayton DG, Bernardinelli L, Montomoli C. Spatial correlation in ecologic analysis. Int J Epidemiol, 1993, 22, 1193–202.
- Detzky. Using economic analysis to determine the resources consequences of choices made in planning clinical trials. *J Chronic Dis*, 1985, 38, 753-765.
- Dunson DB. Practical Advantages of Bayesian Analysis of Epidemiologic Data. Am J Epidemiol, 2001, 153, 1222-1226.
- Freedman L. Bayesian statistical methods. BMJ, 1996, 313, 569-570.
- Gilks WR, Clayton DG, Spiegelhalter DJ, Best NG, McNeil AJ, Sharples LD, Kirby AJ: Modelling complexity in medicine: Applications of Gibbs sampling in medicine. *J R Stat Soc B*, 1993, 55, 39-52.
- Gilks WR, Richardson S, Spiegelhalter DJ, eds (1996). *Markov chain Monte Carlo in practice*. Boca Raton, FL: CRC Press.
- Grieve AP, Krams M. ASTIN: a Bayesian adaptive dose-response trial in acute stroke. *Clinical Trials*, 2005, 2, 340-351.
- Lesko LJ, Rowland M, Peck CC, Blaschke TF. Optimizing the science of drug development: opportunities for better candidate selection and accelerated evaluation in humans. *J Clin Pliarmacol*, 2000, 40, 803-14.

- Lilford RJ, Braunholtz D. The statistical basis of public policy: a paradigm shift is overdue. *BMJ*, 1996, 313, 603-607.
- Maurer W. Creative and innovative statistics in clinical research and development. *Methods Inf Med*, 2005, 44, 551-560.
- Meng CYK, Dempster AP. A Bayesian approach to the multiplicity problem for significance testing with binomial data. *Biometrics*, 1987, 43, 301–11.
- Parmar MK, Spiegelhalter DJ, Freedman LS. The CHART trials: Bayesian design and monitoring in practice. CHART Steering Committee. *Stat Med*, 1994, 13, 1297-312.
- Rousseau A, Marquet P, Debord J, Sabot C, Lachatre G. Adaptive control methods for the dose individualisation of anticancer agents. *Clin Pharmacokinet*, 2000, 38, 315-53.
- Rousseau A, Sabot C, Delepine N, Delepine G, Debord J, Lachatre G, Marquet P. Bayesian estimation of methotrexate pharmacokinetic parameters and area under the curve in children and young adults with localised osteosarcoma. *Clin Pharmacokinet*, 2002, 41, 1095-104.
- Sheiner LB. Learning versus confirming in clinical drug development. *Clin Pharmacol Therapeut*, 1997, 61, 275-91.
- Spiegelhalter DJ, Thomas A, Best NG, Gilks WR, editors (1995). *BUGS: Bayesian Inference Using Gibbs Sampling*. Cambridge: MRC Biostatistics Unit.
- Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Methods in health service research. An introduction to Bayesian methods in health technology assessment. *BMJ*, 1999, 319, 508–12.
- Spiegelhalter DJ, Thomas A, Best NG (1999). WinBUGS version 1.2 user manual. Cambridge' MRC Biostatistics Unit.
- Storer, BE. Design and analysis of phase I clinical trials. *Biometrics*, 1989, 45, 925–937.
- Thall PF, Cook JD, Estey EH. Adaptive dose selection using efficacy-toxicity trade-offs: illustrations and practical considerations. *J Biopharm Stat*, 2006, 16, 623-38.

Chapter 2

# NURSE - PHYSICIAN COMMUNICATION: THE IMPACT OF DISRUPTIVE BEHAVIORS ON FACTORS AFFECTING DECISION MAKING AFFECTING PATIENT OUTCOMES OF CARE<sup>\*</sup>

Alan H. Rosenstein and Michelle O'Daniel

VHA West Coast, California, USA

## ABSTRACT

Disruptive behaviors can have a profound effect on staff perceptions, attitudes and reactions that affect decision making and communication flow. Feelings of anger, hostility and frustration lead to impaired relationships, confused expectations, and unclear roles and responsibilities which can impede the transfer of vital information that can adversely affect patient outcomes of care. Health care organizations need to be aware of the significance of disruptive behaviors and develop appropriate policies, standards and procedures to effectively deal with this serious issue.

## INTRODUCTION

Medical care is a complex system with multiple different parties interfacing and exchanging tasks and information important for patient care. When things go well, everyone benefits. But when things don't go well, the unintended consequences can be severe. Reports from the Institute of Medicine state that nearly 100,000 deaths each year occur because of preventable medical errors [1-3]. The Joint Commission on Accreditation of Hospitals (JCAHO) states that nearly two- thirds of sentinel events (an unexpected medical event that

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

has serious medical consequences) can be traced back to a communication error [4]. The newest initiative from the Institute for Healthcare Improvement (IHI) is unveiling its five million lives saved from harm campaign that focuses on interventions that reduce the likelihood of patient harm [5].

Traditional approaches adopted by the patient safety movement have focused more on improving system related issues than people related issues. While the introduction of new technologies and redesigned safety processes and procedures have made a significant impact, there is still a large opportunity for improvement [6]. But in order to get to the next level, organizations need to be willing and able to address the human factor issues affecting emotions, attitudes and capability to perform the necessary functions needed to provide optimal patient care.

#### BACKGROUND

For several years we have been conducting research on the frequency and significance of physician disruptive behaviors in the hospital sector. The research was originally conducted to see if there were a link between disruptive physician behaviors and nurse satisfaction and turnover as a contributing factor to the hospital nursing shortage. Review of the literature at that time revealed no conclusive findings so we decided to develop our own research survey tool. For the survey instrument disruptive behavior was defined as any inappropriate behavior, confrontation or conflict ranging from verbal abuse to physical or sexual harassment. Results from this research showed a significant linkage between physician disruptive behavior, nurse dissatisfaction and nursing turnover [7,8]. Not only did the research study quantify the frequency, significance and impact of disruptive physician behaviors, it also opened up a series of other important questions around the occurrence of nurse and other staff disruptive behaviors and the overall impact of disruptive behaviors on patient outcomes of care.

The second phase of research built on the findings of the previous research and added questions on the incidence and impact of nurse disruptive behaviors as well as other hospital staff employees. We were also intrigued about concerns raised about the impact of disruptive events on patient outcomes of care [9,10]. When we researched the literature, we could find references attesting to the benefits of more nurse time being spent at the patient's bedside and the lower likelihood of adverse clinical events (adverse clinical events are defined as any unexpected undesirable clinical patient experience, adverse occurrence or medical error that occurred during the hospitalization), and the benefits of good team collaboration on improving patient outcomes of care [11]. Other than a few anecdotal stories, we could not find any data supported evidence of poor clinical outcomes related to disruptive behaviors, poor communication or ineffective teamwork and collaboration [12-14]. With this objective in mind we added a series of questions to the phase two survey which were specifically designed to focus on the impact of disruptive behaviors on human factor issues and behaviors affecting patient outcomes of care.

## **FINDINGS**

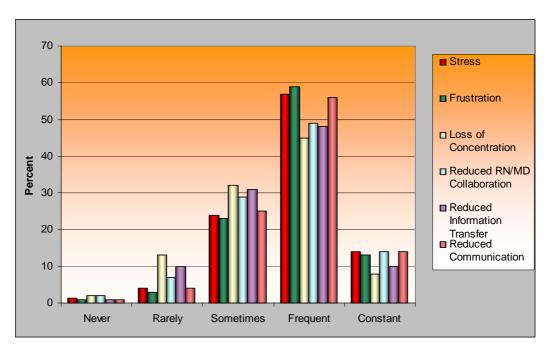
To date our research database has results from over 5,000 participants from more than 150 hospitals across the United States. Overall, 75% of the respondents witnessed disruptive behavior in physicians. The most frequently mentioned medical specialties included General Surgery, Cardiovascular Surgery, Cardiology, Neurosurgery, Neurology, Anesthesiology and Orthopedics. 65% of the respondents witnessed disruptive behaviors in nurses. Most interestingly, more nurses witnessed disruptive behaviors in their peers than other groups witnessing disruptive behaviors in nurses. We also noted a significant concern about disruptive behaviors in respiratory therapists, lab technologists, pharmacists and physical therapists. The current stage (phase three) of the survey focuses on high intensity areas. The first special study was conducted specifically on peri-operative services. The surgical environment represents a rather unique micro chasm of medical care in that it is a high intensity service, conducted in a relatively small spatial confine with strong interdependence on team collaboration and assistance. Our results showed that the frequency and seriousness of disruptive events were more intensified in the surgical setting [10]. Special studies on the Intensive Care units (ICUs) and Emergency Departments (EDs) are planned for the near future.

The phase two and phase three surveys also contained a series of questions designed to assess the association between disruptive behaviors and adverse events. 38% of the respondents reported that they were aware of a potential adverse event that could have occurred as a result of disruptive behavior. 14% reported that they actually witnessed an adverse event as a direct result of disruptive behaviors. No one ever wants an adverse event to occur. The average occurrence rate of adverse events in US hospitals is 2-3%. This makes the 14% occurrence rate even more striking. Some of the comments provided in Table I highlight some of the critical concerns raised by some of the survey participants. The good news is that 80% of the respondents felt that these events could have been prevented.

In an effort to determine if there were a linkage between disruptive behaviors and psychological factors known to impede decision making, we asked a series of questions on the impact of disruptive behaviors on stress, frustration, loss of concentration, reduced team collaboration, reduced information transfer, reduced communication, and impaired nurse physician relations. Responses were recorded on a 10 point Likert scale. The results showed a significant impact of disruptive behaviors on all factors surveyed (see Figure I). Stress and frustration lead to loss of focus and impaired concentration which can lead to problems with communication and transfer of vital information which can ultimately result in a compromise in patient care.

In an effort to assess the relationship between disruptive behaviors and clinical outcomes of care, we asked a question about the perceived linkage between disruptive behaviors and the occurrence of adverse events, medical errors, compromises in patient safety, quality and patient mortality. Responses were recorded on a 10 point Likert scale. The results showed a significant impact of disruptive behaviors on all of the surveyed factors (see Figure II)

The concern raised by these findings is the significant impact disruptive events can have on behavioral factors known to increase stress, anxiety and frustration, which can impair focus and concentration. Reactions stimulated by these events can either directly or indirectly



impair needed exchange of information and carrying out of responsibilities so important for patient care [15].

Figure I. How Often Does Disruptive Behavior Result in the Following?

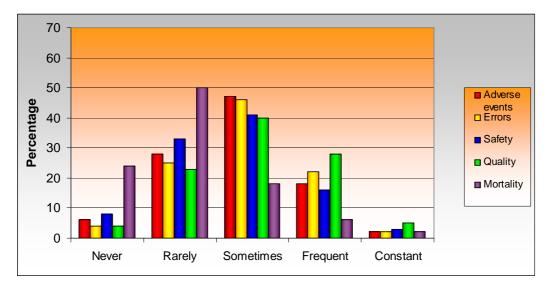


Figure II. Is How Often Do you Think There is a Link Between Disruptive Behaviors and the Following Clinical Outcomes of Care?

## Table I. Selected Comments

- Most nurses are afraid to call Dr. X when they need to, and frequently won't call. Their patient's medical safety is always in jeopardy because of this.
- Staff was afraid to approach MD despite fact order was clearly incorrect. Disruptive behavior policy "wildly ineffective" Complaints delivered to a "black hole:"
- Cardiologist upset by phone calls and refused to come in. RN told it was not her job to think, just to follow orders. Rx delayed. MI extended
- MD became angry when RN reported decline in patient's condition and did not act on information. Patient required emergency intubation and was transferred to ICU. This caused the family much unnecessary heartache and disruption in family grieving process
- Poor communication post-op because of disruptive reputation resulted in delayed treatment, aspiration and eventual demise
- MD was told twice that sponge count was off. She said "they will find it later". Patient had to be re-opened.
- Dr. X has the social skills of a two year old
- The disruptive behavior from nurses is much more upsetting because I expect that behavior from the surgeons NOT the nurses b/c I rely on them as my peers (RN)
- When patient brought to unit for GI bleeding patient saw MD yelling at nurses. Patient asked if that was his doctor. Yes. Patient refused treatment and was transferred to another hospital. I am retiring early and never recommend someone becoming a nurse
- Disruptive behavior results in medication errors, slow response times and treatment errors!
- Specific group of MDs with poor communication habits which adversely affect patient care and nurse role
- Dr. X is so condescending to nurses that we are very reluctant to call him on patient's behalf.
- My concern is that the new nurses are afraid to call about patient problems and issues that truly need to be addressed in a timely manner impacting outcomes.
- It seems there is an increasing lack of respect by nurses and other ancillary caregivers toward practicing clinicians in the hospital environment. Decisions by clinicians are frequently challenged and some orders are flatly disobeyed or at least not carried thru with almost reflexive propensity, and with little or no forethought
- "Are you aware of any specific adverse events ....?" Yes. Death as a result of disruptive behavior. Staff nurses advocated for better patient care but MD would not willing to listen to reason. As a result patient died. The doctor chose to undo all the help that various staff had been working on for weeks to get this patient the help so badly needed.
- Yes, many incidents are preventable if both parties are willing to listen to each other, but many doctors are unwilling to accept a nurse's opinion just as some nurses are unwilling to listen to the opinions of LVNs, techs or CNAs, and it may have to do with the entrenched pecking order that exists at most hospitals.

## **CAUSATIVE FACTORS**

Human nature is complex, health care is a complex science and the health care environment functions in a bed of complexity. This implies that there are multiple factors involved in causing or stimulating disruptive events, none of which are mutually exclusive. Highlighted below are a series of factors known to contribute to disruptive behaviors (see Table II). Causative factors become particularly important when trying to develop and implement programs and strategies that will reduce the likelihood of disruptive behaviors and the resulting downside effects.

#### **Table II. Factors Contributing to Disruptive Behaviors**

•	Setting	
•	Personality	
٠	Gender	
٠	Generation	
•	Culture and Ethnicity	
•	Life experiences	
•	Training	
•	Stakeholder interest	

#### Setting

Disruptive events are inappropriate regardless of the cause or stimulating factor and there is never an excuse or rationalization for sanctioning such events as excusable. We noted that many of the disruptive events were evoked during a stressful situation. A common background would be a patient suddenly taking a turn for the worse or a procedure not going as planned. Another situation leading to a disruptive response is the unexpected call to the physician for clarification of orders or a discussion around plans for treatment or disposition, especially if calls occurred at night and/ or were directed at a physician covering for another physician.

Another factor contributing to the likelihood of occurrence of a disruptive event is either previous experience with the person or place and/ or reputation. Having had a previous bad experience with the individual or place or having prior knowledge of a bad reputation sets the stage for bad expectations which can increase the risk of a disruptive response.

The setting may play a role in setting up the potential for a disruptive outburst, but in most cases the potential lies more in the behind the scenes makeup of the individuals involved.

#### Personality

Suggestions have been made that disruptive events are more likely to occur from inherent personality traits than as a result of a specific inciting event. There are many different

categorizations of personality classifications including Myers- Briggs, Personalysis, and others. For sake of discussion, we will break the personality types into four main groups: Directors, Socializers, Thinkers and Relaters [16].

Directors are described as being very task and results oriented, direct and to the point, demanding, and have a predilection for working independently. Under pressure they become more intense, demanding and controlling, and get results more by applying pressure and bullying than listening.

Socializers are described as being warm and sensitive, more people oriented, have a greater need for recognition and being appreciated, and value communication and collaboration. Under pressure they may become more dramatic, insistent and overbearing in their need to discuss and have their opinions heard.

Thinkers are intelligent, very systematic in their approach to a problem, are very precision and detail oriented, are slow, deliberate and meticulous in their methods, and with enough time, come up with well thought out analytical solutions. Under stress they tend to withdraw, work more independently and will lose interest in outside input.

Relaters are people oriented, tend to get along with everyone, are very agreeable, they don't want to rock the boat, and want to be liked. They shy away from confrontation and prefer to communicate more indirectly. Under pressure they may become more submissive, emotional or act in a passive- aggressive manner.

#### Gender

There has been a lot of research on gender characteristics and the role they play in influencing intention, action and reaction behaviors. If we follow the classic descriptions provided by John Gray PhD. in his text "Men are From Mars Woman are From Venus" we can use his assessment of male and female traits as a point for discussion and relate this to the health care environment [17]. We fully recognize that these are stereotypic descriptions and should not be the basis of conclusion, but our experience has shown us that these characteristics do play out in the medical arena.

Males are described as being very focused, task and goal driven, time dependent, power hungry, self-achievers, who tend to work in an autonomous independent fashion. They look for competency and efficiency in their co-workers. When under stress they either become more dominating and authoritative or withdraw to work things out on their own.

Females are described as being more social minded and communicative, interested in developing relationships, are more accepting and willing to listen and to share, are more sensitive and caring, and they value group consensus and companionship. When under stress they look more to group input and discussion as a way of solving the crisis.

In the health care environment, most of the physicians are male and most of the nurses and other clinical support staff are female. Under stressful situations, you can see how these traits may put males and females at odds [18].

#### Generation

There are several different classes of generation styles, attitudes and preferences based on a person's age [18]. For purposes of this discussion we will divide these generations into four different categories: Veterans (1900-1945), Boomers (1946-1964), Generation X (1965-1980) and Generation Y (1981-1999).

Veterans are known at the "conformists". They are typically very loyal employees, ardent rule followers, they don't challenge the system or the organizational hierarchy and take pride and satisfaction in doing good work,

Baby Boomers are known as "the me" generation. They are loyal hard workers, high achievers, very competitive, self absorbed workaholics who pride themselves on individual accomplishment and personal goal achievement. They want to be the star of the show.

The Generation Xers are referred to as "the survivors". They are self reliant more independent thinkers who value friends and family over work, are more skeptical and less impressed by authority and title, and seek a positive work-life balance.

The Generation Yers are referred to as "the loved ones". They maintain strong family ties, are more comfortable with diversity, are more interested in informal team collaboration than traditional structure, are very confident, possess strong technology savvy and look at life as fun, fast and interactive.

Every generation emerges with a different set of values and experiences as compared to the generation before and after them. One set of values is not better than another, they're just different. These differences can be a real barrier to effective communication. In order to maximize communication and team collaboration efficiency, we need to be conscious of other people's values, styles and priorities and work with them to support a process that maximizes their talents and preferences in a way that results in the best possible outcomes.

#### Culture/ Ethnicity

The United States is one of the most ethnically and culturally diverse countries in the world and many clinicians come from a variety of cultural backgrounds.

Culture, ethnicity, traditions, family values, political and religious beliefs all mold our thoughts, perceptions and interactions with society [20-22]. These values affect ones perceptions, viewpoints and opinions which influence specific actions and reactions to different situations. These values reflect views related to power and authority, roles and responsibilities of men and women, approach to conflict and confrontation and physical and emotional self confidence.

In all interactions, cultural differences can exacerbate communication problems. For example, in some cultures, individuals do not like to be assertive or challenge opinions openly, as loss of face is considered. As a result, it is very difficult for some nurses to speak up if they see something wrong. They often will communicate their concern in very indirect ways, which can inhibit important information transfer. Culture barriers can be furthered hindered by nonverbal communication. For example, some cultures have particular ideas about physical space, the meaning of eye contact, specific facial expressions, touch, tone of voice, and nods of the head. Potential misinterpretation of intentions is a set up for ineffective communication, and is compounded in emergent or stressful situations.

#### Other Life Experiences

In addition to the value set, beliefs and perceptions molded by the factors discussed above, real time life experiences can set the mood for the day. Recent arguments, conflicts, illness, financial troubles, substance abuse or any other disturbance can affect the way we approach and interact with others and influence the final outcome.

#### Training

If one looks at the process involved in training a physician, you will note it encompasses the full gamut of emotions from early training insecurities and abuse to having to make critical life and death decisions under extremely stressful situations.

When one starts out in medical school everyone is very quick to let you know that you don't know anything and your opinions are not wanted or valued. Your job is to do all the "scut work" that no one else wants to do. At the bottom of the chain of command you're abused by staff and as a result feel very isolated and insecure. The training period requires long hours of work and dedication and the associated stress and fatigue doesn't help the situation. The training itself is focused on technical competency rather than personal interactions or leadership development skills. When you do reach a point of responsibility, you need to make immediate authoritative decisions which have a profound effect on patient care. This breeds a composite picture of dominating, autocratic, controlling behavior which lies on top of the insecurity, isolation and independency fostered by earlier conditions. What suffers is the lack of communication, people and team collaboration skills so important for patient care.

#### Stakeholder Interest

There are many participants in the health care delivery process, and they each have their own values and interests.

Nurses want to be respected. They want to feel like they're part of the health care team and are able to have input into the patient care process. They want to feel that their contributions make a difference in quality patient care. Other members of the health care team have similar values.

Physicians want to get the best possible outcome. They want staff to be responsive to their demands and they want to feel that the staff are trustworthy and competent to carry out all their duties.

## **ASSESSMENT, STRATEGIES AND RECOMMENDATIONS**

Given the multiplicity of interacting factors, there is no canned set of strategies or recommendations that can be applied as a cookie cutter solution to the problem. Individual solutions will depend upon the scope of the problem, the underlying culture, leadership and organizational dynamics, and the current status and success of programs already in place. The overall goal is (1) to prevent these types of episodes from occurring, (2) to address acute events in real time to avoid a compromised result, and (3) to develop policies that set appropriate standards of behavior with follow up procedures when these standards are not met. Recommended steps are listed in Table III.

## **Table III. Recommendations**

1.	Organiz	ational awareness		
	a.	Internal assessment		
	b.	Business case		
	c.	Motivation and accountability		
2.	Organiz	ational support and commitment		
	a.	Culture		
	b.	Leadership		
	c.	Champions		
3.	Meeting	ings, discussion groups, education		
	a.	Informal get togethers		
	b.	Committees		
	c.	Department presentations		
	d.	Focus groups		
	e.	Training programs		
4.	Education	onal training seminars and workshops		
	a.	Diversity management		
	b.	Anger management		
	c.	Conflict management		
5.	Communication training			
	a.	Phone etiquette		
	b.	Scripting		
	c.	SBAR		
6.	Team tr	÷		
	a.	Crew resource management		
	b.	Pit crew management		
	c.	Lean process/ Six Sigma		
7.	Compet	ency training		
	a.	Technical competency		
	b.	Language competency		
8.	Interven	tion strategies		
	a.	Assertiveness training		
	b.	6		
	c.	Code White		
9.	Behavio	or policies and procedures		
	a.	Zero tolerance policy		
	b.	Code of Conduct policy (credentialing)		
	c.	Disruptive Behavior policy		
10.	Reporti	ng mechanisms		
	a.	Incident Reporting		
	b.	Incident evaluation		
	c.	Follow up plan		

#### 1. Organizational Awareness

The first step in the process is to assess the current state of affairs. One way to assess the situation is to review complaints and incident reports and/ or listen to the hallway or coffee lounge gossip. Another way to assess the situation is to force the issue with an internal survey. We recommend a combination of these approaches. The added value of the dedicated survey is that it not only gives you an objective assessment of what your staff sees and feels at the organization, it also allows you to pinpoint areas of concern. Our survey tool specifically addresses the frequency of disruptive behaviors by discipline, medical specialty and service unit, the impact of disruptive behaviors on behavioral factors affecting levels of stress, concentration, communication and information transfer, and the linkages of disruptive behavior to adverse events, medical errors, patient safety, quality, mortality, and staff satisfaction. The survey also adds a comments section where participants can go into more detail about their observations, experiences and surrounding issues (see example Table I). As mentioned earlier, these comments provide significant insight into the nature and extend of the problem. The two most important caveats are (1) the assurance that all results will be held confidential and (2) that actions will be taken and follow up feed back is provided.

Once the survey results are accumulated, they are analyzed and compared to other results in the database. The next step is to share the results with the staff at the hospital. The goal is to motivate action and responsibility by making a strong business case that this is an important issue that has series implications and repercussions. The business case developed around phase one of our research focused on the impact of disruptive behaviors on nurse satisfaction, retention and turnover and its contribution to the hospital nursing shortage. This may or may not have had a big impact on the physician audience. Phase two of the research focused on the impact of disruptive behaviors on patient care and this affects everyone in the health service industry. The business case for physicians is that poor patient outcomes may occur because of disruptive events that lead to communication mishaps and /or serious errors resulting in a potentially preventable adverse event and that you have accountability in this process.

#### 2. Organization Commitment and Support

The organization itself sets the stage for expectations in regard to staff roles, responsibilities, behaviors and performance outcomes. Culture is defined as the integrated pattern of human knowledge, belief, and behavior that depends upon man's capacity for learning and transmitting knowledge to succeeding generations [24]. The organization needs to set the priorities by being committed to providing and supporting a culture that demands and reinforces appropriate professional behavioral standards. This needs to be is a top down bottom up commitment. Administrative leaders, Clinical Leaders, Directors, Line Managers, Hospital Employees and Physicians all have responsibilities for supporting this standard and everyone needs to be held accountable for adhering to these cultural and behavioral expectations as a way of doing business.

One of the key successes on the clinical side has been the enthusiasm and advocacy of a clinical champion. For physicians, the clinical champion could be a salaried physician executive (Chief Medical officer (CMO), Vice President of Medical Affairs (VPMA) or

Medical Director), the Chief of Staff, or a physician who so believes in the importance of enforcing appropriate behavioral standards that he or she drives the process on their own. From the nursing side the champion could be the Chief Nursing Officer (CNO), the Vice President of Patient Care Services, a Department Director, or any other nursing staff advocate who is passionate about this issue. For a champion to be successful they must be respected, have a good rapport and reputation with the medical and nursing staff, and be eloquent and persuasive in their actions. The clinical champion usually works in conjunction with the administrative and clinical leadership staff. Some organizations have developed programs that are part of their patient safety, quality, risk management and accelerated performance initiatives.

#### 3. Meetings and Discussions

One of the key principles in issue resolution is to have individuals address issues and concerns with the people involved. The earlier on in the evolutionary course of the disturbance, the greater the opportunity for clearing up expectations, assumptions and misperceptions that led to the problem.

One simple step is to encourage contact. Physicians can do a better job in making efforts to establish a positive nurse- physician relationship. Our research has shown us that while only 2-3% of physicians are overtly disruptive (this 2-3% can have profound effects on the entire organization), another 30-50% have poor communication and social interaction skills. In the hospital environment, physicians appear to often be stressed, hurried and pre-occupied, and want to get in and get out so they can get on to the next thing. If the physician would just take a step back and smile at the nurse or ask a question about how their patient was doing, this would do a lot in improving nurse- physician relations.

Another way to get physicians and nurses together is to set up task forces or committees where they work on mutually important topics. Some organizations have even set up department or town hall meetings where nurses and physicians can address issues of concern in a collaborative manner.

Some organizations have developed dedicated committees to address nurse physician relations. They are responsible for setting policy and addressing issues as they rise.

The topic of disruptive behaviors and its impact on organizational dynamics and patient care need to be discussed as part of the ongoing staff educational process. Topics can be presented at general sessions or ground rounds, at specific department meetings or at off- site events. More in-depth discussions can be conducted as specific focus groups, particularly in areas where problems are more evident. As mentioned previously, there is a growing need to introduce these discussions at the medical school and nursing school level to better prepare and equip students to be knowledgeable about the potential consequences of disruptive behaviors and impaired communication.

#### Educational Seminars and Workshops

In order to address some of the causative factors discussed previously, organizations need to provide more comprehensive educational seminars, workshops and training programs to improve staff skill sets so they can better address these issues. These programs could include a number of different topics such as diversity training, sensitivity training, anger management, conflict management, stress management, or any other topic appropriate for the situation.

#### 5. Communication Training

Special efforts need to be made to improve staff communication skills. We don't think that anyone intentionally plans on being a poor communicator, but some are better than others. Communication is a two way street. The initiator has a specific intent or message that requires a response. The receiver processes the request and gives a response. As discussed previously, the intent and response may be influenced by emotions and attitudes related to events surrounding the current situation and the values already established via gender, age, culture, personality, and life experiences that influence thoughts and actions. Perceptions override reality. It is also important to remember that body language, verbal tone and presentation style have more of an influence on messaging than the actual content of the message itself.

There are a series of tools that have been designed to improve communication efficiency. Equally important are tools that specifically address language capacity and these will be discussed later on under competency.

Scripting is a very efficient way of organizing priorities into a series of straightforward questions designed to achieve desired objectives. One example would be a list of questions originally put together by the Quint Studer Group designed to prepare nurses when making a call to a physician [24]. The basic tenets are to identify who you are, identify the patient, state the problem, be familiar with the patient's chart including a thorough review of the orders, testing results and progress notes, and treat the physician with respect.

Another recently introduced tool that has become very popular is the SBAR tool ((Situation, Background, Assessment, and Recommendation). This tool has been successfully implemented at Kaiser and many other facilities where nurses, physicians and other staff members are trained simultaneously on tool purpose and utility. The tool provides a clear concise consistent process to exchange relevant information needed to make a decision about patient care [26,27].

#### 6. Team Training

Another new innovation recently introduced into healthcare under the umbrella of patient safety and the importance of team collaboration is the adoption of crew pit management techniques as used in the aviation industry. The focus of these types of team training programs is to highlight the importance of understanding your role and everyone else's roles and responsibilities, increasing awareness and anticipation of upcoming events, trusting your team, avoiding conflict and confusion, and speaking up if something appears to be wrong or not proceeding as planned [28-30]. A recent extension of the cockpit crew management technique has been the adoption of the pit crew techniques used by the Ferrari team in the NASCAR races [31].

Many organizations are embracing Lean process design and Six Sigma techniques in an effort to improve efficiency, reduce errors and remove waste in their processes and procedures as a way to accelerate organizational performance. This requires both system redesign and human factor re-engineering which calls for meeting compliance standards around human behaviors.

#### 7. Competency Training

Assuring competency, trust and respect is a key factor in communication efficiency and effectiveness. On one side physicians need to feel confident that nursing and other supporting staff are well trained, knowledgeable and technically competent in handling and carrying out appropriate patent care directives. Likewise, nurses and other clinical staff want to feel that the physician is able and competent in carrying out their duties. Many of these knowledge and technical competencies are mandated through accreditation, training verification and credentialing requirements.

A second area of competency is what we call communication competency. A person could have great knowledge and technical skills but through either a language barrier or an ineffective communication style, there is a risk to communication efficiency and effectiveness. In the previous section we addressed the use of certain tool kits designed to improve communication proficiency. In regard to language, our experience has shown us that in many of the organizations who participated in the survey, in a large percentage of both the clinical (particularly nursing) and medical staff English is a second language. In this regard, some of the organizations have provided specific courses on communication applications and techniques for those individuals where English is not the primary language [32].

#### 8. Intervention Strategies

One of the biggest concerns is to get people to speak up in real time during the time in which a disruptive event is in the process of unfolding. As mentioned before, many individuals are reluctant to speak up because of issues related to personality, gender, culture, organizational hierarchy, reputation, personal experience, or other influencing factors. Think of the case of a new entry female nurse from the orient who is shy and respectful of position faced with an angry domineering foreign trained male physician who is yelling and screaming about an issue of concern.

There was a recent white paper published called "Silence Kills" which highlights the problems that occur when people don't speak up when they see a problem developing [33]. In their book titled *Crucial Conversations* the authors go into detail about the problems that occur from not speaking up and use role play examples to provide techniques that instill confidence on the why, how and when to speak up in an appropriate manner [34]. Another technique is to provide basic assertiveness training classes to all involved staff members.

Some organizations have utilized a briefing/ debriefing technique in an effort to either prevent an event from occurring by upfront discussion, or to address an event that did occur in a post-briefing session. For scheduled encounters the pre-briefing session can discuss objectives, clarify roles and responsibilities and can bring forward any issues of relevance prior to the scheduled event such as surgery, a procedure or group meeting. If someone is having a rough day, they can bring this out up front so people will understand and be more accepting or tolerant of the situation. The post event briefing can do the same. In crisis time some organizations have implemented a "Code White" policy. When a disruptive event is occurring, a call goes out to a team of individuals who then come to the scene to help intervene [35].

#### 9. Behavior Policies and Procedures

Specific policies and procedures need to be developed to support behavioral standards of care. These policies come under the umbrella of Zero Tolerance, Code of Conduct, Code of Ethics or Standard Behavior policies which should outline appropriate behavior standards, expectations and ramifications if these standards are not followed. These policies and procedures need to be applied organization wide. Rather than one policy for physicians, another for nurses, and another for other employees, these policies need to be consistently and equitably applied across all levels of the organization. Staff employees and physicians must agree to abide by these policies as a condition of employment or acceptance of a new position or appointment and/or as part of the privileging and credentialing process.

There also needs to be a distinct Disruptive Behavior policy which includes specific criteria and procedures for handling disruptive individuals. There should a specific committee or group of individuals with representation from all involved disciplines (medical staff, nursing, administration, human resources other ...) who participate in a consistently applied process that evaluates issues and makes recommendations for appropriate action, follow up and feed back.

#### 10. Reporting

Incident reporting is a vital issue. Our research has shown us that more than 50% of the respondents stated that they were reluctant to report in fear of concerns about confidentiality, retaliation, the fact that no one listens, no actions are taken and/ or they never get any feedback about recommendations or resolution.

In order to address these issues, the organization must support a confidential non-punitive reporting policy. Rather than depending upon specific incident reports, written or verbal reports to CMOs, CNOs, supervisors, directors or mangers or other formal or informal documentation streams, it would be preferable to have a consistent process for handling and responding to complaints.

Some organizations have implemented a process where they have one designated committee responsible for receiving and addressing all relevant complaints. The committee has a multidisciplinary composition that as a group (rather than an individual) is responsible for investigating each individual complaint and coming up with an appropriate action and follow up plan. The advantages of using this type of reporting system is that the process is straight forward, all inputs and outputs are held confidential, there is consistency and equity in the review process and decisions are made by a group process.

## CONCLUSIONS

Disruptive behaviors are a common occurrence across all industries, but in health care the risks are more acute as it can adversely affect patient outcomes of care. Until recently, health care as an industry has been reluctant to deal with the issue as they were politically and financially reluctant to confront a high profile physician who voluntarily brings his or her patients to the hospital and is a main source of hospital revenue. Recent issues around workforce shortages, workplace safety, liability and bad patient outcomes have brought new fuel to the system where organizations can no longer afford to look the other way.

Of particular concern is the effect of disruptive relationships on increasing feelings of frustration, anger and antagonism, and its negative impact on willingness to communicate, collaborate and exchange information necessary for patient care. Disruptive relationships can heighten stress, increase distraction and impede concentration, all of which results in a higher likelihood of medical errors. Overall there is a strong linkage between disruptive relationships and the occurrence of adverse events and compromises in patient safety, quality and clinical outcomes of care.

Disruptive behaviors occur for a multitude of reasons. Some of the behaviors are based on deep seated values influenced by gender, age, ethnicity, religious beliefs, family values, training, personal experiences and personality. Some disruptive behaviors are provoked because of real time stresses that affect emotions and disposition.

In developing a program to address these behaviors the focus is to try to prevent such outbursts from occurring, minimizing the impact if they do occur, and taking action on any events which have significant consequences. While it may be difficult to change personalities or personal values, you can change situational awareness, perceptions and reactions through appropriate training and education.

In order to achieve desired results the organization must be committed to a zero tolerance policy, develop and support a professional behavior policy with the commitment to enforce desired standards consistently across all disciplines, provide education and appropriate support services and be prepared to take appropriate action as necessary.

#### REFERENCES

- [1] Institute of Medicine Report To Err is Human: Building a Safer Health System. Washington (DC): National Academy Press; 2000.
- [2] Institute of Medicine Report "Crossing the Quality Chasm: A New Health System or the 21<sup>st</sup> Century" Washington (DC): National Academy Press; March 2001
- [3] Institute of Medicine Report "Preventing Medication Errors" Washington (DC): National Academy Press; July 2006
- [4] "The Joint Commission Guide to Improving Staff Communication" Joint Commission Resources Oakbrook Terrace, IL 2005
- [5] IHI 5 Million Lives Saved Campaign www.IHI.org
- [6] Longo, D., Hewett, J., Ge, B., Schubert, S. "The Long Road to Patient Safety: A Status Report on Patient Safety Systems" *JAMA Vol.294* No.22 December 14, 2005 p. 2858-2865

- [7] Rosenstein, A. "The Impact of Nurse-Physician Relationships on Nurse Satisfaction and Retention" *American Journal of Nursing, Vol.102* No.6, June 2002, p. 26-34
- [8] Rosenstein, A.; Lauve, R., Russell. H. "Disruptive Physician Behavior Contributes to Nursing Shortage" The Physician Executive, November-December 2002, p. 8-11
- [9] Rosenstein, A., O'Daniel, M. "Disruptive Behavior and Clinical Outcomes: Perceptions of Nurses and Physicians" *American Journal of Nursing Publication Vol. 105* No.1 January 2005 p.54-64
- [10] Rosenstein, A., O'Daniel, M. "Impact and Implications of Disruptive Behavior in the Peri-Operative Arena" *Journal of American College of Surgery Vol.203* No.1 p.96-105 July 2006
- [11] Needleman, J., Buerhaus, P, Stewart, M., Zelevinsky, K. and Mattke, S. "Nurse Staffing In Hospitals: Is There A Business Case For Quality?" *Health Affairs, Vol.25* No.1 p. 204-211
- [12] Baggs, J., Schmitt, M., Mushlin, A., Eldredge, D. et. al. "Association Between Nurse-Physician Collaboration and Patient Outcomes in Three Intensive Care Units" *Critical Care Medicine Vol.27* No.9 September 1999 p.1991-1997
- [13] Knaus, W., Draper, E., Wagner, D., Zimmerman, J. "An Evaluation of Outcomes from Intensive Care in Major Medical Centers" *Annals of Internal Medicine Vol104* No.3 March 1986 p.410-418
- [14] Boyle, d., Kochinda, C. Enhancing Collaborative Communication of Nurse and Physician Leadership in Two Intensive Care Units: *JONA Vol.34* No. 2 p.60-70 Feb. 2004.
- [15] Reason, J. "Human Error: Models and Management" British Medical Journal Vol.320 March 18, 2000
- [16] Allesandra, T., O'Connor, M. The Platinum Rule: Discover the Four Basic Business Personalities and How They Can Lead You to Success Warner Books New York 1996
- [17] Gray, J. Men and Venus in the Workplace Harper Collins Publishers New York 2004
- [18] Zelek, B., Phillips, S "Gender and Power: Nurses and Doctors in Canada International Journal for Equity in Health Vol.2 No.1 February 11, 2003 p.1
- [19] Lana caster, L., Stillman, D. When Generations Collide Harper Collins Publishers New York 2002
- [20] Peterson, Brooks. *Cultural Intelligence: A Guide to Working With People from Other Cultures*. Boston: Intercultural Press, 2004.
- [21] Taylor SL, Lurie N. The Role of Culturally Competent Communication in Reducing Ethnic and Racial Healthcare Disparities. *American Journal of Managed Care*. 2004 Sep; 10 Spec No:SP1-4
- [22] Early, P. Christopher and Ang, Soon. *Cultural Intelligence: Individual Interactions Across Cultures*. Stanford University Press, 2003
- [23] Kohls, L. Robert and Knight, John M. A Cross-Cultural training Handbook. 2nd Edition. Boston Intercultural Press, 1994
- [24] Webster's Ninth New Collegiate Dictionary Merrian- Webster Inc. Springfield, Massachusetts 19889
- [25] Personal communication The Quint Studer Group Gulf Breeze, FL
- [26] Leonard, M., Graham, S., Bonacum, D. "The Human Factor: The Critical Importance of Effective Teamwork and Communication in Providing Safe Care" *Qual Saf Health Care Vol.13* Suppl.1 2004 p.85-90

- [27] Sutcliffe, K., Lewton, E., Rosenthal, M. "Communication Failures: An Insidious Contributor to Medical Mishaps" Academic Medicine Vol.79 No.2 p.186-194 February 2004
- [28] Grogan, E., Stiles, R., France, D., Speroff, T. et. Al. "The Impact of Aviation- Based Teamwork Training on the Attitudes of Health- Care Professionals" *The American College of Surgeons Vol.199* No.6 December 2004 p.843-848
- [29] Leming-Lee, S., France, D., Feistritzer, N., Kuntz, A. et. al. "Crew Resource Management in Perioperative Services: Navigating the Implementation Road Map" *JCOM Vol.12* No.7 July 2005 p.353-358
- [30] Sexton JB, Thomas EJ, Helmreich RL. "Error, Stress, and Teamwork in Medicine and Aviation: Cross Sectional Surveys. *BMJ Vol.320* March 2000 p.745–9.
- [31] Naik, G. "A Hospital Races to Learn Methods of Ferrari Pit Stop" *WSJ* November 14, 2006 p.D1-D2.
- [32] Personal communication Peachy Hains RN, Christopher Ng MD, Cedars-Sinai Medical Center Los Angeles, California
- [33] "Silence Kills" American Association of Critical Care Nurses www.silencekills.com VitalSmarts Provo, UT 2005
- [34] Patterson, K., Grenny, J., McMillan, R., Switzler, A. Crucial Conversations: Tools for Talking When Stakes are High McGraw – Hill New York 2002
- [35] "Building the Nurse- Physician Partnership: Restoring Mutual Trust, Establishing Clinical Collaboration" The Health Care Advisory Board Washington, D.C. 2005

Chapter 3

## TIME AND RISK IN HEALTH-RELATED DECISIONS\*

## Monica Ortendahl<sup> $\dagger$ </sup>

Stanford University School of Medicine, Department of Medicine, Division of Immunology and Rheumatology, California 94304, USA

## ABSTRACT

Classical decision theory, which deals with situations in which a decision-maker can choose between different actions, is based on the effect of probability of outcome and value of outcome on judgments and decisions. The importance of the time factor has been ignored in decision theories, and in preventive procedures and clinical decision-making as well the impact of time has largely been neglected. Most work on inter-temporal choice is based on a trade-off between two outcomes of different values over different periods. Although risk is known generally to have an important impact on judgments and decisions, the effects of time have not been studied to the same extent. However, the concepts of time and risk are both important to health-related decisions. The intricate relationships between time and risk imply that subjects are risk aversive, with the individual's risk aversion intensifying the sooner the outcomes of choices will occur in the future. Time-related aspects of risk have a potentially large impact on health-related decisions and responses to forecasts of morbidity and mortality from disease in the future.

Keywords: time, risk, aversion, personal, general, morbidity, mortality

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence concerning this article should be addressed to: Monica Ortendahl, Malma Backe 3 H, 756 47 Uppsala, Sweden. Phone home (Sweden): 46 18 302549; E-mail: monicaortendahl@hotmail.com.

## **INTRODUCTION**

The impact of time on many decisions has increasingly been recognized as an important factor in the health field as it has become apparent that many medical situations involve the exchange of present-day costs for future benefits. Traditional decision-analytic paradigms weigh the probability of outcomes and the value of outcomes. Such analyses are incomplete if they do not consider the time of the outcome as well.

The evaluation of routine preventive care, population screening programs, and therapies for chronic disease often works toward remote goals and includes a trade-off between costs in the present and benefits in the future (Ariely & Zakay, 2001; Walker & Kumaranayake, 2002). Such programs attempt to change behavior and foster investment in the future instead of consumption in the present. However, the benefits often occur so far in the future that they may seem to be of little value to the individual relative to the immediate costs.

Inter-temporal models, that is the choice between something now and something later, generally assume that people are impatient. The concept of temporal discounting signifies that the value of a future outcome diminishes because of the time factor; the diminishing of value over time is called positive discounting. If there is a preference for living for the present rather than investing in the future, that preference might not optimize the outcome or maximize the utility of a treatment program (Chapman, Brewer, Coups, Brownlee, Leventhal & Leventhal, 2001).

Most existing work on inter-temporal choice has been restricted largely to a trade-off between two outcomes of varying values over different periods. The role of risk has not been studied to the same extent, although it is known that risk generally has a large impact on judgments and decisions made either by the individual or by society. A decision will often have to be made about *when*, rather than *whether*, to undertake a risk-reducing behavior. A similar judgment is also required about when a preventive procedure should be adopted or a clinical decision made. Introducing features of time and risk provides a more complex picture that should influence health policy and clinical judgments.

## THE CONCEPT OF RISK

The concept of "risk attitude" was introduced to describe inter-temporal health choices (Gafni & Torrance, 1984). Risk attitude incorporates three effects, as follows:

- A pure time preference effect; this represents a preference structure in which the individual prefers his health gains sooner rather than later.
- A gambling effect; this depends on the individual's attitude toward risk (also called uncertainty). Uncertainty in this context refers to the possibility that she will not receive the benefit of the promised future pleasure.
- A quantity effect; this refers to a diminishing marginal utility of consumption. A quantity effect is a process in which an individual who is offered a choice between various durations of life extension in a chronic dysfunctional health state gives each succeeding year less value because he or she is simply growing weary or fed up with the situation.

The concept of risk encompasses several characteristics. The discount rate for a particular risk is one of the features discussed in this. Issues related to discount rate are risk in money and health, risk aversion, and personal versus general risk. Discount rates differ for different types of risk, and the varying results obtained for the discount rate might be explained in part by a difference in the value of risk reduction. In the past, empirical work on risk reduction has focused mostly on risk in the present or has used a-temporal models. However, proposed time discounting utility-based theories of the value of life have been including, to a greater extent, both a-temporal and inter-temporal life-cycle models in which the timing of risks is relevant (Roelofsma & Van der Plight, 2001).

## **MONEY AND HEALTH**

Previous research has mostly focused on time discounting and monetary values. However, health problems, such as those due to addictive behavior and the difficulties of implementing preventive measures, have become increasingly important. The acceptance of a current cost for future benefits could be regarded as an investment. This is an opinion initially developed by Becker (1964) with the theory of investments in human capital, and with the application of this theory specifically to health. This gives a framework for explaining differences in willingness or ability to make health investments due to different time perspectives (Drummond, 2002).

Time-discounting effects are often discrepant and cannot easily be generalized between the two fields of money and health (West, McNabb, Thompson, Sheldon & Evans, 2003). The positive discount rate generally is higher for money than for health, and there are problems with treating death and other health effects as monetary equivalents (Groot, Van Den Brink & Plug, 2004). The normative rationale for using the same discount rate for health and money are that they are fungible. Chapman (2002) speculates that this condition may be less easily met for individual decisions than for societal decision, suggesting that health and monetary discount rates may be similar for social choices. Moreover, people are more risk aversive with regard to choices involving years of life than they are to those involving monetary outcomes (Hoel, 2003).

## **RISK AVERSION**

Risk aversion implies that one prefers the expected value, whereas risk seeking implies that one prefers a gambling situation with an uncertain outcome of choice. Divergent explanations for a risk-aversive attitude have been offered (Cher, Miyamoto & Lenert, 1997). One explanation refers to the commonly observed non-linear concave shape of the utility function. Another is probability weighting, with a preference for \$50 for sure instead of a 50-percent chance of \$100. A third is the relationship between risk attitude and temporal discounting, where, due to non-linear utility or temporal discounting, later years are valued less than earlier ones.

Individuals are usually risk aversive (Rosen, Tsai & Downs, 2003; Wakkar, 2004). With health outcomes, such as years of life gained, risk aversion is increased if benefits are non-

transferable. In the judgment process, the non-tradable risk must be accounted for in some way. Different measures of health can be used (Salomon & Murray, 2004) that incorporate risk aversion or risk preference with regard to remaining life years or quality-adjusted life years. For outcomes that are calculated without risk, preventive programs will be overvalued in an absolute sense. However, compared with the more risky alternative, preventive programs might be undervalued. Because some programs may be undervalued due to a net risk reduction, this factor should be taken into consideration when preventive projects are evaluated.

When the outcome of a choice is in the near future, subjects exhibit greater risk aversion than when the outcomes are in the remote future (Langford, Marris, McDonald, Goldstein, Rasbash & O'Riordan, 1999). For smokers, the level of perceived risk has been found to decline over time for those who failed in their attempt to quit smoking (Gibbons, McGovern & Lando, 1991). In order to sustain motivation to stop smoking, therefore, educational efforts need to counteract the tendency of risk perception to decline.

Risk willingness increases with disengagement, where a smaller value for the consequences leads to decisions that more closely approximate expected value (Dixon, 2005). People will take bigger risks in connection with smaller consequences because in fact there isn't very much at risk. Moreover, Bayesian studies concerned with examining information processing in judgment have shown that conservatism in judgments increases for rare events (Richardson, Thomson, Best & Elliott, 2004). Conservatism in judgments is consistent with an increased tendency to take risks when outcomes of the decisions are delayed. The salience of both negative and positive consequences diminishes with increased distance to the goal, though in fact the saliency decreases more rapidly for negative outcomes (Ortendahl, 2005).

## **PERSONAL AND GENERAL RISK**

Individuals judge that their personal risks from adverse health behaviors are smaller than the same risks for people in general; the difference between personal and general risk is related to perceived control (Sjoberg, 2001). When people believe they have control over a risk-filled situation, they judge their personal risk to be smaller than general risks. Presumably, this is because people do not consider others to be equally capable or willing to protect themselves from risk.

Time discounting may be in value, in probability, or in both. Regardless of whether or not expected value is discounted, probability discounting is used if the risk is perceived to be controllable (Hakes & Viscusi, 1997). For perceived controllable risks, probability represents a more general mechanism than value discounting. People's tendency to discount future consequences might be counteracted, then, by lowering the amount of perceived control.

Paradoxically, perception of decreased control may lead to greater exercise of control. If an educator or clinician emphasizes factors which oppose the idea that temporal delay will result in increased control, the discount rate will be lowered. For example, with smokingrelated behaviors, the possibilities for future corrective actions are limited because of the detrimental effects of smoking; some if its adverse effects, such as emphysema, are irreversible. Thus, the relationship between risk perception and time horizons is important. If risk-aversive persons moved toward positive time preferences and risk-taking persons toward negative time preferences, then their judgments would more closely approximate expected value. However, this area has not been thoroughly and systematically investigated.

## **IMPLICATIONS IN PREVENTION**

In risk perception and in insurance policies, risk is composed of the two concepts of probability and consequence. An additive combination of these two factors, probability and consequence, would imply that they do not interact and that each component could be studied separately for its influence on risk perception. While risk has mostly been defined as consequences, it could also be defined as seriousness of risk (i.e., risk level). Thus the choice of definition has important implications at the societal level for communications about risk and for preventive programs. This applies to both primary and secondary prevention.

Motivation for risk reduction by individuals is only weakly related to perceived level of risk (Sjoberg, 2000); perceived consequences appear to be much more important for motivation. Consequences have more influence on insurance value and on the importance given to prevention in public policy than risk level does.

As noted above, there is a need to distinguish between personal and public risks. In issues related to risk policy personal risks generally are most important for technical risks, while the general risk is most important for matters of lifestyle (Sjoberg, 2003).

An additional distinction could be made between judging the value of insurance and judging the importance of preventive actions. The author of this chapter sees the time aspect as more salient for preventive procedures, because a longer time horizon is involved. Individuals determine the value of insurance by personal consequences and by perceived competence to prevent the damage from occurring. The importance of preventing damage has been found to be determined by personal consequences and by perceived worry (Sjoberg, 1998).

These are all intricate relationships, which hypothetically could relate to a dimension of risk aversiveness and risk-seeking behavior and to a preference for gains rather than losses (Ortendahl & Fries, 2002; Tversky & Kahneman, 1991). If we may indulge in speculation, the time variable should have a greater impact on personal risks than on public risks.

Overall, the likelihood of a specific adverse outcome is a parameter that affects the estimate of future public risk and its consequences (Dellve, Lagerstrom & Hagberg, 2003). Another parameter is the fact that temporal distribution of risk is not homogenous throughout the working life of particular individuals.

Moreover, the degree of severity may not be the same for everyone who experiences the event. Occupational back pain, for example, may be relatively mild in some persons but disabling in others (Lahiri, Markkanen & Levenstein, 2005). Specific individual factors modify the risk for a specific person in a specific job, and person-specific modifiers are likely to be distributed differently in time. Discounting for future events could be based on an event in the future having less importance than one in the present, and where the adverse outcome may have different weights.

There is also a trade-off in the timing of risk. Time is a key element when risks with different time profiles compete (Cropper, Aydede & Portney, 1994). There will be a trade-off in the timing of risk, rather than a trade-off of risk for money, related to the present value of a

risk reduction that will take place in the future. Risk has been studied in many settings, and evidence suggests that risk-risk trade-offs are more stable than responses to risk-money choices (Krupnick & Cropper, 1992).

For programs that reduce long-term health risks, the market view involves placing less emphasis on reducing long-term health risks in favor of greater emphasis on programs that achieve more rapid reductions in health risks (Gravelle & Smith, 2001). This conclusion is valid, even if those programs may save fewer statistical lives in the long run than an alternate program that produces less immediate results, if time discounting is present. For some risks, the median discount rate may be close to the real market discount rate; for other risks it may not. Consequently, we can reject the hypothesis that the discount rates for all risks are equal.

How resources are allocated over time has an impact on mortality and morbidity, where morbidity and mortality have different costs (Johannesson & Meltzer, 1998). If your risk of death declines now, you gain some lifetime (Sogaard & Gyrd-Hansen, 1998). But when does the gain occur? Does it occur at the end of your expected life-span or before risk reduction? Or is it even occurring now, at the very moment of risk reduction?

The necessity of deferred gratification in health decisions is a central issue, and the expected benefits are reductions in the probability of morbidity and mortality from diseases in the future. Fries (2003) proposed the Compression of Morbidity paradigm, which emphasizes delaying the onset of morbidity and thereby shortening its duration, with the intent to reduce lifetime illness and morbidity. This paradigm envisions reduction of lifetime infirmity and medical-care costs by compressing the period of morbidity between an increasing average age at onset of disability and the average age of death. In senior populations, fractures associated with osteoporosis are a major cause of morbidity; preventive interventions aimed at reducing the age-specific incidence of fracture are crucial in reducing the morbidity resulting from these fractures. The effect of preventive interventions is likely to postpone the age of onset of morbidity.

Another time-risk interaction in this context is the optimum timing of different preventive actions. If the probability or risk for certain outcomes is introduced, the relationship between morbidity and mortality might therefore be conceptually extended both for public policy makers and for individuals who receive the information or care. The fact that the years added by preventive measures may not be purely healthy ones suggests that the quantitative aspect of these years is important.

The relationship between probability or risk for certain outcomes and morbidity and mortality is important, since the majority of deaths under age 65, including those from premature diseases, injuries and other types of morbidity, are preventable. "Avoidable deaths" have been mapped to investigate the impact of preventive procedures (James, Manuel & Mao, 2006). Lists of causes of death have been compiled that, in certain age groups, serve as indicators of the outcome of medical care intervention.

## **IMPLICATIONS IN CLINICAL DECISIONS**

Time-related problems occur both with diagnostic decisions and with treatment decisions. Is there a difference between patients' and doctors' estimate of time aspects and risk aspects? Patients might show a higher estimation of risk with a more conservative attitude, since decision outcomes could be more important to them. It is one's own life that is affected, and consequently patients might be more risk aversive. However, as noted above, people underestimate their personal risks.

Recently decision research related to risk has increasingly stressed the emotional and affective components in risk perceptions (Slovic, Finucane, Peters & MacGregor, 2004; Loewenstein, Weber, Hsee & Welch, 2001). In formal normative risk analysis, based on a calculation of expected utility, emotional and affective components have been regarded as irrational.

Emotional reactions at risk-related decisions generally are different from cognitive judgments of risks (Jardine, Hrudey, Shortreed, Craig, Krewski, Furgal & McColl, 2003). An increased competence when handling risks includes emotional and affective components such as limited consciousness, erroneous presentation of risk and optimistic bias (Buchanan & O'Connell, 2006).

There could also be a discrepancy in time perspective between the doctor and the patient, both objectively and subjectively. That discrepancy might affect the patient's level of compliance with prescriptions and treatments. Therefore, features of risk and time should increasingly be stressed in shared decision-making where health-care decisions are based on mutual agreement (Feldman, Chen, Hu & Fleischer, 2002; Montgomery & Fahey, 2001).

It is clear that beliefs vary about risk modification and preference for risk counseling (Ryan & Skinner, 1999). In first-degree relatives of breast cancer patients who have a potential genetic risk for developing breast cancer, there was a misunderstanding about risk. Because they confused risk factors with causes, the relatives, if they knew a breast cancer victim without risk factors or with protective factors, discounted the scientific validity of the risk-factor information they were given. Most of the relatives thought that lifestyle factors contributed to level of personal risk, believing that risk could be reduced by lifestyle modifications.

According to the present author, the process of attaining shared goals in health-related decisions based on time and risk aspects should be taken into greater account in clinical and health-policy decision-making. One issue, then, is how clinical inferences are generally arrived at when a judgment is being made about future health. Much clinical inference is characterized by backward reasoning, in that diagnosticians often attempt to link observed effects to prior causes (Kempainen, Migeon & Wolf, 2003). In contrast to this post hoc explanation, statistical prediction entails forward reasoning, because it is concerned with forecasting future outcomes given observed information. Clinical inference utilizes information from prior periods to make a statement about today. Prediction uses data to reason forward about the future. This can lead to different results in decision-making. The time factor may thus have an influence on future health in several respects.

Many people tend to overestimate how much they know, even about the easiest knowledge tasks. Overconfidence (i.e., more certainty than circumstances warrant) leads to overestimating the importance of occurrences that confirm one's hypothesis (Klayman, Soll, Gonzales-Vallejo & Barlas, 1999). This impedes learning from environmental feedback and hence results in deleterious effects on future predictions. In many decision settings, inexperienced practitioners and even naive laboratory subjects perform as well (or as poorly) as performers with more experience (Lewis, Robinson & Wilkinson, 2003).

Clinical inference tends to consider error as a nuisance variable. The statistical approach, on the other hand, accepts error as inevitable, and in so doing makes fewer errors in

prediction for periods extending over a relatively long time. Studies of clinical decision are often based only on the factors of value and probability. Clinical decision-making based on time and the distribution of risk over time provides another dimension. Therefore, timerelated aspects of health should be taken into consideration for clinical inference and realistic decision analysis.

## CONCLUSION

The issues raised in this chapter emphasize the importance of time-related problems in health-care work. The chapter documents an intriguing relationship between the time dimension and the risk dimension and suggests that this relationship has applications to preventive medicine and clinical decision-making. Introducing features of risk like the impact of risk aversion, and the difference between personal and general risk, yields a more intricate and also more complex description of variables that influence decisions.

Framing of preventive health messages may be important. Due to the limited tradability of health, when eliciting its time preference rate, it has been concluded (Ortendahl & Fries, 2005) that inter-temporal choices should be framed in such a way as to resemble as closely as possible those choices facing health planners and decision makers. The desire for behavior change in preventive health programs would, therefore, benefit from considering the usefulness of psychological factors in the models.

Time considerations are salient for individual as well as societal choices that impact longterm health outcomes. Time periods for health programs are usually defined on a societal basis, and tension may arise between risk and time periods for the individual and for society. If tension between individual and societal time aspects is low, health outcomes may be improved (Ortendahl & Fries, 2006).

Many choices in health, at both an individual and societal level, involve decisions with a trade-off between something now and something later. The evaluation of health risk reduction, routine preventive care, and population screening programs includes health measures that improve future health. Inter-temporal choices, which are involved or implied in many medical situations on a societal level, such as preventive programs on smoking, depend in part on the exchange of present-day costs for future benefits. However, for the individual, the value of not smoking vs. smoking is small, because the acquisition of the benefits of the former (or the disadvantages of the latter) takes place in the remote future. A distinction has to be made, therefore, between the personal and the political in the relationship between exchange and power.

One could speculate, from the perspective of political science, about whether there is what could be termed a common good. One issue here is how individuals view distributive justice with fair and equitable distribution of limited resources. Time profiles might be supposed to vary with reference to people's age, gender, creed, and ethnic group. I argue therefore that health policy should be presented in an explicit and transparent manner so that individuals can apply different risks and time horizons if they wish to.

These issues seem to have received little focus, and they need further attention in matters of different health behaviors in different contexts. In our efforts to explore health behaviors, we need to integrate research from medicine, economics, psychology, education, sociology and political science.

The perceived risks and time horizons for various outcomes of health procedures could be explored in future studies. Increasing numbers of health care professionals and clinical epidemiologists recognizes the future costs of smoking and the benefits of quitting the habit, especially during pregnancy. On a societal level the health education goal is to decrease smoking, which has to be implemented on an individual basis. What is the perceived risk for various outcomes of smoking? When will the outcomes occur? How do these outcomes affect behavior?

Clinicians must daily make many decisions where patients are involved. In the clinical decision-making situation the norms are set by society, and they are to be implemented by the individual. Various aspects of risks and the time dimension should be taken into account in prescriptions and proposals. But while the clinician generally has knowledge of risk and different time aspects in health, this may be not the case for the patient, and the level of compliance with prescriptions and proposals might increase when the physician provides information about risks and the time perspective. It would be very appropriate to conduct studies examining these two variables in clinical decision-making from the perspectives of both physician and patient, since their psychological experience of risk and time might differ.

Rheumatoid arthritis is a chronic disease that demands life-long treatment, with a high risk of constant deterioration in several aspects of health. The treatment involves compliance and persistence over a person's lifetime, where risk and time effects must be quite high, with subsequent difficulty in behavior change (Ortendahl & Fries, 2000). An educational aspect should be included, since framing health messages as large, important outcomes with long delays might induce greater compliance (Ortendahl & Fries, 2005).

Classical decision theory is based on the two factors of probability of outcome and value of outcome. The time effect has not been extensively taken into account in attempts to explain decisions related to health. This applies to policy-level decisions, clinicians' decisions, and the behavior of patients and the general public in relation to preventive health or treatment. When resources are being allocated, time factors must be included if a complete analysis of decisions is to be made. A decision analysis model based on time of outcome and chance of outcome is a theoretical and practical approach that has an intuitive appeal when it comes to health-care interventions. However, we are only at the very beginning of this field of research, and extensive research remains to be done to assess the impact of time on health-related judgments and decisions.

### REFERENCES

- Ariely, D., & Zakay, D. (2001). A timely account of the role of duration in decision making. Acta Psychologica, 108, 187-207.
- Becker, G.S. (1964). Human Capital. New York: Colombia University Press.
- Buchanan, L., & O'Connell, A. (2006). A brief history of decision making. *Harvard Business Review*, 84, 32-41.

Chapman, G. (2002). Your money or your health: time preferences and trading money for health. *Medical Decision Making*, *22*, 410-416.

- Chapman, G.B., Brewer, N.T., Coups, E.J., Brownlee, S., Leventhal, H., & Leventhal, E.A. (2001). Value for the future and preventive health care. *Journal of Experimental Psychology. Applied*, 7, 235-250.
- Cher, D.J., Miyamoto, J., & Lenert, L.A. (1997). Incorporating risk attitude into Markovprocess decision models: importance for individual decision-making. *Medical Decision Making*, 17, 340-50.
- Cropper, M.L., Aydede, S.K., & Portney, P.R. (1994). Preferences of life saving programs: how the public discounts time and age. *Journal of Risk and Uncertainty*, *8*, 243-265.
- Dellve, L., Lagerstrom, M., & Hagberg, M. (2003). Work-system risk factors for permanent work disability among home-care workers: a case-control study. *International Archives* of Occupational and Environmental Health, 76, 216-24.
- Dixon, B. (2005). Risk aversion. The Lancet Infectious Diseases, 5, 599.
- Drummond, M.F. (ed). (2002). *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press.
- Feldman, S.R., Chen, G.J., Hu, J.Y., & Fleischer, A.B. (2002). Effects of systematic asymmetric discounting on physician-patient interactions: a theoretical framework to explain poor compliance with lifestyle counseling. *BMC Medical Informatics and Decision Making*, 2, 8.
- Fries, J.F. (2003). Measuring and monitoring success in compressing morbidity. Annals of Internal Medicine, 139, 455-459.
- Gafni, A., & Torrance, G., W., (1984). Risk attitude and time preference in health. *Science, Medicine and Man, 30*, 440-451.
- Gibbons, F., McGovern, P.G., & Lando, H.A. (1991). Relapse and risk perception among members of a smoking cessation clinic. *Health Psychology*, 10, 42-45.
- Gravelle, H., & Smith, D. (2001). Discounting for health effects in cost-benefit and costeffectiveness analysis. *Health Economics*, 10, 587-599.
- Groot, W., Van Den Brink, H.M., & Plug, E. (2004). Money for health: the equivalent variation of cardiovascular diseases. *Health Economics*, 13, 859-72.
- Hakes, J.K., & Viscusi, W.K. (1997). Mortality risk perception: a Bayesian reassessment. *Journal of Risk and Uncertainty*, 15, 135-150.
- Hoel, M. (2003). Allocating health care resources when people are risk aversive with respect to life time. *Health Economics*, 12, 601-608.
- James, P.D., Manuel, D.G., & Mao, Y. (2006). Avoidable mortality across Canada from 1975 to 1999. BMC Public Health, 6, 137.
- Jardine, C., Hrudey, S., Shortreed, J., Craig, L., Krewski, D., Furgal, C., & McColl, S. (2003). Risk management frameworks for human health and environmental risks. *Journal of Toxicology and Environmental Health*, 6, 569-720.
- Johannesson, M., & Meltzer, D. (1998). Some reflections on cost-effectiveness analysis. *Health Economics*, 7, 1-7.
- Kempainen, R.R., Migeon, M.B., & Wolf, F.M. (2003). Understanding our mistakes: a primer on errors in clinical reasoning. *Medical Teacher*, 25, 177-181.
- Klayman, J., Soll, J.B., Gonzales-Vallejo, C., & Barlas, S. (1999). Overconfidence: it depends on how, what and whom you ask. *Organizational Behavior and Human Decision Processes*, 79, 216-247.
- Krupnick, A.J., & Cropper, M.L. (1992). The effect of information on health risk evaluations. *Journal of Risk and Uncertainty*, 5, 29-48.

- Lahiri, S., Markkanen, P., & Levenstein, C. (2005). The cost effectiveness of occupational health interventions: Preventing occupational back pain. *American Journal of Industrial Medicine*, 48, 515-529.
- Langford, I.H., Marris, C., McDonald, A.L., Goldstein, H., Rasbash, J., & O'Riordan, T. (1999). Simultaneous analysis of individual and aggregrate responses in psychometric data using multilevel modeling. *Risk Analysis*, 19, 675-683.
- Lewis, D.K., Robinson, J., & Wilkinson, E. (2003). Factors involved in deciding to start preventive treatment: qualitative study of clinicians' and lay people's attitudes. *British Medical Journal*, 327, 841.
- Loewenstein, G.F., Weber, E.U., Hsee, C.K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127, 267-86.
- Montgomery, A.A., & Fahey, T. (2001). How do patients' treatment preferences compare with those of clinicians? *Quality in Health Care, 10*, 139-143.
- Ortendahl, M. (2005). Using time framing effects to improve health education. *International Journal of Health Promotion & Education*, 44, 92-96.
- Ortendahl, M., & Fries, J.F. (2000). Factors influencing length of time taking Methotrexate in Rheumatod Arthritis. *The Journal of Rheumatology*, 27, 1139-1147.
- Ortendahl, M., & Fries, J.F. (2002). Time-related issues with application to health gains and losses. *Journal of Clinical Epidemiology*, 55, 843-848.
- Ortendahl, M., & Fries, J.F. (2005). Framing health messages based on anomalies in time preference. *Medical Science Monitor*, 11, 253-256.
- Ortendahl, M., & Fries, J.F. (2006). A low tension between individual and societal time aspects in health improved outcomes. *Journal of Clinical Epidemiology*, 59, 1222-1227.
- Richardson, S., Thomson, A., Best, N., & Elliott, P. (2004). Interpreting posterior relative risk estimates in disease-mapping studies. *Environmental Health Perspectives*, 112, 1016-25.
- Roelofsma, P.H., & Van der Plight, J. (2001). On the psychology of time preference and human decisions: introduction to the special issue. *Acta Psychologica*, *108*, 91-93.
- Rosen, A.B., Tsai, J.S., & Downs, S.M. (2003). Variations in risk attitude across race, gender, and education. *Medical Decision Making*, 23, 511-517.
- Ryan, E.L., & Skinner, C.S. (1999). Risk beliefs and interest in counseling: focus-group interviews among first-degree relatives of breast cancer patients. *Journal of Cancer Education*, 14, 99-103.
- Salomon, J.A., & Murray, C.J. (2004). A multi-method approach to measuring health-state valuations. *Health Economics*, 13, 281-90.
- Sjoberg, L. (1998). Worry and risk perception. Risk Analysis 18, 85-93
- Sjoberg, L. (2000). Consequences matter, 'risk' is marginal. *Journal of Risk Research, 3*, 287-295.
- Sjoberg, L. (2001). Limits of knowledge and the limited importance of trust. *Risk Analysis*, 21, 189-198.
- Sjoberg, L. (2003). The different dynamics of personal and general risk. *Risk Management: An International Journal, 5,* 19-34.
- Slovic, P., Finucane, M.L., Peters, E., & MacGregor, D.G. (2004). Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, 24, 311-322.
- Sogaard, J., & Gyrd-Hansen, D. (1998). Counting and discounting gained life-years. *Developments in Health Economics and Public Policy*, 6, 51-74.

- Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A referencedependent model. *The Quarterly Journal of Economics, Nov.*, 1039-1061.
- Wakkar, P.P. (2004). On the composition of risk preference and belief. *Psychological Review*, *11*, 236-241.
- Walker, D., & Kumaranayake, L. (2002). Allowing for differential timing in cost analyses: discounting and annualization. *Health Policy and Planning*, 17, 112-118.
- West, R.R., McNabb, R., Thompson, A.G., Sheldon, T.A., & Grimley Evans, J. (2003). Estimating implied rates of discount in healthcare decision making. *Health Technology Assessment*, *7*, 1-60.

Chapter 4

# DATA-DRIVEN DECISION-MAKING IN CLINICAL SETTINGS: APPLICATIONS TO UNIQUE POPULATIONS\*

# Jill D. Stinson, James C. Coleman, Niels C. Beck and Anthony A. Menditto

Fulton State Hospital, Fulton, Missouri, USA

# ABSTRACT

Traditional assessment instruments designed to predict risk, identify appropriate treatment interventions, and inform placement decisions typically rely on heterogeneous normative samples that generalize to a broad range of clinical settings. While this approach is useful in providing empirical data for clinical decision-making, it lacks the specificity that is sometimes necessary in order to address the needs of unique client populations. An additional problem relates to the methodological techniques used to derive prediction equations, which primarily involve linear statistical models and univariate comparisons. We advocate the use of multivariate models and innovative statistical methods to more fully address the complex factors which drive the behaviors of interest in a clinical population.

Here we discuss several examples which are relevant to this issue, including prediction of aggression within a forensic psychiatric institution, assessment and treatment of problematic sexual behavior in individuals with intellectual and developmental disabilities, and prediction of elopement from public psychiatric facilities. For each of these, we address the use of heterogeneous samples to inform decision-making and how this process may be enhanced through the use of creative statistical procedures (e.g., latent class analysis, logistic regression) and sampling techniques more specific to institutional needs.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

Decision-making in psychology involves a complex process of integrating clinical, statistical, and policy concerns in order to address a variety of issues. Important decisions are made regarding selection of assessment instruments and interpretation of results, determination of appropriate treatment interventions, evaluation of treatment efficacy, considerations of placement in secure or treatment-intensive settings, or estimation of risk of some future behavior, just to name a few. As psychologists and other mental health professionals are called upon to provide the best possible outcome in these situations, it becomes all the more critical for us to have available the most accurate and reliable data to inform these decisions.

Current and historical psychological trends have recommended the use of empirical data to improve the precision of decision-making. Early discussion prompted by Meehl (1954) claimed the superiority of empirical, actuarial prediction methods for a variety of outcomes over clinical judgment. Extensive follow-up of his original hypotheses has repeatedly demonstrated the usefulness and greater accuracy of empirical prediction techniques over those relying on clinical knowledge or experience alone (Grove & Lloyd, 2006; Hilton, Harris, & Rice, 2006; Marchese, 1992). While these empirically-derived decision-making techniques are typically applied to the prediction of risk, they could be successfully employed in treatment outcome research, including studies of mediators and moderators, efficacy, and overall response to treatment in much the same manner. A growing number of clinicians, researchers, and policy makers have recognized the potential of empirical data to enhance clinical decision-making and could apply actuarial prediction to a variety of clinical concerns.

However, we must consider possible methodological limitations present in the empirical research underpinning these decision-making processes. A common criticism from those who rely on research findings to make decisions is the perceived lack of applicability to the "real world" of clinical practice (e.g., Magill, 2006). Are the empirically-derived strategies currently available generalizable to specific or unique populations of individuals? Are the statistical methods those capable of addressing the complex and multivariate nature of the questions being asked? In this chapter, we will review some of the limitations of sampling and statistical techniques that need to be taken into consideration when using the empirically-derived decision-making techniques currently available. We will also discuss several ways in which to address these issues, including suggestions for improved research design and expanded use of statistical modeling. We will conclude the chapter with several examples of how special populations and creative statistical methodology have been used to refine assessment and prediction in a specific treatment environment.

# LIMITATIONS OF CURRENT SAMPLING PRACTICES – WHO IS INCLUDED IN THE EMPIRICAL RESEARCH?

Efforts at empirically validating decision-making processes in psychology include a number of steps, such as operationally defining and measuring outcomes, selecting relevant indicators which play a role in the decision being made, and identifying appropriate samples for testing hypotheses and generalizing results. Sampling procedures are a critical component of this process, as they largely determine the standard against which decisions are made.

For example, if we were interested in measuring the effectiveness of anti-psychotic medications in treating symptoms of psychosis in patients with schizophrenia, we would need a) participants with schizophrenia, b) participants who have demonstrated symptoms of psychosis, and c) participants who have taken or who are eligible to take anti-psychotic medication. While the reasoning behind this may seem all too apparent, too often we are faced with attempts at empirical validation that are not applicable to the population of interest or may exclude large portions of those affected by the decisions being made. Expanding on the previous example, this would mean excluding individuals without schizophrenia or psychosis but who also take anti-psychotic medications, or excluding a number of people who meet the criteria but are significantly different from the rest of the individuals in the sample (e.g., inpatient vs. outpatient participants). When this is done, it causes us to question the accuracy of the empirical findings and our ability to make valid decisions or predictions regarding the outcome. As decision-makers, we have information regarding the outcome of interest, but may not be able to confidently or reliably apply the information to those in our own client population. Here we will further discuss these issues related to sampling procedures and how they can impact decision-making with specific groups. We will also explore potential solutions for addressing the limitations of current sampling practices in empirical research.

#### Heterogeneous Normative Samples

Traditional instruments, methods, or tools designed to predict risk, identify appropriate treatment interventions, and inform placement decisions typically rely on heterogeneous normative samples that generalize to a broad range of clinical settings. This allows for the most representative sample available. However, while this approach may be useful in providing empirical data for clinical decision-making, it lacks the specificity that is sometimes necessary in order to address the needs of unique client populations.

First, this may distort base rates of certain outcomes to such an extent that it is difficult to apply resulting empirical formulas to another population. Prediction of certain behaviors – such as violence, treatment attrition, or treatment failure – relies on an understanding of the normal occurrences of these outcomes in a given population. Large samples are often used in this type of research to provide enough information to detect low-base-rate phenomena. Not all populations though evidence the same or similar base rates for these outcomes. The standard base rate of violence among inmates in a maximum security prison, for example, is likely higher than that in a minimum security facility. However, a study attempting to identify the expected rate of violence among prison inmates might include both groups in the sample, thus inflating the base rate for violence among minimum security inmates and minimizing the potential violence among those in maximum security. This concern is highly relevant for those making decisions regarding risk of criminal recidivism, medication non-compliance, self-harm, or other behaviors which may carry significant consequences for the individual and concerned others.

A second issue involves other assumptions about expected outcomes derived from heterogeneous comparison groups. Returning to the example involving empirically-derived decision making for patients with schizophrenia who are taking anti-psychotic medications, let's say that the majority of a large, multi-site sample of participants showed improvement in symptoms of psychosis, treatment motivation, work attendance, and interpersonal interactions following the recommended regimen of medication therapy. This sample included a mix of patients from an inpatient facility, a short-term day treatment center, and an outpatient clinic. A closer look at the sample might reveal significant differences in these outcomes for the three different groups of participants that become muddled once they are combined into one group. This would then limit our ability to apply outcome expectations to a single client who has never before received treatment in any of these settings, or a client who has received both outpatient and inpatient services within the past several months.

These and other related problems reflect the underlying conflict many researchers face when seeking to maximize the external validity of their research. Considerations of external validity – as evidenced by how widely representative and therefore generalizable a sample may be – are critical for ensuring that the findings may be applicable to a variety of settings and individuals and not based entirely on a small, relatively biased sample. As important as this may be, there are those times, as described above, when we need to focus a sample more narrowly in order to increase the precision, accuracy, and meaningfulness of our results.

A recent illustration of these considerations involves the search for variables to predict future violence or dangerousness in those who have a history of engaging in similar behaviors. A number of researchers have focused on developing reliable, empiricallyvalidated instruments for determining potential risk of aggressive behaviors prior to placement in a less restrictive or community setting (e.g., Monahan, Steadman, Appelbaum, Robbins, Mulvey, Silver, Roth, & Grisso, 2000; Monahan, Steadman, Robbins, Appelbaum, Banks, Grisso, Heilbrun, Mulvey, Roth, & Silver, 2005; Quinsey, Rice, & Harris, 1995; Rice & Harris, 1997; Steadman, Silver, Monahan, Appelbaum, Robbins, Mulvey, Grisso, Roth, & Banks, 2000). Some of the more recent research (Monahan, et al., 2000; Steadman, et al., 2000) has utilized a sample of nearly 1000 acute, civilly-committed inpatients from the MacArthur Violence Risk Assessment Study (Monahan, Steadman, Silver, Appelbaum, Robbins, Mulvey, Roth, Grisso, & Banks, 2001) for identifying relevant risk factors and determining appropriate cutoffs on measures of potential risk. The findings by these researchers have been cross-validated on a similar sample of participants, with formulas then incorporated into a software package called the Classification of Violence Risk (COVR; Monahan, Steadman, Appelbaum, Grisso, Mulvey, Roth, Robbins, Banks, & Silver, 2006) to be used in placement decisions involving acute inpatients in a civil psychiatric setting.

Though these authors have cautioned against applying their findings to other samples of individuals, it would obviously be tempting to many policy- and decision-makers to use an already-established risk assessment measure, normed on a broad sample of psychiatric patients, to aid in the determination of future risk. However, there are serious problems inherent in applying these statistically-derived methods to an inappropriate comparison group. Base rates of prior violence in the original and cross-validation samples were relatively low, nearing 20%, which may be significantly lower than reported rates of violence among individuals in forensic psychiatric or incarcerated populations. Therefore, the original weighted measures using infrequency of violent behavior as a major predictor would be problematic and could inflate estimates of potential dangerousness in non-comparable groups. Other related issues include the fairly low rates of certain categories of mental disorder in the MacArthur data sample (e.g., schizophrenia in approximately 17%; Monahan, et al., 2000), which would be perhaps vastly different from rates of these disorders in long-term psychiatric institutions. Once again, one must be extremely cautious when applying empirically-

established formulas to a population that is so markedly different from the normative sample. The statistical weights derived through empirical analysis and any subsequent actuarial formulas might not represent other targeted groups.

#### **Exclusionary Criteria**

In addition to concerns of heterogeneous sampling methods and their potential for creating bias when applied to unique groups, there is also the question of the exclusionary criteria used in empirical research and their impact on any subsequent findings. This is perhaps an extension of the issues described above, where the heterogeneity of normative groups is not the actual problem, but the way in which these groups may be created. In order to increase the generalizability, extreme variants of behavior may be excluded from the research sample. In some cases, the same variants of behavior or types of participants may be routinely excluded, consistently biasing the research in one direction.

This is particularly problematic when considering the possible result of such exclusions. Let us again return to our original example related to research on outcomes in participants with schizophrenia following the use of antipsychotic medications. What if this research routinely excluded individuals demonstrating one category of symptom behavior, such as catatonia, due to difficulties with generalizing to the rest of the sample? We would have difficulty applying the results to or predicting outcomes with individuals with catatonic schizophrenia, regardless of how impressive or promising the empirical findings may be in support of these medications.

This, of course, does not mean that empirically-validated techniques should never be used for situations or individuals beyond those included in the original validation research, but it does limit our ability to understand and predict the outcome. The costs of this may be high. Perhaps the treatment in question is quite expensive, or carries risks for the patient. Maybe the behavior which we are trying to predict is one with significant risk to the individual (e.g., suicide) or the community (e.g., aggression). When unique populations of individuals are excluded from or poorly represented in the research samples, we lose some of the benefit that is gained through empirical study. At that point, the predictive factors we are considering have moved beyond that which is empirically known.

An example illustrating these points can be found in the literature regarding the assessment and treatment of sexual offenders with developmental disabilities and severe and persistent mental illness. These individuals are often excluded from the mainstream research in this area. Assessment instruments frequently used to establish norms for sexual interests, historical behaviors, and paraphilic diagnoses have not been normed using samples of offenders with developmental disabilities (e.g., MSI-II; Nichols & Molinder, 2000). Other instruments, such as the polygraph, were only recently empirically evaluated using samples of intellectually/developmentally disabled sexual offenders and have demonstrated questionable validity with these groups (e.g., Branaman & Gallagher, 2005). Until this research was made available, it was only assumed that sex offenders with developmental disabilities would perform in the same manner as their non-disabled counterparts on this instrument.

Some treatment research in the sex offender literature excludes participants with severe and persistent mental illness, which may interfere with the effects of treatment and limit the generalizability of the results to a non-psychiatric sex offending population. While these concerns of external validity are important, the exclusion of these groups may prevent us from gaining critical knowledge related to treatment response, indicators of risk, and future treatment needs in a community setting.

These issues are especially troubling given evidence suggesting that groups of sexual offenders with either developmental disabilities or severe and persistent mental illness (or both) may be a larger group than what is frequently observed in the research (e.g., Barron, Hassiotis, Banes, 2002; Lambrick & Glaser, 2004; Lindsay, 2002). While this is not known with any great certainty, it would clearly highlight deficiencies in the way we empirically validate our hypotheses for use in decision-making tasks. If these individuals are not included in the research yet represent a significant number of the offenders in our institutional settings, we have not adequately addressed the empirical question regarding etiology, assessment, treatment, and risk.

#### Addressing Sampling Concerns

The question then remains as to what we can do to remedy or at least alleviate this problem when applying empirical findings to specialized groups. One such solution is to establish "local" norms for special populations. If considering an actuarial instrument designed to predict risk that was normed on non-psychiatric prison inmates with a history of violence, decision-makers in a psychiatric facility could use similar empirical methods to develop their own norms using clients from their facility. This would provide more accurate information as to the base rates of targeted outcomes or relevant historical factors and allow for precise estimates of the importance of these factors within the given sample. Thus, empirically-derived methods which have demonstrated success in larger, heterogeneous samples, but which are perhaps not applicable to more specialized, unique groups due to the heterogeneity of the normed participants, could still be used in informing clinical decisionmaking. This speaks to the overall importance of cross-validation, where outcomes derived from the empirical research are validated or examined in light of other similar or even unique samples. The formulas derived for risk prediction, for example, could be applied to numerous samples to confirm not only the strength or validity of the formula itself, but to establish population-specific base rates, statistical weighting, or probability estimates in unique groups.

Other sampling strategies must take into account the special needs or challenges of those within a unique clinical population. The development of empirically-derived decision-making tools or techniques must incorporate these challenges in order to ensure the most accurate and efficacious results possible. For example, empirical research related to treatment response or treatment attrition in a large mixed sample may not fully address the challenges inherent in administering treatment in a long-term, inpatient psychiatric care facility for individuals with serious mental illness. Considerations of severity of the illness, previous treatment efforts, and other co-morbid psychiatric symptomotology, must be incorporated within an empirical design. If not, the resulting decisions based on non-specific empirical research may grossly overestimate the ease or simplicity of addressing treatment responsiveness issues.

A second example of how to address this issue draws upon the risk assessment literature, where actuarial measures are developed with little consideration of the causal mechanisms of risk. While these mechanisms may be seemingly unimportant in determining a risk probability score, they can be critical in determining risk for a unique population of individuals excluded from the original sample. One actuarial measure of risk of sexual violence, for instance, includes an item related to the establishment and maintenance of long-term intimate relationships prior to incarceration or institutionalization (i.e., Static-99; Hanson & Thornton, 2000). The lack of such relationships is one factor predicting a higher risk of future sexual violence. However, this was originally determined using samples of inmates, who likely had opportunity to form such relationships prior to their incarceration, and not psychiatric patients or institutionalized developmentally disabled individuals, who had less opportunity to form these relationships due to lengthy periods of hospitalization or other supervisory limitations. Inclusion of such groups in future research would allow us to better understand the causal mechanisms – Is it because of personality or motivational issues that they lacked relationships, and this predicts risk, or is it because of lack of opportunity due to severity of cognitive deficits or mental illness that they lacked long-term relationships, and this predicts risk? We would then be able to apply these instruments more confidently those within these samples.

Finally, another method of addressing these concerns with unique populations calls for the inclusion of important policy considerations affecting the specified group. It is irrelevant, for example, to develop risk assessment methodologies for a sample of individuals who will never be given realistic opportunity to attempt a desired outcome. Some research has suggested that individuals who have committed violent crimes, for example, have lower recidivism rates for similar criminal behavior when compared to nonviolent offense recidivism (e.g., Holland, Holt, & Beckett, 1982). Similarly, psychological and criminological research has demonstrated that individuals are not "specialists" with regards to one specific type of criminal behavior and can thus be expected to engage in a range of criminal or antisocial acts, a small percentage of which may be violent (e.g., Gottfredson & Gottfredson, 1990; Gottfredson & Hirschi, 1990; Simon, 1997). However, this does not mean that low recidivism for violent vs. nonviolent offenses, nor the fact that they might be more likely to engage in nonviolent criminal behavior, should lead to the unquestioned release of these individuals. Other factors are taken into account, many of which are populationspecific, given institutional policies, best practices within the mental health field, societal concerns, and individual needs. Not only does this suggest that we should incorporate these concerns into our empirical research, but also that we should sample our population accordingly and examine those for whom prediction, assessment, and treatment may be most relevant. Inclusion of the most relevant groups in the development of empirically-derived decision-making processes will allow us to address these issues.

# STATISTICAL DESIGN AND Empirically-Derived Decision-Making

Just as sampling practices can profoundly impact the findings of empirical research and its applicability to decision-making, statistical design and other related methodological issues shape the way in which information is gathered and subsequently used in clinical settings. Statistical design is perhaps at the heart of empirical research, where outcomes are based on standardized methodological practices and at times complex statistical techniques which can help us identify important relationships between variables and derive formulas to aid in predicting desired outcomes. The use of advanced statistical techniques has greatly aided in decision-making for mental health professionals who consult the relevant literature. However, there are instances in which current statistical practices may be insufficient to address the questions being asked, and newer methodologies must be considered.

Traditional statistical methods in clinical research fall into several categories. The first of these involves univariate, linear comparisons between independent and dependent variables. One example of this is most commonly seen in treatment outcome research, where two or more similar groups are administered different forms of treatment and then compared after a period of time. In this type of study, Group A receives the active or experimental treatment, whereas Group B receives another type of treatment (e.g., treatment as usual) or a placebo as a control group. Decisions may then be made based on the overall effectiveness of the treatment, its relative cost, and other clinically relevant factors. However, this type of study may not fully capture the other processes at work, such as moderating variables which affect treatment response, interaction effects, and the potential effectiveness of combined treatment efforts which may be more commonly seen in broader or more comprehensive clinical settings (e.g., Rohrbaugh, Shoham, & Racioppo, 2002).

Another type of statistical procedure typically used in the clinical empirical literature involves one or more variables predicting one specified outcome in a linear sequence. For example, a clinician might be interested in the impact of motivation, engagement in treatment, and understanding of treatment materials on responsiveness to treatment after a given period. Standard procedures might include linear regression, where the predictive relationships of each of these three variables can be determined. However, many clinicians will recognize that each of these – motivation, engagement, and understanding – may change throughout the treatment process and are therefore not strictly linear. Other statistical techniques may be needed in order to make the most informed decisions regarding these and other relevant nonlinear relationships.

A similar problem arises when we want to evaluate the combined effects of several independent variables on one or more dependent variables. The standard linear regression procedures may be capable of examining multiple predictor variables, but they may lack the sophistication to identify underlying relationships (e.g., latent factors) or interaction effects between these variables. In these instances, more complex methodologies must be considered that can account for such potential relationships.

A second type of methodological procedure commonly employed in the empirical literature involves the statistical classification of groups. Often, different groups of individuals are identified and classified based on aggregated counts of targeted outcomes or behaviors. For example, assessment instruments designed to predict dangerousness among a group of individuals rely on an aggregated count of the number of violent or aggressive behaviors committed by each individual in the sample. These counts are then labeled as high or low, depending on overall base rates within the sample. Those above a certain cut-off are considered "high-risk," while others below that cut-off point are labeled "low-risk" for these behaviors. While this makes inherent sense, in groups of individuals with higher base rates of violent behaviors, less obvious cut-off points, or unevenly distributed acts of aggression during the measurement period, other statistical procedures may be needed in order to determine the relative risk between groups of participants.

This also leads to other concerns not addressed via these means. What if everyone in the sample has committed a violent act, but some acts are more violent than others? Do

aggregated counts adequately reflect the difference between an assault resulting in no injury and one resulting in grievous bodily harm to the victim? "Weighting" the outcome measures in an aggregated count is usually not possible within the context of traditional statistical methods. This is an important consideration though for individuals seeking empirical validation of methods within a unique population.

In addition to the questions of relative risk between groups and severity of the outcome behaviors, aggregated counts of specific behaviors at given time points also do not fully explain predicted patterns of behavior over time. These counts do not reflect the actual behavior of individual participants within the group. Measures of violent recidivism after a 10-year follow-up, for example, merely tell us how many individuals were re-arrested after 10 years. They do not indicate other important features of this group, such as the fact that perhaps only 10% recidivated within six months of release from an institution, whereas the remaining 90% remained arrest-free for another three years or more. This type of knowledge would perhaps dramatically change discharge planning for those individuals who fall within the 10% who were immediately arrested for another violent offense.

A similar issue has been raised by others (e.g., Monahan & Steadman, 1996) who have questioned the attempts at predicting behaviors in the seemingly far distant future rather than what we can successfully know after several months of observation and follow-up. The idea that predictive variables are dynamic and may change is one relatively recently introduced to the actuarial prediction literature. Statistical methods of determining risk should be able to account for changing relationships between predictors over time. These considerations, as well as those mentioned above related to aggregated counts and classification into various groups, are important if one is responsible for making associated decisions.

A final concern with the statistical classification of groups involves the way in which these groups are created. In many cases, groups are identified and then classified on the basis of one outcome variable. Returning for a moment to the empirical literature involving the prediction of dangerousness, groups are often labeled as high- or low-risk according to their rate of arrest following release from an institution. While that may be the issue of greatest concern to the researcher, it may not fully capture other concerns relevant to clinical decisionmaking. For example, are they at high risk if they are unable to find a job, sustain a supportive social network, obtain adequate housing, or afford prescribed medications? Many clinicians and policy-makers would say yes, as they are concerned with a variety of issues beyond that of the original research question (i.e., Will they commit another violent act?), or they may recognize the potential relationship between these factors and a "high" risk of committing a violent act in the future.

This then leads to the question of what can be done to remedy some of these limitations in current statistical methods in the clinical literature, as they are applied to decision-making practices within clinical settings. Below we will discuss a variety of statistical techniques which may prevent some of the problems previously described, as well as conclude with a section describing the implementation of some of these procedures in a unique clinical setting.

As discussed above, in some instances the traditional univariate linear regression procedures may be insufficient to adequately address the clinical research question being posed. Dichotomous or otherwise categorical dependent variables, curvilinear relationships between variables, or evaluation of multiple independent and dependent variables might all necessitate the use of multivariate nonlinear regression techniques. This may include logistic regression, which allows for the researcher to examine nonlinear relationships between independent and dependent variables, allowing for the simultaneous evaluation of multiple independent factors on the outcome of interest. Additionally, this procedure may be particularly useful in clinical decision-making contexts where the outcome is dichotomous or polytomous, such as treatment success vs. failure, or high vs. medium vs. low risk. The results obtained from this analysis, typically presented in the form of an odds ratio, may assist in more accurate classification of individuals into clinically meaningful groups than techniques which allow for the evaluation of only one potential probability or outcome and restrict the decision-maker to only those variables which demonstrate linear relationships. A similar procedure is ordinal regression, which allows for the evaluation of ranked categorical variables (e.g., 100% treatment compliant, sporadically treatment compliant, treatment noncompliant) and also produces odds ratios to assist in classifying individuals into meaningful groups.

Latent variable modeling procedures are also being increasingly used by clinicians and researchers to evaluate underlying factors which may not be identified by standard univariate techniques. These methods not only describe the relationships between observed variables and outcomes and latent variables which are not immediately evident or available to the clinician, but they can also aid in classifying individuals utilizing procedures beyond the aggregate count methods described earlier in this section. The most commonly known form of latent variable modeling is factor analysis, which has provided useful contributions to diagnostic assessment procedures and treatment efficacy research (e.g., Bagby & Parker, 2001; Rogers, Jackson, Sewell, & Salekin, 2005; Rojahn, Matson, Naglieri, & Mayville, 2004). Other forms of latent analysis which show promise for clinical decision making include path analysis, which provides causal explanations for certain behaviors within a population, and latent growth modeling, which assigns individual cases into classes or categories and describes changes in the behavior or outcome of interest over a given time period. This presents advantages over the traditional classification procedures previously discussed in that they highlight underlying relationships which may not be directly observable and can identify trajectories rather than single-point cross-sectional data of expected behavior.

Another statistical technique that has been used with regards to prediction of risk in clinical settings involves iterative classification trees (Breiman, Friedman, Olshen, & Stone 1984), a stepwise regression procedure which identifies those variables which most differentiate individuals into specified outcome groups. Here, variables are considered in an iterative fashion, eliminating those which are irrelevant when preceded by far more statistically significant predictors. This presents advantages not only in that it provides a type of predictive equation to aid in classification, but also in that it allows for multiple thresholds (e.g., threshold for high risk vs. threshold for low risk; Monahan, et al., 2000). Additionally, iterative classification trees simplify the process of clinical decision making when compared to other regression techniques – traditional regression methods require responses for all predictor variables of interest, where iterative classification trees only require contingent responses, thus eliminating the acquisition of perhaps unnecessary information.

In sum, the statistical procedures described thus far are only several possibilities which may improve the accuracy and precision of clinical decision making. Some of the traditional methods used in empirical research may be inadequate for evaluating the complex questions which commonly arise within clinical settings. Now we will turn to several illustrations from unique clinical populations to demonstrate the points outlined thus far.

# APPLICATIONS OF THESE PRINCIPLES TO UNIQUE POPULATIONS

In this first example (Menditto, Linhorst, Coleman, & Beck, 2006), administrators and clinicians in a public mental health system had become concerned with elopements from state hospitals. These concerns were heightened as the population of patients residing in the four long-term hospitals in their state had become increasingly forensic, with many of these patients having committed serious violent crimes. Administrators sought guidance at predicting patients at greatest risk for elopement as they undertook a revision of state hospital operations and policies in an attempt to reduce elopements, especially of violent offenders. A review of the literature revealed many demographic and clinical characteristics of patients that have been associated with elopement from psychiatric hospitals in one or more studies. In fact, so many were the variables identified to render their use impractical—nearly all patients would be identified as posing some risk of elopement, with little certainty in differentiating the level of risk among the patients in this specific population. Thus, researchers were asked to study the problem utilizing existing databases within the state mental health system.

The prediction of elopement poses a classic dilemma within decision theory, involving the prediction of a low base rate behavior with enough sensitivity to be useful and enough discrimination to avoid unduly restricting patients who pose little risk. Nonetheless, comprehensive databases existed that provided extensive demographic, clinical, and contextual variables as well as historical records of elopement. The authors initially conducted a series of traditional bivariate analyses to describe the populations of individuals who had eloped during the previous year and those who had not. Then, to predict elopement, as well as to illustrate the advantages of a multivariate analysis approach over the traditional bivariate analyses, the authors developed logistic regression models and accompanying classification tables. (For more details, we refer the reader to Menditto, et al., 2006).

Data for this study were collected on all patients receiving care in the four long-term-care public psychiatric facilities in Missouri on July 1, 1995. Importantly, patients residing in the state's only maximum security forensic center were excluded from the study sample. This was done to avoid introducing undesired sampling bias into the analyses, because only two patients had ever successfully eloped from this center in its history. Complete data were available for 636 patients, of which 58 had eloped or attempted to elope during the previous year (base rate = 9%).

Even though predicting elopement for the entire patient sample was potentially informative, it was recognized that the perceived and actual risk to the community in the event of an elopement would be affected by a patient's commitment status and associated criminal history. To illustrate the effect on the prediction of elopement among groups differing on this important dimension of risk, authors divided the sample into a Violent Offenders group (236 patients that had a committed a serious violent offense such as assault, murder, rape, kidnapping, robbery), and Civilly Committed group (400 patients that were admitted voluntarily by self or guardian, were civilly committed, or represented forensic commitments pursuant to less serious offenses). Logistic regression analyses were subsequently conducted separately for the entire sample, the Violent Offender group, and the Civilly Committed group. A stepwise selection method was used in deriving the logistic regression models, using a .05 level of significance criterion for a variable to initially enter the prediction equation, and a retention criterion of .15.

When conducting the logistic regressions, it was also recognized that decision-makers may be interested in selecting from a number of different classification cutoff values tailored to the question they may be facing. For example, a more conservative cutoff value that emphasized the identification of patients that actually eloped (i.e., "true positives"), perhaps even at the cost of increasing the number of errors in predicting elopement for those patients that did not elope (i.e., "false positives"), may be desirable when attempting to predict elopement among the Violent Offenders group. The authors therefore examined the accuracy of classification in predicting elopement in the entire sample, the Violent Offenders group, and the Civilly Committed group, using a conservative cutoff value of .05 as well as a more moderate cutoff value of .20. Further, they identified the best-fit cutoff value that maximized classification accuracy for each group and examined its potential impact on decision-making.

Using traditional bivariate analyses such as t-tests and tests of chi-squared to examine the entire sample, significant differences were found between patients who eloped and patients who did not elope on 11 of the 23 demographic, clinical, and placement variables. In contrast, only 3 of these 11 variables demonstrated predictive power when the logistic regression analysis of the entire sample was conducted, with a fourth significant predictor variable also being identified that did not differentiate elopers from non-elopers in the traditional bivariate analyses. Examining the data in the context of a multivariate analysis allowed for statistical control among the variables included in the prediction equation, yielding more practical and meaningful results than those provided by the traditional bivariate analyses alone.

The prediction of elopement in the Violent Offenders group and in the Civilly Committed group revealed the importance of separately deriving logistic regression models for these two groups. The model for the Violent Offenders group identified three variables that significantly contributed to the prediction of elopement among these patients. Three variables also demonstrated predictive relations with elopement among the Civilly Committed group. However, only one variable (previous history of elopement) predicted elopement in both groups.

Overall classification accuracy for the entire sample (92.3%), the Violent Offenders group (90.7%), and the Civilly Committed group (93.3%) were by definition maximized whenever the best-fit cutoff value for each group was used. However, the nature of the decisions administrators often face may make even the cutoffs that provide optimal classification less than desirable. For example, in the Violent Offender group, only 50% of those patients that eloped were predicted to elope when the best-fit cutoff was used. Prediction of elopement among this group of serious offenders that is accurate only half the time would likely represent too great of a risk to public safety for an administrator to accept. The use of a more conservative cutoff value, while decreasing the overall accuracy of classification among the Violent Offenders group (44.5%), resulted in a substantial increase in the number of patients predicted to elope whom actually eloped (86.7%). There were obvious costs to this conservative approach that increased the ability to accurately identify elopers in this group—many more patients were predicted to elope who did not elope than when using a best-fit cutoff value or a moderate cutoff value. The advantages of using multivariate logistic regression in this manner, however, allows decision-makers the flexibility to focus their analysis on substantive outcomes of interest. Another clear advantage of this approach is the weighting of outcomes, where the costs of Violent Offender elopement vs. the costs of Civilly Committed elopement were considered, thus accounting for realistic clinical and administrative concerns.

A second example and more recent project in which the authors have been engaged pertains to seclusion and restraint use (Beck, Durrett, Stinson, Coleman, Stuve, & Menditto, submitted for publication; Durrett, et al, in preparation). Given increasing nationwide concerns with aggression within psychiatric facilities and the application of preventative measures involving forcible seclusion and restraint, administrators, clinicians, and researchers have been interested in identifying those clients for whom seclusion and restraint may be most frequently applied (e.g., Forquer, Earle, Way, & Banks, 1996).

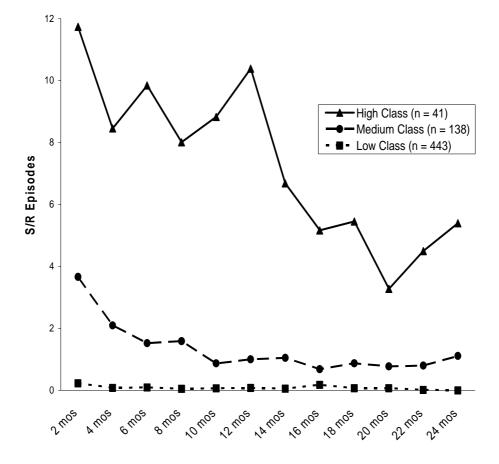
This study was conducted at a rural midwestern state psychiatric facility which contains a high percentage of forensic psychiatric patients. The facility includes the state's only maximum security unit, which comprise 200 of the facility's 500 beds. An additional 200 beds are classified as intermediate security. The sample being studied is comprised of seclusion/restraint records for 622 patients admitted to the facility since September of 2001, and seclusion and restraint use is being tracked over a two-year time frame.

Data analysis involved a two-step process, the first of which was comprised of a latent class analysis, which revealed the presence of three highly discrete classes of seclusion and restraint use (Beck, et al., submitted; Figure 1). The largest class, which made up 70% of the sample, had a very low incidence of seclusion and restraint use (less than one incident per month) over the entire course of the study. The smallest class (7% of the sample) averaged approximately six seclusion and restraint incidents per month during the first two months following admission, and then decreased over the remainder of the study period to a rate of approximately two events per month by the end of the second year. A third class fell between the first two, beginning at a rate of two events per month and decreasing to a rate of .5 events per month by the end of the study.

A polynomial regression procedure was then applied to patient diagnostic, demographic, and other characteristics available at the time of admission, in order to build a predictive equation which might provide clinicians with information about a given patient's expected level of seclusion and restraint during the hospital course; an additional regression model was developed which included a variable which represented each patient's first two months of seclusion and restraint use (Durrett et al., in preparation).

The regression model was built with variables available at admission and included age, admission status, previous hospitalizations, marital status and employment history, as well as diagnoses relating to alcohol and drug abuse, borderline personality disorder, antisocial personality disorder, and intermittent explosive disorder. Across the entire 3x3 classification matrix, this equation produced statistically significant classification rates, with an overall correct classification rate of 75%. The second model included diagnostic and demographic variables in addition to each patient's first two months of seclusion and restraint use. The overall correct classification rate for this model was 88%, again a statistically significant classification rate. Use of the variable comprised by the first two months of seclusion and restraint use resulted in an appreciable increase in predictive accuracy.

Results of this study have obvious implications for clinicians and administrators charged with the responsibility for making decisions about placement of newly-admitted patients, and their eventual transfer to lower levels of security, as well as targeting high-risk individuals for specialized treatment programming designed to reduce seclusion and restraint use. Conversely, these results could be used to identify and "fast track" other individuals who are



unlikely to require seclusion and restraint use for transfer to less restrictive treatment environments.

Figure 1. Latent class trajectories of seclusion/restraint episode frequencies among 622 patients up to two years following hospital admission.

From the standpoint of generalization and cross-validation issues, the results of this study give rise to several points of discussion. First, with regard to cross-validation of this model to other treatment facilities, an issue arises related to seclusion and restraint rates, since the rates that are encountered in higher security psychiatric facilities are almost certainly substantially greater than those which may be encountered in other, lower security treatment environments (Forquer, et al., 1996). This is the same problem described in our previous discussion of base rates, relative to the COVR and rates of prior violence (Monahan, et al., 2000; Monahan, et al., 2006). Previous empirical studies describing seclusion and restraint rates, as well as associated client or environmental characteristics, may be insufficient and poorly applicable to situations which are not comparable in terms of client behavior and institutional policy. As is the case with so many other behavioral prediction issues discussed in this chapter, generalization and cross-validation are key concerns. Methodologically, it can be argued that cross-validation is an absolutely essential step in judging the predictive utility of any model, since it is only upon cross-validation that one will be able to determine the true generalizability/ predictive validity of a given model.

A second issue pertains to the length of time the patients in the current sample were studied, and the fact that the high security nature of the environment in which these predictive equations were developed is often correlated with a longer length of stay, as contrasted with many other psychiatric units, which are short-term and acute in nature, and typical length of stay is measured in days, as opposed to months or years. Additional issues pertain to the distribution and nature of the diagnostic and demographic characteristics included in the predictive equations. Once again, given the largely forensic and at-risk nature of the current sample, questions arise as to whether, when cross-validating these findings to a different population, rates would differ and the predictive power of the equations would be greatly affected.

### **SUMMARY AND CONCLUSIONS**

Advancements in empirical research within the field of psychology have led to improved clinical decision making and a greater ability on the part of decision-makers to understand the causal mechanisms behind relevant client behaviors. However, we must exercise caution in interpreting the vast array of empirical findings that are available within the scientific literature. Considerations of appropriate sampling, cross-validation of empirical findings, and methodological practices must be evaluated before research results can be applied with precision and confidence to unique or specific clinical populations. The suggestions and examples offered here are only the initial stages toward increasing our ability to develop research techniques that will best inform policy and practice in clinical psychology.

#### REFERENCES

- Bagby, R.M., & Parker, J.D.A. (2001). Relation of rumination and distraction with neuroticism and extraversion in a sample of patients with major depression. *Cognitive Therapy & Research*, 25 (1), 91-102.
- Barron, P., Hassiotis, A., Banes, J. (2002). Offenders with intellectual disability: The size of the problem and therapeutic outcomes. *Journal of Intellectual Disability Research*, 46 (6), 454-463.
- Beck, N.C., Durrett, C.A., Stinson, J.D., Coleman, J.C., Stuve, P., & Menditto, A.A. (submitted for publication). Trajectories of seclusion and restraint use at a state psychiatric hospital.
- Branaman, T.F., & Gallagher, S.N. (2005). Polygraph testing in sex offender treatment: A review of limitations. *American Journal of Forensic Psychology*, 23 (1), 45-64.
- Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1984). *Classification and regression trees*. Pacific Grove, CA: Wadsworth and Brooks/Cole.
- Durrett, et al. (in preparation authorship to be determined). Variables predicting trajectories of seclusion and restraint use in a state psychiatric hospital.
- Forquer, S.L., Earle, K.A., Way, B.B., & Banks, S.M. (1996). Predictors of the use of restraint and seclusion in public psychiatric hospitals. *Administration and Policy in Mental Health and Mental Health Services Research*, 23 (6), 527-532.

- Gottfredson, S.D., & Gottfredson, D.M. (1990). *Classification, prediction, and criminal justice policy*. Washington, DC: National Institute of Justice.
- Gottfredson, M.R., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press.
- Grove, W.M., & Lloyd, M. (2006). Meehl's contribution to clinical versus statistical prediction. *Journal of Abnormal Psychology*, 115 (2), 192-194.
- Hanson, R.K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior*, 24 (1), 119-136.
- Hilton, Z.N., Harris, G.T., & Rice, M.E. (2006). Sixty-six years of research on the clinical versus actuarial prediction of violence. *Counseling Psychologist*, 34 (3), 400-409.
- Holland, T.R., Holt, N., & Beckett, G.E. (1982). Prediction of violent versus nonviolent recidivism from prior violent and nonviolent criminality. *Journal of Abnormal Psychology*, 91 (3), 178-182.
- Lambrick, F., & Glaser, W. (2004). Sex offenders with an intellectual disability. *Sexual* Abuse: A Journal of Research and Treatment, 16 (4), 381-392.
- Lindsay, W.R. (2002). Research and literature on sex offenders with intellectual and developmental disabilities. *Journal of Intellectual Disability Research*, 46 (Suppl1), 74-85.
- Magill, M. (2006). The future of evidence in evidence-based practice: Who will answer the call for clinical relevance? *Journal of Social Work*, 6 (2), 101-115.
- Marchese, M.C. (1992). Clinical versus actuarial prediction: A review of the literature. *Perceptual and motor skills*, 75 (2), 583-594.
- Meehl, P.E. (1954). *Clinical vs. statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis, MN: University of Minnesota Press.
- Menditto, A.A., Linhorst, D.M., Coleman, J.C., & Beck, N.C. (2006). The use of logistic regression to enhance risk assessment and decision making by mental health administrators. *Journal of Behavioral Health Services and Research*, 33 (2), 213-224.
- Monahan, J., & Steadman, H.J. (1996). Violent storms and violent people: How meteorology can inform risk communication in mental health law. *American Psychologist*, 51 (9), 931-938.
- Monahan, J., Steadman, H.J., Abbelbaum, P.S., Grisso, T., Mulvey, E.P., Roth, L.H., Robbins, P.C., Banks, S., & Silver, E. (2006). The Classification of Violence Risk. *Behavioral Sciences and the Law, 24* (6), 721-730.
- Monahan, J., Steadman, H.J., Appelbaum, P.S., Robbins, P.C., Mulvey, E.P., Silver, E., Roth, L.H., & Grisso, T. (2000). Developing a clinically useful actuarial tool for assessing violence risk. *British Journal of Psychiatry*, 176, 312-319.
- Monahan, J., Steadman, H.J., Robbins, P.C., Appelbaum, P.S., Banks, S., Grisso, T., Heilbrun, K., Mulvey, E.P., Roth, L.H., & Silver, E. (2005). An actuarial model of violence risk assessment for persons with mental disorders. *Psychiatric Services*, 56 (7), 810-815.
- Monahan, J., Steadman, H.J., Silver, E., Appelbaum, P.S., Robbins, P.C., Mulvey, E.P., Roth, L.H., Grisso, T., & Banks, S. (2001). The MacArthur study of mental disorder and violence. New York: Oxford University Press.
- Nichols, H., & Molinder, I. (2000). *Manual for the Multiphasic Sex Inventory, 2<sup>nd</sup> edition*. Tacoma, WA: Crime and Victim Psychology Specialists.

- Quinsey, V.L., Rice, M.E., & Harris, G.T. (1995). Actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence, 10* (1), 85-105.
- Rice, M.E., & Harris, G.T. (1997). Cross-validation and extension of the Violence Risk Appraisal Guide for child molesters and rapists. *Law and Human Behavior, 21* (2), 231-241.
- Rogers, R., Jackson, R.L., Sewell, K.W., & Salekin, K.L. (2005). Detection strategies for malingering: A confirmatory factor analysis of the SIRS. *Criminal Justice and Behavior*, 32 (5), 511-525.
- Rohrbaugh, M.J., Shoham, V., & Racioppo, M.W. (2002). Toward family level attribute x treatment interaction research. In H.A. Little, D.A. Santisteban, R.F. Levant, & J.H. Bray (Eds). *Family psychology: Science-based interventions* (pp. 215-237). Washington, DC: American Psychological Association.
- Rojahn, J., Matson, J.L., Naglieri, J.A., & Mayville, E. (2004). Relationships between psychiatric conditions and behavior problems among adults with mental retardation. *American Journal on Mental Retardation*, 109 (1), 21-33.
- Simon, L. (1997). Do criminal offenders specialize in crime types? *Applied and Preventive Psychology*, 6 (1), 35-53.
- Steadman, H.J., Silver, E., Monahan, J., Appelbaum, P.S., Robbins, P.C., Mulvey, E.P., Grisso, T., Roth, L.J., & Banks, S. (2000). A classification tree approach to the development of actuarial violence risk assessment tools. *Law and Human Behavior, 24* (1), 83-100.

Chapter 5

# CLINICAL, PSYCHOLOGICAL AND ETHICAL DILEMMAS: THE CASE OF UNRELATED BONE MARROW TRANSPLANTATION IN ADULT THALASSEMIC PATIENTS<sup>\*</sup>

# G. Caocci<sup>1,†</sup>, S. Pisu<sup>2</sup>, A. Vacca<sup>1</sup>, E. Piras<sup>1</sup>, A. Ledda<sup>1</sup>, R. Littera<sup>3</sup>, C. Carcassi<sup>3</sup> and G. La Nasa<sup>1</sup>

<sup>1</sup>Bone Marrow Transplantation Centre, "R. Binaghi" Hospital, Departmenf of Hematology, University of Cagliari, Italy <sup>2</sup>Department of Legal Medicine, University of Cagliari, Italy <sup>3</sup>Department of Medical Genetics, University of Cagliari, Italy

# ABSTRACT

Thalassemia mayor is a chronic, progressive and, in the long term, life-threatening condition. Extensive ongoing medical care in this pathology (regular transfusions and iron-chelation) can have a major impact on the physical, psychosocial well-being and quality of life (QoL) of patients and their families. Allogeneic bone marrow transplantation (BMT) from an HLA identical sibling is currently the only treatment available to cure these patients. Unfortunately, the probability of finding an HLA-identical donor within the family is less than 30%. In all other cases, it is necessary to search for a compatible donor among volunteer donors worldwide. Unrelated BMT in adult thalassemic patients is burdened by a transplant-related mortality of about 30%. For this reason, patients without a sibling donor need to make a difficult and radical decision. They can either continue traditional transfusion and chelation therapy, with no chance of

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence concerning this article should be addressed to: Dr. Giovanni Caocci, Cattedra di Ematologia, Centro Trapianti di Midollo Osseo, P.O. "R. Binaghi", Via Is Guadazzonis, 3 09126 Cagliari, Italy. Tel: ++39 070 6092800; Fax: ++39 070 6092936; e-mail: gcaocci@alice.it.

complete recovery, or, they can accept the high risk of BMT in the hope of obtaining a definitive cure for the disease. This circumstance gives rise to several social, psychological and ethical problems.

We report the results obtained in 31 adult thalassemic patients transplanted from unrelated donors. To better understand the decision making process, we investigated the motivation factors, the communication strategies, the post-transplantation quality of life (QoL) and discussed the ethical issues of this therapeutic approach. Twenty-four patients (77.4%) are alive and transfusion independent. Seven patients (22.6%) died from transplant-related causes. The surviving patients were given two questionnaires to evaluate pre-transplantation motivation, communication factors and psychological aspects of decision- making. The global QoL was also assessed. All patients were strongly motivated and satisfied with the communication modalities of physicians before transplantation. None of the patients regretted their choice and would make the same decision if reliving the experience. All patients had a good Qol.

The central ethical issue of BMT in a non malignant disease can be framed in terms of a conflict between two fundamental principles of medical ethics: non-maleficence and beneficence. Given the fact that many situations present in medicine where benefits cannot be achieved without at the same time inflicting harm, the question arises as to who will be responsible for the decision and according to what criteria can a fair balance of benefits and burden be struck. The other topic is based on respect for patient's autonomy.

The issue is not simply "to cure" but "to take care" of the patient as a whole person. The multi-disciplinary approach used in this study seems to offers a valuable contribution to the decision-making that surrounds the choice of treatment with a high mortality risk in a chronic non malignant disease.

#### INTRODUCTION

#### Thalassemia and Survival

Thalassemia mayor is a hereditary anemia resulting from several genetic mutations that lead to absent or reduced synthesis of the ß-chain of hemoglobin. The consequences of this syndrome affect multiple organs and are associated with considerable morbidity and mortality. In fact, this common hereditary disorder requires regular blood transfusions and chelation therapy for the iron overload resulting from transfusions and increased gastrointestinal absorption. Iron deposition in the heart, liver, and multiple endocrine glands results in severe damage of these organs. Life-threatening cardiotoxicity, impairment of growth and endocrinopathies are other common features of thalassemia. Despite the severe clinical picture, most cases of childhood thalassemia can now be successfully managed and some can actually be cured. Regular transfusion therapy to maintain hemoglobin levels of at least 9 to 10 g per deciliter allows for improved growth and development and also reduces hepatosplenomegaly as well as bone deformities [1].

Survival of patients with thalassemia mayor continues to improve, especially for patients born shortly before or after the availability of iron chelation. A large Italian study of 977 patients born since 1960 showed that 68% of the patients are still alive at the age of 35 years and that the survival of patients born from 1960 to 1984 continued to improve [2].

Currently, the problems of the older thalassemic patients are represented by poor compliance with iron chelation therapy (especially desferrioxamine), endocrine problems such as diabetes, hypoparathyroidism and failure of puberty, the high frequency of chronic hepatitis, cardiotoxicity besides the psychological and social problems due to this chronic illness.

#### The Dilemma of BMT in Thalassemia

Optimization of erythrocyte transfusion support and regular iron chelation therapy has produced a remarkable improvement in the life expectancy of patients with thalassemia mayor [2-4]. However, extensive ongoing medical care in this chronic, progressive and, in the long term, life-threatening condition can have a major impact on the physical, psychosocial well-being and quality of life (QoL) of patients and their families [5]. Allogeneic bone marrow transplantation (BMT) from an HLA identical sibling is currently the only treatment available to cure patients with thalassemia mayor [6,7]. The drawback of this therapeutical procedure is that the probability of finding an HLA-identical donor within the family is less than 30% [8]. In all other cases, it is necessary to search for a compatible donor among volunteer donors worldwide. Several reports suggest that a more precise characterization of HLA alleles, using high-resolution typing for both class I and class II molecules, can reduce the risk of developing immune-mediated complications and fatal events [9,10]. A previous pilot study conducted on unrelated BMT in thalassemia patients showed that there was considerable overlap between the overall survival (93%) and thalassemia-free survival (80%) of the low and medium risk groups (class 1 and 2) and the outcome of transplantation from an HLA-identical family donor [11]. In the high risk (class 3) group (generally adult patients with hepatomegaly, presence of portal fibrosis and a history of low quality iron chelation), transplant-related mortality was elevated (29%). However, this risk was not different from that observed in transplantation from HLA-identical family donors [12,13]. For this reason, adult class 3 thalassemia patients without a sibling donor need to make a difficult and radical decision. They can either continue traditional transfusion and chelation therapy, with no chance of complete recovery, or, they can accept the high risk of BMT in the hope of obtaining a definitive cure for the disease. The patients and their families who are faced with this decision need to have a realistic picture of all aspects related to transplantation outcome, the recovery of health and the global quality of life in terms of physical and psycho-social well-being.

This study combines the results of unrelated transplantation performed in 31 adult class 3 thalassemia patients with the assessment of pre-transplantation communication and motivation factors, post-transplantation QoL, and the overall satisfaction felt for the decision to undergo BMT. The study also addresses ethical problems with particular emphasis on the physician-patient healing relationship and the moral aspects of clinical decision-making. In fact, this dilemmatic decision requires accurate assessment of three key points. Firstly, to carefully balance the unquestionable beneficence of therapy with the same indisputably bad outcomes: is the beneficence worth the heavy risk of irreparably damaging the patient? Secondly, who are the candidates for this type of intervention, and what are the criteria for the choice? Finally, the need to make the patient aware so that an informed consent or, as proposed in the discussion, a "conscious" consent can be obtained. For these reasons, the relationship between physician and patient in BMT is an important part of the healing relationship. The capabilities of modern medicine, its frightening costs, its ethical challenges,

and its influence on society and culture force us to grapple with the philosophical issues: what is the real purpose of medicine? What are the essential requisites of the healing relationship? In the light of the constant changes within the context of medicine, the answers to these questions are extremely important. Scientific revolution is transforming our conception of disease, health and therapy and the ultimate decisions are taking more account of the values and autonomy of the patient than the clinical evidence or judgment of the physician. It is impossible to make a decision without involving the patient in that choice. The patient needs to be aware of all the factors within a particular scenery. It is the physician who must have the responsibility of the awareness of the patient.

## **PATIENTS AND METHODS**

#### Patient Features

From November 1992 to May 2006, 31 consecutive class 3 thalassemia patients (16 males and 15 females age >16, mean 21 years, range 16 - 37) transplanted from unrelated donors were enrolled in the study. Prior to transplantation, all patients underwent a complete check-up and liver biopsy and were assigned to risk class 3 according to the criteria proposed by Lucarelli et al. [12]. Risk factors were: 1) hepatomegaly; 2) liver biopsy revealing the presence of portal fibrosis; 3) low quality of pre-transplantation iron chelation. Pretransplantation patient characteristics are shown in Table 1. All patients showed marked hepatomegaly, portal fibrosis and had a history of reduced compliance with regular iron chelation. Portal fibrosis was defined in each patient as mild, moderate or severe [14]. Fourteen out of 31 patients (45%) were hepatitis C virus (HCV)-RNA positive. Twenty-two patients (71%) had positive cytomegalovirus (CMV) serology. Approval for the study was obtained from the local institutional review board.

#### **Donor Selection**

All bone marrow donors (15 females and 16 males; age range 19-50 years, mean 35 years) were identified inside the Italian Bone Marrow Donor Registry (IBMDR), except 2 who were found in the National Marrow Donor Program (NMDP) and in the Cyprus Bone Marrow Donor Registry. The median interval between donor search and transplantation was 5 months (range 1-9 months).

In all donor/recipient pairs, the alleles at the HLA-A, B, Cw, DRB1, DRB3, DRB4, DRB5, DQA1, DQB1 and DPB1 loci were identified by Polymerase Chain Reaction-Single Strand Polymorphism (PCR-SSP) (Dynal, Oslo, Norway) and sequence-based typing. All donor/recipient pairs were matched at the HLA-DRB1, DRB3, DRB4, DRB5, DQA1 and DQB1 loci. One pair had disparity for an allele at the HLA-B locus and 2 pairs for an allele at the HLA-Cw locus. Eleven patients were matched with their donors for both HLA-DP alleles, 9 patients had a single disparity, and 7 patients had disparity for both HLA-DP alleles.

#### **Transplantation Procedures**

Eleven patients were transplanted after a myeloablative regimen including Busulfan (BU), and Cyclophosphamide (CY). In 16 patients, the conditioning regimen included Thiotepa (TT). In the remaining 4 patients the conditioning regimen included Threosulphan (TRSF), Fludarabine and TT. Marrow was infused 36 hours after the last dose of chemotherapy. The median bone marrow nucleated cell dose was  $2.8 \times 10^8$ /kg (range 1.8 - 7.5) of recipient weight (Table 1).

All patients received Cyclosporine intravenously and short term Methotrexate for Graft Versus Host Disease (GVHD) prophylaxis. Three patients were also given anti-thymocyte globulin in order to reduce the risk of both graft rejection and GVHD.

Acute and chronic GVHD were graded according to the Seattle criteria [15].

#### Table 1. Patients, donors, pre-transplantation characteristics and transplant outcome

Variable	Frequency	%
Median donor age (years, range)	35 (19-50)	
Median recipient age (years, range)	21 (16-37)	
Gender		
Male	16	52
Female	15	48
HCV RNA		
Negative	17	55
Positive	14	45
CMV serology		
Negative donor / Negative recipient	2	6
Positive donor / Negative recipient	6	19
Negative donor / Positive recipient	3	10
Positive donor / Positive recipient	20	65
Donor/recipient sex combination		
Female donor / Male recipient	10	32
Other combinations	21	68
Liver fibrosis		
None	0	0
Mild	6	19
Moderate	17	55
Severe	8	26
Iron Chelation		
Regular	0	0
Irregular	31	100
Median number of nucleated cells infused		
$(10^8)$ / Kg of recipient (range)	2.8 (1.8-7.5)	
Conditioning regimens		
BU14 / Cy 160	1	3
BU14 / Cy 120	10	32
BU14 / TT 10 / CY 120	4	13
BU14 / TT 10 / CY 120	12	39
Thrf/Fluda/TT	4	13

Variable	Frequency	%
Median donor age (years, range)	35 (19-50)	
Median recipient age (years, range)	21 (16-37)	
GVHD prophylaxis		
CSP + MTX	28	90
CSP + MTX + ATG	3	10
Death*	7	23
Rejection	1	3
Alive with graft	24	77
Acute GVHD II-IV	10	32
Chronic GVHD	8 / 28	29
Limited	6	
Extended	2	

#### **Table 1. Continued**

CMV = Cytomegalovirus; BU = Busulphan; CY = Cyclophosphamide; TT = Thiotepa; Thrf = Threosulphan; Fluda = Fludarabine; CSP = Cyclosporine; MTX = Methotrexate; ATG = Anti-Thymocyte Globulin. \*including the patient who rejected.

#### Questionnaires

After obtaining written informed consent, the questionnaires were either mailed or given personally to patients. All 24-surviving patients (12 males and 12 females, mean age 22 years) were administered two different questionnaires 300 or more days from the BMT procedure: a questionnaire to evaluate the pre-transplantation motivation and communication factors and the EORTC QLQ-C30 questionnaire to assess the global QoL.

#### Communication/Motivation Factors and Satisfaction for the Choice of BMT

The first questionnaire contained eleven items and was developed in collaboration with an expert psychologist to investigate communication and motivation factors and the global satisfaction felt for the decision of BMT. Patients were asked to recall the time when they were first informed of the possibility of BMT. The first item of the questionnaire measured whether satisfaction for the decision to undergo BMT was excellent, good, satisfactory, inadequate or poor. Other 10 items with dichotomous response categories, such as "yes" or "no", investigated whether the information about the procedure and the high mortality risk had been clearly and exhaustively presented; the BMT decision had been taken by the patient or his/her relatives; the patient had been given the opportunity to ask questions about his/her transplant; the patient had asked for psychological support; the physician had talked about quality of life expectancy; the patient regretted the decision to undergo BMT or he/she would still make the same decision if reliving the same experience.

#### Quality of Life

To evaluate the global QoL, patients received the QLQ-C30 questionnaire created by the European organization for research and treatment of cancer (EORTC) quality of life study group, which is a validated multidimensional 30-item questionnaire to assess self reported QoL [16]. This questionnaire is composed of scales evaluating physical, emotional, cognitive, role and social functions, and the global quality of life. Three symptom scales evaluated nausea and vomiting, pain and fatigue, while six single items assessed dyspnea, diarrhea, constipation, appetite loss, sleep disturbances and financial difficulties. The response categories were either dichotomous such as "yes" or "no", or with four categories: "not at all", "a little", "quite a bit", "very much" or distributed on a modified visual analogue scale from 1 to 7. All scores were linearly transformed to a 0 to 100 scale. A cross-sectional assessment of QoL was performed.

#### Statistical Analysis

Survival probability of this cohort of patients was estimated by the product-limited method of Kaplan and Meier [17].

Descriptive statistics were used to evaluate the communication and motivation factors, QoL and the satisfaction for the choice to undergo BMT. The data collected were summarized by reporting the proportion of patients answering each item. On the EORTC QLQ-C30 questionnaire, higher scores on the function scales and the global quality of life scale indicate better functioning: results were summarized as very poor (0-20), poor (21-40), intermediate (41-60), good (61-80) and very good (81-100). Higher scores on the symptom scales reflects problems: results were summarized as none to slight (0-29), moderate (30-69) and severe (70-100) [18]. Student's Test was used to compare the global quality of life and the physical, role and social functions in patients with and without chronic GVHD. Comparisons were also made with transplanted patients of other studies [18,19].

### RESULTS

BMT Outcome

Figure 1 shows the Kaplan-Meier probabilities of survival (77.4%), thalassemia-free survival (74.2%), and Figure 2 shows probabilities of rejection (3.2%) and aGVHD (32.3%) for the 31 patients studied.

Seven patients (22.6%) died after a mean of 165 days (range 17-470) from the BMT procedure for transplant-related causes: hepatic veno-occlusive disease (day + 17), heart failure (day +63) without any sign of engraftment (this patient was also classified as rejection), cerebral hemorrhage due to acute GvHD (day + 95), CMV interstitial pneumonia (day 119), cerebral hemorrhage related to liver failure (day + 125), intestinal obstruction occurring as a complication of intestinal GvHD (day + 269), and liver failure due to chronic GvHD (day + 470). Five of the seven dead patients had been given a conditioning regimen

with BU-TT-CY, while two patients had been treated with BU-CY. One of these two patients died with no take after prolonged marrow aplasia. In 24 patients (77%) the transplantation was successful with complete allogeneic reconstitution. Ten of 31 transplanted patients (32%) developed grade II-IV acute GVHD. Among the 28 evaluable cases, 8 (28.6%) developed chronic GVHD, limited in 6 cases and extended in 2 (Table 1).

The median follow-up of surviving patients was 41 months (range 10-136).

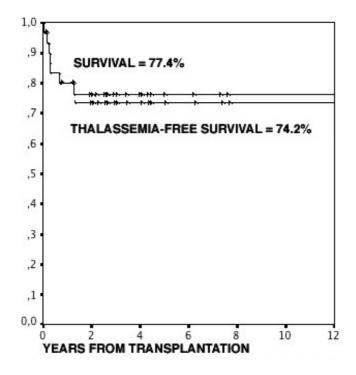


Figure 1. Kaplan-Meier probabilities of survival, thalassemia-free survival, in 31 adult class 3 thalassemia patients transplanted from HLA-matched unrelated donors.

#### Questionnaires

The two questionnaires administered to the 24 surviving patients were returned after a mean of 39 days. The questionnaires were compiled at a median time of 864 days after BMT (range 300 - 4029). The mean age of responding patients at transplantation day was 21 years, while the mean age at the time of the questionnaire was 24 years.

### Communication/Motivation Factors and Satisfaction for the Choice of BMT

Table 2 presents the results obtained with the questionnaire used to investigate communication and motivation factors. Twenty-one of 24 patients (87%) reported excellent or good levels of satisfaction for the decision to undergo BMT (the two categories were collapsed). The majority of patients reported that the physician had provided them with exhaustive information on the transplantation procedure and all of them were satisfied with

the information concerning the high mortality risk. All patients reported that the information had been clearly presented and that they felt to be protagonists of the BMT decision. Nineteen patients (79%) felt that they had been able to ask questions about their BMT procedure. Twenty-two patients (92%) reported that they had been well informed by the physician about the quality of life expectancy. Sixteen patients (67%) reported the need to discuss the decision for BMT with relatives and 17 (71%) expressed a need for psychological support before, during and after the transplantation procedure. None of the transplanted patients regretted his/her choice to undergo BMT and all patients would make the same decision if reliving the experience.

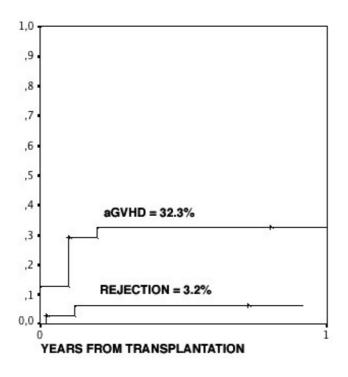


Figure 2. Kaplan-Meier probabilities of rejection and aGVHD, in 31 adult class 3 thalassemia patients transplanted from HLA-matched unrelated donors.

Quality of Life

The results of the QLQ C30 survey are shown in Table 3. The mean global QoL was good (mean 79.5). Sixteen of the 24 patients (67%) enjoyed a very good global QoL (score 81-100), 3 (12%) had a good score (61-80) and 4 (17%) had an intermediate score (41-60). Only one patient had a poor global score (Figure 3).

Very good scores were obtained for the physical, emotional, cognitive, role and social function scales. Patients reported minimal distress symptoms such as fatigue, nausea/vomiting, pain, dyspnea, constipation, diarrhea, appetite loss, insomnia and financial difficulties. Four patients (17%) suffered severely from at least one symptom and 9 (37%) from two or more moderate symptoms.

C/M	Items of the questionnaire	Answer	Patients	%
	-		N°	
	Satisfaction for BMT choice	excellent-	21	87
		good	3	13
		satisfactory	0	0
		inadequate	0	0
		poor		
С	Patient felt that the information about BMT had	yes	21	87
	been exhaustively presented	no	3	13
С	Patient had been given sufficient information	yes	24	100
	about high mortality risk	no	0	0
С	Patient felt that the information about BMT had	yes	24	100
	been clearly presented	no	0	0
М	Patient felt to be protagonist of the decision to	yes	24	100
	undergo BMT	no	0	0
С	Patient needed to discuss the choice of BMT	yes	16	67
	with relatives	no	8	33
С	Patient was able to ask questions about BMT	yes	19	79
		no	5	21
С	Patient needed the support of a psychologist	yes	17	71
		no	7	29
С	Physicians talked about the quality of life	yes	22	92
	expectancy	no	2	8
М	Patient would still make the same decision if	yes	24	100
	reliving the same experience	no	0	0
М	Patient regretted the decision of BMT	yes	0	0
		no	24	100

# Table 2. Communication (C) and motivation (M) factors in 24 adult class 3 thalassemiapatients at mean 864 days after unrelated BMT

Table 4 presents the results obtained in the groups with or without chronic GVHD. In the group of patients that developed chronic GVHD, the global QoL was good (65.3); in those that did not develop GVHD the score was very good (mean 84.3). The difference in global health status between the two groups was statistically significant. All functions yielded lower scores in the group of patients with chronic GVHD but significant differences between the two groups were only obtained for the physical function scale. Patients who developed chronic GVHD reported a higher number of distress symptoms with significant differences in pain and insomnia scales.

864 days after unrelated BNI I			
Domain	Mean (standard error)		
Physical function	91.5 (3.0)		
Role function	80.6 (6.9)		
Emotional function	79.5 (5.0)		
Cognitive function	91.0 (4.0)		
Social function	86.8 (4.0)		
Global quality of life	79.5 (4.6)		
Fatigue	25.0 (6.0)		
Nausea/vomiting	3.5 (2.5)		
Pain	10.4 (5.5)		
Dyspnea	5.6 (2.8)		
Insomnia	9.7 (6.0)		
Loss of appetite	9.7 (3.5)		
Constipation	9.7 (3.5)		
Diarrhea	4.2 (2.5)		
Financial difficulties	9.7 (4.1)		

Table 3. EORTC QLQ-C30 domains in 24 adult class 3 thalassemia patients at mean864 days after unrelated BMT

Higher scores on the function scales and the global quality of life scale indicate better functioning: results were summarized as very poor (0-20), poor (21-40), intermediate (41-60), good (61-80) and very good (81-100). Higher scores on the symptom scales reflects problems: results were summarized as none to slight (0-29), moderate (30-69) and severe (70-100)

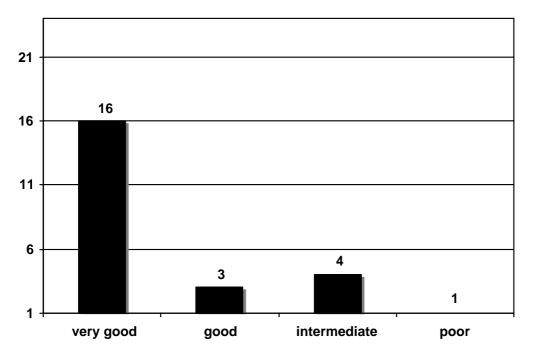


Figure 3. Global quality of life assessed in 24 transplanted patients. Very good: score 81-100, good: score 61-80; intermediate: score 41-60; poor: score<40.

	With cGVHD	Without cGVHD	
Domain	N = 6	N = 18	P-value
	Mean (standard error)	Mean (standard error)	
Physical function	82.2 (8.9)	94.6 (1.9)	0.05
Role function	66.7 (17.2)	85.2 (5.8)	ns
Emotional function	66.7 (13.6)	83.8 (3.6)	ns
Cognitive function	86.1 (10.9)	92.6 (3.2)	ns
Social function	80.6 (10.9)	88.9 (3.1)	ns
Global quality of life	65.3 (11.3)	84.3 (3.6)	0.05
Fatigue	38.9 (12.7)	20.4 (5.5)	ns
Nausea/vomiting	8.3 (8.3)	1.9 (1.2)	ns
Pain	33.3 (17.2)	2.8 (1.4)	0.005
Dyspnea	5.6 (5.6)	5.6 (2.9)	ns
Insomnia	27.8 (18.1)	3.7 (3.5)	0.05
Appetie loss	11.1 (7.0)	9.3 (5.6)	ns
Constipation	5.6 (5.6)	11.1 (3.6)	ns
Diarrhea	5.6 (5.6)	3.7 (2.4)	ns
Financial difficulties	16.7 (11.4)	7.4 (3.2)	ns

## Table 4. Mean scores of EORTC QLQ-C30 in 20 adult class 3 thalassemia patients with and without cGVHD following unrelated BMT

cGVHD = Chronic Graft versus Host Disease; ns = not significant.

Higher scores on the function scales and the global quality of life scale indicate better functioning: results were summarized as very poor (0-20), poor (21-40), intermediate (41-60), good (61-80) and very good (81-100). Higher scores on the symptom scales reflects problems: results were summarized as none to slight (0-29), moderate (30-69) and severe (70-100)

## DISCUSSION

#### **Clinical Outcome**

In our study, the outcome of unrelated BMT in adult class 3 thalassemia patients was similar to that reported in the literature for the same risk class transplanted from an HLA identical sibling donor [12,13]. Rejection and mortality rates were 3% and 22.6%, respectively. Seventy-seven percent of the patients are alive with sustained engraftment of donor hematopoiesis and a projected thalassemia-free survival of 74%. A higher incidence of deaths was observed in the group of patients conditioned with the protocol including Thiotepa compared to patients conditioned with BU-CY alone. Although this difference was not significant, a conditioning regimen with three drugs (BU-TT-CY) is likely to be too toxic for this category of patients. A recent report on a group of class 3 thalassemia patients, aged less than 17 years and transplanted from HLA identical siblings, has demonstrated the efficacy of a new and less aggressive conditioning regimen [20]. It can be postulated that this new regimen may also improve the outcome of unrelated BMT in adult class 3 thalassemia patients.

The incidence of acute and chronic GVHD in our patient sampling was 32% and 29% respectively. Rejection occurred only in one case. This relatively low incidence of rejection and GVHD can probably be attributed to the careful immunogenetic selection of the donor/recipient pairs [9-11].

Although this study was performed on a relatively small cohort of patients, the data obtained seem to suggest that BMT from unrelated donors in adult class 3 thalassemia patients may be able to achieve a success rate comparable to that offered using HLA-identical family donors.

#### **Communication and Motivation Factors**

The patients were asked to remember emotional experience and their motivation to undergo BMT. Exploration of communication strategies between physician and patient-relatives may lead to a better understanding of factors playing a role in transplantation and its general outcome. In fact, patients who are satisfied with communications are more capable of coping with treatment-related stress [21]. The importance of imparting individually tailored information has been emphasized by several authors, [22-24] but so far no study has been performed on thalassemia patients to investigate communication and motivation factors and patient satisfaction for BMT. All patients reported that they had received sufficient information about the high risk of mortality. According to all patients, the information received had been clearly presented particularly that concerning the quality of life expectancy and the cure rate. However, in 13% of patients some information had been perceived as not completely exhaustive and another 21% of patients felt that they had not been able to ask questions about their transplant. These data suggest that communication between the physician and patient-relatives could be improved. Particular attention should be paid to the patient's level of understanding and preferences.

By exploring motivation factors, we found that the patients of this study were satisfied with the choice to undergo BMT. The thalassemia patients enrolled in our study were all strongly motivated and felt to be protagonists of the choice to undergo BMT. None of the patients regretted their choice and would make the same decision if reliving the experience. Only 33% of the patients did not report the need to discuss the decision with their relatives. Nevertheless, social and family support were recognized as a valuable coping resource [5].

#### Quality of Life

Good QoL is a state of physical, psychosocial well being, in which the individual is able to perform everyday activities and report satisfaction with daily function [19]. Numerous reports have directly focused on the QoL of patients with a diagnosis of cancer undergoing BMT or thalassemia mayor patients treated with chronic blood transfusions and iron chelation therapy [5,18,19,25-30]. However, no studies have been performed on the QoL in transplanted thalassemia patients and no data exist on their "return to normal" following BMT. QoL is recognized as an important aspect of BMT, particularly in adult class 3 patients with an increased mortality risk and a worse outcome of the transplant procedure [12,13]. The mean global QoL in our patient sampling was good (mean 79.5) and 67% of patients enjoyed

a very good QoL. In a previous report, Arboretti et al. investigated QoL in a group of 867 thalassemic patients treated with parental iron-chelation therapy and blood transfusion: the perceived QoL was high in 40% of respondents, moderate in 39% and fair or poor in 21% of patients [26]. However, these data were not comparable with ours because of the different questionnaire used to evaluate QoL. In the group of patients that developed chronic GVHD, the global QoL was 65.3 whereas in patients without GVHD it was 84.3, and this difference was statistically significant. All functions were lower in the group of patients with chronic GVHD. In particular, significant difference was obtained for the physical scale. Additionally, pain and insomnia symptoms were significantly higher in the chronic GVHD group (Table 4). These data confirm that chronic GVHD has an important role in reducing the posttransplantation global QoL. Other reports of long-term transplant survivors describe persisting fatigue as a frequent finding, capable of impeding normal functioning [27]. In our study, fatigue was reported as moderate in the group of patients with GVHD (mean 38.9) and as slight (mean 20.4) in patients without it. When compared with data reported by Epstein et al., [19] who administered EORTC QLQ-C30 to a sample of 50 consecutive patients with different hematological diseases at 100 days after related or unrelated BMT, our results showed higher statistical performance on global quality of life (mean 79.5 vs. 50.5) and on all function scales (p < 0.05). Additionally, the symptoms observed in our study were significantly lower than those reported in Epstein's study (p < 0.05). No significant differences in global quality of life (mean 79.5 vs. 85) and other function scales were found when comparing our results to those of Worel et al. who assessed EORTC QLQ-C30 in 106 long term allogeneic or autologus stem cell transplanted patients alive and in complete remission after two years [18].

#### The Ethical Issues

When any medical treatment is proposed it must be considered whether the proposal therapy is effective, futile or even dangerous. The choice is usually not easy, particularly for two reasons: 1 - the definition of each concept lacks widespread consensus; 2 - the boundary between concepts is not clear and often there is a grey zone where it can be very hard to take a decision in a specific case. From 1987 to 1996, the concept of "medical futility" was debated in the medical community with vehemence [31].

Heated discussions gave shape to two big groups of thought: the first emphasized the responsibility of the doctor in the definition of the concept; the second, on the contrary, pointed out the role of the patient to decide about futility. The discussion about futility captures the central point of the relationship: the struggle between freedom and responsibility. Those who outline the responsibility of the physician attempt to define medical futility in quantitative terms (for example, if there have been no successes in 100 consecutive cases, the treatment must be considered futile) or through empirical data such as the threshold, expressed in terms of the physician's prediction of the chance of survival [32]. Other authors instead start to focus on the issue of who has the right or, better said, the power to decide whether medical care is futile. These professionals assume that futility is not an objective entity, but that it must be determined in the light of the subjective view and goals of the patients. Veatch, for example pointed out that, though physicians may be more qualified than patients in technical judgment making about medical treatments, they have no particular

expertise in decision making about such subjective matters as futility [33]. Some researchers speak about the concept of "futile value" when the goal is not worthy of achieving, even when the therapy could have some physiological effect [34]. In fact, the advancement of modern medicine, particularly with the introduction of controlled clinical trials, has led to the distinction between benefits and effects and is, by this time, determinate [35].

Furthermore, the treatment could be detrimental. The majority of therapies are in some way dangerous. In fact, the beneficial effects of all drugs and therapeutic devices are counterbalanced by undesired consequences. Therefore, the physician and the patient need to speak about the appropriateness of the cure. The physician must be able to balance beneficence and no maleficence when considering the feasibility of treatment. Nevertheless, definitions of futility based on the patient's values and desires alone could be perilous. In fact, it is not possible to classify the utility of a given therapy referring only to the wish of the patient. At the same time, doctors are responsible for careful assessment of the benefits and burdens. In the past, little or no credit was given to the will of the patient. This behaviour was called paternalism and was the most important feature of the relationship in the past. It was based on the moral authority of the physician, rather than his superiority. Albert Jonsen gave an evocative portrait of the doctor: «He was gentle but firm, honest but discreet, a thoroughly trustworthy confidant. He expended energy and time, often compensated only with gratitude. He practiced alone, his medical judgement unfettered by any supervisor. The doctor's ethics, even more than his power to cure, made him, if not an American folk hero, at least a paradigm of moral probity» [36]. The origin of paternalism can be sought in Hippocratic thought: «The doctor-philosopher is like God, and between medicine and wisdom there is no difference». The dramatic shift toward patient self-determination in recent times has almost certainly gained much of its force from society's backlash against physician paternalism. Furthermore, philosophical and political concerns about individual rights and respect elevated the principle of autonomy to a position in ethics that it had not previously held. Today ethics and the law give primacy to patient autonomy, defined as the right to be a fully informed participant in all aspects of medical decision making and the right to refuse unwanted, even recommended and life saving, medical care. The informed consent is one of the most important features, arising as a consequence in twentieth century health care. Unfortunately, both patients and physicians alike confuse the signing of the form with the reality of giving informed consent. Too often, physicians think that they are receiving consent by obtaining the patient's signature. Contemporarily, the patient lets the physician have his way without understanding well what is happening. So, there may be a serious lack of understanding between them. The misunderstanding is unacceptable because the consent is a cornerstone of ethical medicine. It is particularly unacceptable when the physician is faced with clinical problems that can have serious consequences on the health of the patient, such as is the case in BMT.

The central ethical issue of BMT in a non malignant disease can be framed in terms of a conflict between two fundamental principles of medical ethics, i.e., non-maleficence and beneficence. In the light of these two principles, doctors and medical investigators must avoid harming the patient (*primum non nocere*), or, at least, make sure that any harm incurred by the patient can be compensated by a benefit that is proportionately greater than the harm inflicted. The relation between the two principles is a debated one in ethics. The harm of unrelated BMT in adult class 3 thalassemia patients may be significant: about 22.6% of these patients may die within a year from the time of transplantation. On the other hand, the

benefits of BMT are no less real (disease free survival = 74%). Given the fact that many situations present in medicine where benefits cannot be achieved without at the same time inflicting harm, the question arises as to who will be responsible for the decision and according to what criteria can a fair balance of benefits and burden be struck. Traditionally, the physician has been held responsible for the decision. According to the new "liberal" paradigm, decisions in medicine are based on respect for patient's autonomy and its derivative rules. As a result, autonomous actions are analysed "in terms of normal choosers who act: 1 – intentionally; 2 – with understanding; 3 – without controlling influences" [37]. When seen in relation to the case of thalassemia patients, the principle of autonomy aims at restoring the patients' ability to choose freely. At the same time, autonomy must be seen within the larger relational context of the physician-patient encounter. When defined in terms of a healing relationship, such an encounter bears the phenomenological mark of a covenantal exchange rather than a contract: it draws on the physician's virtues and his ability to ensure that the patient he encounters in an obvious position of power maintains the dignity of a person [38]. Thus the issue is not simply "to cure" but "to take care" of the patient as a whole person [39]. Within this relationship the patient wants to understand what the doctor is doing, why he is doing it, what the alternatives may be, and what the future will mean for the things he values most. In the end, one might say that although the decision is ultimately the patient's, the physician has the responsibility to probe whether the patient is suitable for BMT. All efforts should be made by the physician to raise the patient's awareness of all the variables involved in the decision. One must be mindful of Hans Jonas' warning to make the patient's consent autonomous and conscious [40]. On this ethical ground it might be reasonable to suggest that unrelated BMT represents a possible cure in adult class 3 thalassemia patients without a sibling donor.

## CONCLUSIONS

This study reports the clinical outcome of unrelated BMT in a cohort of adult risk class 3 thalassemia patients and contemporarily investigates communication strategies, motivation factors, quality of life and ethical issues using an observational method. The lack of a control population for comparison of QoL was compensated by the reference values of the EORTC questionnaire. Overall, the patient sampling is relatively small even though the 31 enrolled patients represent the total number of adult class 3 thalassemia patients transplanted from unrelated donors in Italy over the past 10 years. Another limitation is that patients were asked to recall previous experiences and that the administered items cannot exhaustively explore all the motivation and communication factors involved. Because the QoL assessment was cross-sectional and not prospective, only one assessment at one point in time was undertaken.

A multi-disciplinary approach (medical, psychological and ethical), may further help to resolve the dilemma that surrounds the choice of unrelated BMT in a chronic non malignant disease [41-43].

The Authors agree with Mazur in identifying at least six dimensions in the decisionmaking process: [44]

- 1. *Harms*. Any adverse outcomes associated with drugs or medical strategies, including nonintervention, intervention at a later time and watchful waiting;
- 2. Benefits. With a basic distinction between survival and quality of life benefits;
- 3. Scientific evidence. Science supplies the proof or basis for any clinical decision;
- 4. *Clinical experience*. When there are no randomized clinical trials available, these clinical experiences form the basis of care decisions.
- 5. *Estimation experience*. The ability to estimate the chance of an event occurrence. It may be related to science evidence, clinical experience and the art of medicine.
- 6. *Psychological experience involving estimates.* It includes the basic frameworks that individual physicians use in making their estimates related to their individual patients.

To these six dimensions, the Authors would add a seventh dimension:

7. Conscious consent. Often the physician provides a multitude of information but the patient loses track of the central point; sometimes there is enough information but the physician is not sufficiently involved and so forth. Hence, the most important goal of the patient-physician relationship is not the information but rather the understanding. The supply of information is a condition to obtain awareness but it does not coincide. One can give consent without understanding, while the aim is comprehension. Understanding means information, empathy between physician and patient and creation of the most comfortable possible situation. The pursuit of conscious consent is paramount and the physician needs to be educated to learn that respect of the patient's freedom is part of the goal of his profession and consciousness is the only ground where freedom can be respected.

## REFERENCES

- [1] Rund D, Rachmilewitz E. ß-Thalassemia. N Engl J Med. 2005; 353: 1135-46.
- [2] Borgna Pignatti C, Rugolotto S, De Stefano P, Zhao H, Cappellini MD, Del Vecchio GC, et al. Survival and disease complications in thalassemia major. *Ann N Y Acad Sci.* 1998; 850: 227-231.
- [3] Ehlers KH, Giardina PJ, Lesser ML, Engle MA, Hilgartner MW, et al. Prolonged survival in patients with Beta Thalassemia major treated with deferoxamine. *J Pediatr*. 1991; 118: 540-545.
- [4] Modell B, Khan M, Darlison M. Survival in beta-thalassemia major in the UK: data from the UK Thalassemia Register. *Lancet*. 2000; 355: 2051-2052.
- [5] Atkin K, Ahmad W. Living a "normal life": young people coping with thalassemia major or sickle cell disorder. *Soc Sci Med.* 2001; 53: 615-626.
- [6] Thomas ED, Buckner CD, Sanders JE, Papayannopoulou T, Borgna-Pignatti C, De Stefano P, et al. Marrow transplantation for thalassemia. *Lancet*. 1982; 2: 227-229.
- [7] Lucarelli G, Galimberti M, Polchi P, Angelucci E; Baronciani D, Giardini C, et al. Bone marrow transplantation in patients with thalassemia. *New Engl J Med.* 1990; 322: 417-421.

- [8] Anasetti C, Etzioni R, Petersdorf EW, Martin PJ, Hansen JA. Marrow transplantation from unrelated volunteer donors. *Ann Rev Med.* 1995; 46: 169-179.
- [9] Sasazuki T, Juji T, Morishima Y, Kinukawa N, Kashiwabara H, Inoko H, et al. Effect of matching of class I HLA alleles on clinical outcome after transplantation of hematopoietic stem cells from an unrelated donor. N Engl J Med. 1998; 339:1177-1193.
- [10] Petersdorf EW, Gooley TA, Anasetti C, Martin PJ, Smith AG, Mickelson EM, et al. Optimizing outcome after unrelated marrow transplantation by comprehensive matching of HLA class I and II alleles in the donor and recipient. *Blood.* 1998; 92:3515-3520.
- [11] La Nasa G, Giardini C, Argiolu F, Locatelli F, Arras M, De Stefano P, et al. Unrelated donor bone marrow transplantation for thalassemia: the effect of extended haplotypes. *Blood*. 2002; 99: 4350-4356.
- [12] Lucarelli G, Giardini C, Angelucci E. Bone Marrow Transplantation in thalassemia. In: Winter JN ed. *Blood stem cell transplantation*. Kluwer Academic: Boston, 1997, pp 305-315.
- [13] Lucarelli G, Clift R, Galimberti M, Polchi P, Angelucci E, Baronciani D, et al. Marrow transplantation for patients with thalassemia: results in Class 3 patients. *Blood*. 1996; 87:2082-2088.
- [14] Angelucci E, Baronciani D, Lucarelli G, Baldassarri M, Galimberti M, Giardini C, et al. Needle liver biopsy in thalassemia: Analyses of diagnosic accuracy and safety in 1184 consecutive biopsies. *Br J Haematol.* 1995; 89: 757-761.
- [15] Shulman HM, Sullivan KM, Weiden PL, McDonald GB, Striker GE, Sale GE, et al. Chronic graft-versus-host syndrome in man. A long-term clinico-pathologic study of 20 Seattle patients. *Am J Med.* 1980; 69: 204-217.
- [16] Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. For the European Organization for Research and Treatment of Cancer. Study Group on Quality of Life. The EORTC QLQ-30: a quality of life instrument for use in international trials in oncology. *J Natl Cancer Inst.* 1993; 85: 365-376.
- [17] Kaplan ER, Meier P. Non parametric estimation from incomplete observation. J Am Stat Assoc. 1958; 53: 457-481.
- [18] Worel N, Biener D, Kalhs P, Mitterbauer M, Keil F, Schulemberg A, et al. Long-term outcome and quality of life of patients who are alive and in complete remission more than two years after allogeneic and syngenic transplantation. *Bone Marrow Transplant*. 2002; 30: 619-626.
- [19] Epstein JB, Phillips N, Parry J, Epstein MS, Nevill T, Stevenson-More P. Quality of life, taste, olfactory and oral function following high-dose of chemotherapy and allogenic hematopoietic cell transplantation. *Bone Marrow Transplant*. 2002; 30: 785-792.
- [20] Sodani P, Gaziev D; Polchi P, Erer B, Giardini C, Angelucci E, et al. A new approach for bone marrow transplantation in class 3 thalassemic patients aged less than 17 years. *Blood.* First edition paper, pre-published online, March 23, 2004.
- [21] Schofield PE, Butow PN, Thompson JF, Tattersall MHN, Beeney LJ, Dunn SM. Psychological responses of patients receiving a diagnosis of cancer. *Ann Oncol.* 2003; 14: 48-56.
- [22] Sardell AN, Trierweiler SJ. Disclosing the cancer diagnosis. Procedures that influence patient hopefulness. *Cancer*. 1997; 72: 3355-3365.

- [23] Ptacek JT, Eberhardt TL. The patient-physician relationship. Breaking bad news: a review of the literature. *JAMA*. 1996; 276: 492-496.
- [24] Rutter DR, Iconomou G, Quine L et al. Doctor-patient communication and outcome in cancer patients: an intervention. *Psych Health*. 1996; 12: 57-71.
- [25] Caro J, Ward A, Craig Green T, Huybrechts K, Arana A, Wait S, et al. Impact of Thalassemia major on patients and their families. *Acta Haematol*. 2002; 107: 150-157.
- [26] Arboretti R, Tognoni G, Alberti D. Pharmacosurveillance and quality of care of thalassaemic patients. *Eur J Clin Pharmacol*. 2001; 56: 915-922.
- [27] Hjermstad MJ, Holte H, Evensen SA, Fayers PM, Kaasa S. Do patients who are treated with stem cell transplantation have a health-related quality of life comparable to the general population after 1 year? *Bone Marrow Transplant*. 1999; 24: 911-918.
- [28] Hjermstad MJ, Loge JH, Evensen SA, Kvaloy SO, Fayers PM, Kaasa S. The course of anxiety and depression during the first year after allogenic or autologous stem cell transplantation. *Bone Marrow Transplant*. 1999; 24: 1219-1228.
- [29] Zittoun R, Suciu S, Watson M, Solbu G, Muus P, Mandelli F, et al. Quality of life in patients with acute myelogenous leukemia in prolonged first complete remission after bone marrow transplantation (allogeneic or autologous) or chemotherapy: a crosssectional study of the EORTC-GIMEMA AML 8A trial. *Bone Marrow Transplant*. 1997; 20: 307-315.
- [30] Sutherland HJ, Fyles GM, Adams G, Lipton JH, Minden MD, Meharchand JM, et al. Quality of life following bone marrow transplantation: a comparison of patient reports with population norms. *Bone Marrow Transplant*. 1997; 19: 1129-1136.
- [31] Helft PR, Siegler M, Lantos J. The Rise and Fall of Futility Movement. *JAMA*. 2000; 343: 293-296.
- [32] McCrary SV, Swanson JW, Youngner SJ, Perkins HS, Winslade WJ. Physicians' quantitative assessment of medical futility. J Clin Ethics. 1994; 5: 100-105.
- [33] Veatch RM. Why physicians cannot determine if care is futile. *J Am Geriatr Soc*. 1994; 42: 871-874.
- [34] Mohindra RK. Medical futility: a conceptual model. J Med Ethics. 2007; 33: 71-75.
- [35] Schneiderman JL, Jecker NS, Jonsen AJ. Medical Futility: Its Meaning and Ethical Implication. Ann Intern Med. 1990; 112: 949-954.
- [36] Jonsen AJ. The Birth of Bioethics. Oxford University Press, New York-Oxford, 1998; 3-4
- [37] Beauchamp TL, Childress JF, eds. *Principles of Biomedical Ethics*. New York, NY: Oxford University Press; 1994: 123.
- [38] Engel E. The clinical Application of the Biopsychosocial Model. Am J Psychiatry. 1980; 137: 535-544.
- [39] Sulmasy D. What's so Special About Medicine. *Theor Med*. 1993; 1: 27-42.
- [40] Jonas H, ed. Philosophical Essay. From Ancient Creed to Technological Man. Chicago, IL: The University of Chicago Press;1974.
- [41] Lucarelli G, Weatherall DJ. For debate: bone marrow transplantation for severe thalassemia. *Br J Haematol*. 1991; 78: 300-303.
- [42] Weatherall DJ. The treatment of thalassemia: slow progress and new dilemmas. *New Engl J Med.* 1993; 329: 877-879.
- [43] Giardini C. Editorial. Ethical issue of bone marrow transplantation for thalassemia. Bone Marrow Transplant. 1995; 15: 657-658.

[44] Mazur DJ. *The new medical conversation*. Rowman & Littlefield Publishers, Boston, 2003; 99-100.

Chapter 6

# PERSONAL MEANINGS AND HEALTH CARE DECISION MAKING – THE CASE OF TOTAL KNEE REPLACEMENT<sup>\*</sup>

 F. Toye<sup>1</sup>, J. Barlow<sup>2</sup>, C. Wright<sup>3</sup> and S.E. Lamb<sup>4</sup>
 <sup>1</sup>Physiotherapy Research Unit Nuffield Orthopaedic Centre, Windmill Road, Oxford, OX3 7LD
 <sup>2</sup>Interdisciplinary Research Centre in Health, Coventry University, School of Health & Social Sciences, Whitefriars Building, Whitefriars St., Coventry, CV3 1FB
 <sup>3</sup>School of Health Sciences, University of Birmingham, Birmingham, B15 2TT
 <sup>4</sup>. Warwick Clinical Trials Unit, Warwick University, Coventry, CV4 7AL

## ABSTRACT

There seems to be a mismatch between the severity of knee symptoms and the decision to undergo (total knee replacement) TKR. Personal meanings are important because decisions regarding need for TKR do not seem to be explained by symptoms alone. This study explores patients' personal meanings of knee osteoarthritis (OA) and TKR. 18 semi-structured interviews were conducted with a purposive sample of respondents who were listed for TKR at one specialist orthopaedic hospital. Data were analysed using Interpretive Phenomenological Analysis. Results suggest that the decision to undergo TKR is not related to symptoms alone, but to personal meanings. Some of these personal meanings may not be useful in accurately assessing need for TKR, and may result in mis-targeting of treatment.

This chapter focuses on the following findings:

1. Perceptions of need in relation to Levanthal's Self Regulatory Model.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Medicine and Health Care*, edited by Thomas E. Lynch, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

Differences in disease representations are important as they may explain some of the variability in decision-making regarding need for TKR. Findings support *identity, consequences, time-line* and *controllability* as important representations affecting the decision to undergo TKR.

- 2. Moral narratives in the construction of need for TKR. Priority status for TKR was related to a person's 'good virtue'. Undesirable traits such as: letting the disease 'beat you', having a 'bad attitude' towards the doctor, 'letting yourself' get worse, and not having a 'positive age attitude', were linked to lower perceived priority for TKR.
- 3. Gender differences in health care decision making There were some important gender differences found, that might explain the finding that women seem to be listed later in their disease course

## BACKGROUND

A person's perceptions of need for treatment are likely to be related to their decision to undergo treatment. However, the concept of need for TKR is unclear, and there may be a range of factors that contribute to the decision that a person needs a TKR. There is currently no agreed method in the UK for deciding who needs a TKR, and who should have priority (Dieppe et al.1999). Patients' perceptions of TKR have only been explored to a very limited extent. Woolhead et al. (2002) is the only study to have explored patients' views on priority for TKR (2002). At the time of this study, no other study had explored patients' perceptions of need for TKR, and factors that influence these perceptions.

This is of particular importance, not only in view of recent government policy to involve users in health-care decisions (National Health Service Plan 2000, Thomson et al. 2001), but also because patients' perceptions are likely to affect their use of TKR. The decision to undergo TKR is complex and likely to be qualitatively different for different people. For example, decisions regarding need for TKR do not seem to be explained by symptoms alone (Toye et al. 2006, Tennant et al. 1995, Hawker et al. 2001). Patients' construction of need may also vary in important aspects from need as perceived by their doctor (Lambert et al. 2000).

#### Mismatch between Symptoms and Utilisation of TKR

In a review of studies reporting prevalence of knee pain and osteoarthritis in older patients in the UK (Peat et al. 2001), the prevalence of knee pain in older adults in the UK ranged from 13% to 28%. Prevalence of knee pain associated with disability ranges from 8% to 15% (McAlindon et al. 1992, Tennant et al. 1995, O'Reilly et al. 1996, Urwin et al. 1998), and severely disabling symptoms related to x-ray changes affect 1.6% of adults over the age of 55 years (Tennant et al. 1995). Differences in estimates may partly be explained by differences in the questions used to screen for knee pain (O'Reilly et al. 1996). There are also significant differences between studies depending on the age and gender of the sample used. For example, in a large population based sample in the UK, Urwin et al. (1998) found that

23% of women between the ages of 45 to 64 had knee pain, compared to 35% of women over the age 75. Fewer men (21% and 27%) experienced knee pain. In the Urwin et al. study, disability associated with knee pain also increased with age. Despite differences in estimations of pain and disability related to knee osteoarthritis, the World Health Organisation (WHO) has indicated that knee osteoarthritis is a major cause of disability (Cooper et al. 2000).

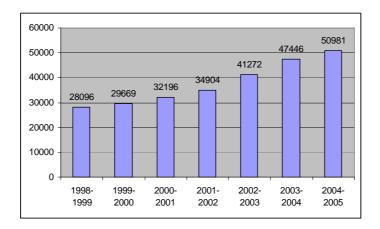


Figure 1. Utilisation of primary TKR (with and without cement) in the UK from 1998/9-2004/5 (information from the Hospital Episode Statistics 2007). www.hesonline.nhs.uk

Recent figures (figure 1) from Hospital Episode statistics suggest that the number of TKRs (with and without cement) performed in the UK National Health Service has increased dramatically since 1998/1999 (Hospital Episode Statistics 2007). A wide regional variation in the use of primary TKR in the UK has been reported both in and outside the UK (Williams et al. 1994, Peterson et al. 1992, Van Walraven et al. 1996, Coyte et al. 1997). This wide variation in utilisation of TKR cannot be explained by differences in prevalence of knee pain. It also seems that in the UK there is a mismatch between the number of TKRs performed and the number of people in the population with severe knee pain and disability (Tennant et al. 1995, Juni et al. 2003). This mismatch between supply and potential demand for TKR in the UK, is currently managed using a waiting list system. However, there are concerns that this system does not ensure that patients with the greatest need are treated first (British Medical Association 1998). In the UK, waiting time for surgery does not match disease burden (Harry et al. 2000), and there is a wide range of disease burden in those listed for TKR surgery (Roy and Hunter 1996, Harry et al. 2000, Kennedy et al. 2003).

Harry et al. (2000) not only found a large range in symptom burden, but also that there was no correlation between clinical priority and waiting time for hip and knee joint replacement in a UK hospital. This is supported by another UK survey of 97 patients waiting for lower limb surgery (40 of which were TKRs), which found that there was no significant difference in walking speed between urgent and non-urgent patients on the waiting list (Roy and Hunter 1996). There was also no significant difference in mobility, household activities, social activities, activities of daily living, pain and anxiety between urgent and non-urgent patients. This has also been found outside the UK. For example, in a prospective population-based study of 553 patients waiting for hip and knee replacement in Canada, Kelly et al.

(2000) found that pain and function as measured by the Western Ontario and McMasters Universities Osteoarthritis Index (WOMAC – Bellamy et al. 1988), and the Short Form-36 Health Survey questionnaire (SF36 – Ware et al. 1993), were not related to waiting time for hip and knee replacement surgery. Although it is important to consider that the dynamics of a surgical waiting list may be different outside the UK, it has been estimated that in Canada, orthopaedic waiting lists may overestimate the number of patients that need surgery by as much as 30% (Lewis et al. 2000).

In summary, there seems to be a large range of symptom burden in those waiting for joint replacement surgery. This may suggest that some patients do not actually have severe pain or loss of function at the time they are listed. It may be that some are strategically listed for surgery.

#### What is Health Care Need?

Health care decision making is complex and likely to be related to definitions of health care need. However, the definition of health care need can be problematic.

Clinicians and patients may describe need for health care in different ways. Clinicians may describe health care need as "a person's ability to benefit from health care" (Stevens and Raftery 1994 :14). If defined in this way, it implies that need only exists if there is currently a treatment available. However, it could be argued that a person still has a need irrespective of the availability of treatment. Stevens and Raftery (1994) distinguish need (ability to benefit from treatment) from demand, which they define as what people *want* to have. This is perhaps misleading as it implies that the patient may only *think* that they have a need, and therefore they may demand something *in error*, whereas the doctor defines true need (Good 1994). It could be argued that this denies the importance of the lay view in medical decision-making.

However, although clinicians often describe need as the ability to benefit, this definition cannot always be supported in practice, particularly as there is currently limited evidence related to predictors of outcome. The system of surgical waiting lists in the UK recognises that some patients have priority for treatment, and this priority may not necessarily reflect ability to benefit from surgery (British Medical Association 1998, NHS 2000). For example, we may feel that those with more pain or less function have a greater *need* for TKR, even if we do not know whether this group of patients would derive the most benefit (Fortin et al. 1999). Other factors may also be considered to determine relative need. For example, the amount of time that a person has been waiting for surgery is considered an important criterion in terms of government waiting list policy (British Medical Association 1998). Therefore, in terms of needs assessment, the definition of need as *the ability to benefit from TKR* is problematic.

When deciding on health care, it is also important to incorporate patient's perceptions of need for TKR. Bradshaw's distinction between different types of need (cited in Rogers et al 1999: 13), is useful as it recognises that need can be defined from patient or professional perspective. He proposes four types of health care need:

- 1. Normative need what the professional thinks a person needs;
- 2. Felt need what a person feels they need;

- 3. Expressed need what the person feels they need and actually seeks out;
- 4. Comparative need how a person's need compares with others.

This model is useful, as it recognizes not only the importance of patient and clinician perspective, but also the idea of relative priority. However, in terms of needs assessment, it does not consider what constitutes the different types of need. For example, normative need may be based on a range of factors, such as, professional evaluation of the person's burden of symptoms, the likelihood of benefit from treatment, the relative benefit expected compared to other patients, and even psychosocial factors. A needs assessment model needs to consider all of these factors.

A difficulty with defining the term 'need' is that some define need as the problem that is interfering with a person's life (For example – 'I *need* to get rid of my pain'), and others define it as the ability to benefit from a specific treatment (For example – 'I *need* surgery to get rid of my pain'). If need is defined as the problem (pain), then it is possible to have this need even if there is no solution (Callaghan 1991, Doyal and Gough 1991). If however, need is defined as the ability to benefit from treatment, a person must need *something* (Hadorn 1991a, Hadorn and Brook 1991, Sheaff 1996). There are practical and philosophical challenges associated with these definitions of need: If need is defined as the problem, then need is defined by the person with the problem; if need is defined as the treatment, then need is defined by the person who suggests the appropriate treatment. In terms of health care decision making, it may be more useful to consider both perspectives: Firstly, the symptom burden, which would include the effects of the health problem on the person's life. Secondly, the ability to benefit from a particular treatment, which would include the doctor's opinion on whether or not a treatment is likely to benefit the person. It is also likely to incorporate the patient's opinion regarding treatment outcome.

There is another important concept related to needs assessment that is often not considered – significant net benefit. It is clear that health care need is not necessarily the same as demand. However, it is less clear what distinguishes need from benefit. It could be argued that health care need is more than just the ability to benefit from an intervention, and that *necessary* treatment should provide *significant net benefit* (Hadorn and Brook 1991). To define health care need we must therefore answer two questions: firstly, who benefits from treatment; secondly, what is *significant net benefit*? The answer to the first question is not known for every health situation. For example very little is known about who benefits most from TKR (Dieppe et al.1999). However, although difficult, it may be possible, at some time in the future, to answer the first question, and find out who benefits from treatment, through targeted research on outcome predictors. The second question – What is *significant* net benefit? - is likely to be more problematic.

Hadorn (1991b and 1991c) argues that it is possible to determine significant net benefit by developing objective criteria that will distinguish treatment that is *necessary*, from treatment that is *beneficial*. Sheaff (1996) also argues that to be defined as having a need, a person must satisfy certain criteria. However, others argue that the search for *objective* criteria to define need may be misplaced, and that the concept of health care need is *subjective* (Callaghan 1991). These would argue that significant net benefit is socially constructed, and therefore a matter of qualitative judgement. Kaplan (1991) argues that although the definition of health care need is subjective, rather than reflecting objective biomedical criteria, some definitions would be almost universal. For example, most people would want to be free of pain, and it seems unlikely that any individual would want to undergo any procedure that did not carry significant net benefit. The difficulty is that in any definition, significant net benefit depends upon the perspective of the person who is defining it.

## Psychosocial Factors Determining Utilisation of TKR

Assessing need for TKR can be difficult, as there is limited professional consensus related to need for TKR, and a lack of concordance between x-ray findings and symptoms (Toye 2004). It is clear that factors other than symptom burden may affect a person's decision that they need a TKR. For example, Hudak et al. (2002) found that people decided that they did not need a TKR for several different reasons: they explained their symptoms as an inevitable part of ageing; they felt that there were others worse off than themselves; they were waiting for the doctor to suggest TKR. In view of this, the assessment of need for TKR is not straightforward. Evidence suggests that people with similar levels of pain and function vary in their decision of whether they need or want a TKR. However, evidence for the psychosocial factors contributing to this differential utilization of TKR is limited. It is likely that psychosocial factors have a role in health care decision making. In a large Canadian, population-based survey, Hawker et al. (2001) found that there was no significant difference in pain and function between those who had, and those who had not been recommended a joint replacement. This may be because many people with knee pain may not feel that they need a TKR, or may be unwilling to undergo surgery. It may also be that the health care professional will reach a different decision depending on other factors. Zola (1973) found that the doctor treated different cultural groups differently, even though there was no objective difference in the symptoms.

Several possible psychosocial factors have been suggested. It may be that there are gender differences in the utilization of joint replacement, although the mechanism for this is unclear. Kennedy et al. (2003) found that women listed for TKR had higher disease severity than men. It may be that doctors are less likely to initiate discussions about TKR with women, or vice versa. Hawker et al. (2000) found that although women with potential need for joint replacement surgery were less likely to be on the list for TKR, it was not willingness to undergo TKR that affected this, but whether or not they had discussed TKR with their doctor.

In a volunteer sample of people with hip or knee osteoarthritis, Dexter and Brandt (1993) found that, of those with lower function, those who had a high school education, were 4 times more likely to be under the care of the physician for OA (p<0.003) than those without a high school education. Hawker et al. (2001) found that only 8%-15% of those who had potential need for TKR were definitely willing to undergo surgery. In multivariate analysis, greater willingness to have surgery was related to being younger, and having spoken to a physician about TKR. There are some difficulties with interpreting the findings of Hawker et al. (2001). However, need for TKR was defined by a cut-off score that this might not represent need for TKR. Hawker et al. (2002) also found that although lower income and less education were independent predictors of greater need for TKR, a person's willingness to undergo surgery was not explained by their income or education, but by whether a person had discussed TKR

with their doctor. These studies suggest that health professionals may be more likely to discuss treatment with patients with a higher income or more formal education.

There may also be ethnic or cultural differences in perceptions of TKR. In a survey of 596 elderly male patients with hip or knee pain recruited from primary care in the USA (Ibrahim et al. 2002), black people were less likely to have friends or family with a total joint replacement, or to have a "good understanding" of total joint replacement. They were more likely to believe that hospital care after surgery would last more than 2 weeks, and to expect both moderate to extreme pain, and moderate to extreme difficulty with walking following TKR. Cultural factors may explain why white men are 3 to 5 times more likely to undergo TKR than black men are (Wilson et al. 1994). In a cross-sectional population study of 829 American women aged 40-53 years with osteoarthritis, x-ray changes were a more sensitive predictor of pain among African-American women than for Caucasian women. This may also suggest cultural differences in pain experience, and triggers that make a person seek health care (Zola 1973, Lachance et al. 2002). In the UK, inequalities in health care have been recognized since the Black Report was published in 1980 (Whitehead et al. 1992), and continue to be the focus of government policy. It may be that socio-economic factors such as race, gender, education, income have an effect on the decision to perform or to undergo surgery, and that there is an inequality in the provision of TKR.

Social factors are likely to have a role in health care decision making. Some argue that our social network does not just influence decision-making, but actually shapes the decisions that we make (Blumer 1969). Rao et al. (1998) found that in a group of patients with various types of arthritis, family and friends as well as doctors were the most important source of information about treatment. Friedson (1970) proposed that people use the social network to determine whether they are ill, and to decide what to do. It could be argued that society drives the process of categorising and treating illness (Garro and Mattingly 2000), and that therefore our social network has a large effect on what we decide to do when we are ill (Friedson 1970, Good 1994, Pescosolido 1992). If many people agree about a pattern of symptoms, its explanation and treatments, it has been described as an "illness prototype" or "folk illness" (Helman 2000:86). It could be argued that osteoarthritis is an example of an illness prototype, where decisions about classification of disease and subsequent behaviour are socially determined (Turner 1992). This does not deny the biophysical reality of disease (Williams 1999). The value of taking this sociological perspective, is that it can highlight that even though we experience disease as if it is an objective category, disease categorisation is inseparable from the society. This may help to explain why people with a similar level of symptom burden, may either decide that they need a TKR, or not (Tennant et al. 1995).

In summary, the utilisation of TKR is not related to severity of symptoms alone. It seems that factors other than symptom burden may affect a person's decision that they need a TKR. In view of this, health care decision making in the case of TKR is not straightforward. There is some professional consensus that pain, function and x-ray findings determine 'need' for TKR. However, psychosocial factors also seem to influence decisions regarding TKR.

# THE STUDY

## Sample

Thirty-four patients listed for TKR at one specialist orthopaedic hospital who scored below average on *both* the pain and the functions domain of the Western Ontario and McMasters Universities Osteoarthritis Index (WOMAC – Bellamy et al. 1988), were sent a letter inviting them to take part and were given details of the research. The WOMAC is widely used in knee osteoarthritis trials, and has been extensively validated in this patient group (Bellamy 2000, McConnell). The average score was taken from the results of a postal questionnaire sent to all patients on the waiting-list for TKR and is similar to average scores found in other waiting-list samples (Toye et al. 2006). This sample was chosen for several reasons:

- 1. In qualitative research, the aim of sampling is to *purposively* choose patients who demonstrate a particular characteristic (Charmaz 2000, Silverman 2001). Other research has explored the perceptions of patients with severe knee symptoms who were not willing to undergo joint replacement surgery (Hudak et al. 2002). No study has explored the perceptions of patients listed for TKR who have lower than average disease burden as measured by validated disease specific measures.
- 2. In the context of demand management for TKR, it would be useful to explore the construction of need and priority in those with a lower disease burden, who may fall below a certain threshold of clinical need, as determined by validated measures of disease burden. This group includes patients who might not be on the waiting list for TKR surgery, or who would be given a low priority if quantitative measures were used to determine access to, or priority for TKR. In interview data regarding pain and functional levels in patients with patellofemoral osteoarthritis, Campbell et al. (2003) found less than 50% concordance between WOMAC pain and function data, suggesting that quantitative measures of function may not always identify those with functional loss.
- 3. Those with lower than average symptom burden, were likely to provide an insight into factors other than pain and function, that were considered to be important in patients' perceptions of need.

Six women (aged 60-76) and 12 men (aged 54-77) were interviewed. Five out of six women were widowed and living alone and one was living with her husband. Ten men were living with their wives, one was a widower, and the other was single. The group consisted of people from a wide range of occupations, including manual work (e.g. cleaning, factory work), and professional occupations (e.g. teaching and engineering), as well as a range of other employments (e.g. footballer, farmer, cleaner, mechanic, office worker). To protect the anonymity of the participants, each was given a pseudonym. Each interview lasted for approximately 1 to  $1\frac{1}{2}$  hours and was audio-taped and transcribed with consent from each participant.

#### Interviews

Patients were sent a letter inviting them to take part in an interview, and were given details of the research aims. Ethical approval was obtained from the local research ethics committee.

Semi-structured interviews were used to explore the concept of need and priority for TKR surgery. This is a useful approach in research adopting a phenomenological philosophy (Bowling 1997, Smith 1995), where the aim is to explore personal meanings. This type of interview allows the person to "to tell his or her own story" (Smith 1995: 12). To ensure that the interview covered areas of interest, an interview schedule was used (figure 2) to give some structure. However, the schedule was used only as a prompt, and patients were given freedom to discuss areas that they felt were important. This semi-structured format allowed the flexibility to follow leads opened by patients, and not to be constrained by structured questioning.

## Interpretive Phenomenological Analysis (IPA)

Interpretive Phenomenological Analysis (IPA) was used to explore the data from semistructured interviews (Smith 1995, Smith 1996a, Smith et al. 1999). A strength of IPA is that it is influenced by both Phenomenology and Constructivism (Smith 1996a, Smith et al. 1999). The aim of Phenomenology is to "explore in detail the participants view of the topic under investigation" (Smith et al. 1999:218). The researcher aims to get close to the personal world of the participant, and to gain an "insider perspective" (Smith 1996a: 264) of the phenomenon - in this case, the patients' perceptions of need and priority for TKR. Husserl (1970) argued that the social world is not "out there" or distinct from a person's interpretation of it (Gubrium and Holstein 2000:489), but that interpretation actually constitutes the social world. Therefore, to understand human action, we need to understand personal meanings (Schwandt 2000). A person is likely to act in a certain way because of the personal meaning that the situation has for them. For example, Hudak et al. (2002) showed that a person's decision to undergo TKR surgery is likely to be related to personal meanings about need and priority for surgery.

However, a criticism of Phenomenology is that meaning is not directly accessible; to find meaning requires *interpretation*. It could be argued that we cannot assume one interpretation is superior, or more "true" than another (Schwandt 2000:195). Some argue that a difficulty with interpreting interview narratives, is that narrative is not just a *medium* for presenting reality about personal meanings, but that meaning is actually constructed within the interaction and will therefore be context specific (Riessman 1993, Silverman 2000). Bruner (1986) argues, that a person telling a story will not always be specific about their meaning, but will leave the meaning open to interpretation, thus forcing the listener to interpret what he or she hears (Bruner 1986, Denzin and Lincoln 2000). Bruner refers to this process as - "subjunctivising reality" (Bruner 1986:26) – it can enable a person to *mean* more than, or even to mean the opposite to what is said, thus making interpretation problematic. A strength of IPA is that it is also influenced by Constructivism and therefore recognizes that meaning is negotiated during the interaction (Smith 1996a).

There are always likely to be conceptual difficulties in interpreting interview data. Patients are likely to have beliefs about the aims of the research that will structure the content of the interview. Kleinman (1988) argued that a person's interpretation of the purpose of the interview would influence the telling of the story. This difficulty of interpretation arises even in a highly controlled experiment; this is recognised in the use of blinding in scientific research. A person's interpretation of the interview agenda is likely to be related to the context of the interview. For example, some patients may have believed that the interview would have a bearing on where they were positioned on the waiting list. Other patients may have agreed to be interviewed in the hope of gaining access to physiotherapy services prior to TKR. In this study, I interviewed patients at home if possible, so that they would feel more at ease. I also made every attempt to create rapport between the patient and myself, and to distance myself from my role as a health professional. I made it clear that agreeing to be interviewed would not affect a persons' access to treatment.

A feature of IPA is that the first steps of analysis begin early in the research process with initial data coding. In IPA, analysis does not follow data collection, but is simultaneous to data collection. This first stage of coding is also referred to as 'open', 'line by line' or 'action' coding (Charmaz 2000:515). Analysis began early with simultaneous transcription of each interview in full. Transcripts were listened to and read several times in order to become familiar with the accounts, and transcripts were coded line by line. Each interview was transcribed and coded prior to conducting the next interview. This type of initial coding is important to theory development but also helps the researcher remain close to the data (Charmaz 2000). The act of coding a transcript line by line helps the researcher to challenge any a priori assumptions, and to redefine categories based on what is found in the data. One method of challenging a priori assumptions used in this study was "deviant-case analysis" (Seale 1999: 75). This is a method that specifically looks for cases that challenge arising themes. A *deviant case* is found when the data do not fit with developing theory. Searching for deviant cases helps the researcher to challenge assumptions about the data (Seale 1999). Seale argues that during the research process, the researcher should expect to change his or her mind about something.

The second stage of data coding used in this study was to make a list of emerging themes and look for connections between them, as suggested by Smith et al. (1999). After the initial coding, a list of key words and themes was recorded in order to "capture the essential quality" of findings in the interview data (Smith et al. 1999:221). The aim of this second stage of coding is to make connections between segments of data and to develop arising theory (Smith 1996a, Smith et al.1999, Charmaz 2000). This is comparable to the Constant Comparison used in Grounded Theory (Charmaz 2000, Strauss and Corbin 1998). Constant Comparison describes a process whereby the researcher constantly compares first and second stage codes, both from one individual, and from different individuals, in order to develop theory. However, although second stage coding is more concerned with theory development, it is still important to keep in touch with the data, and to constantly refer to initial interviews and codes. Smith stresses the importance of continuing to check back to primary source material (Smith 1996a, Smith et al. 1999). Division between first and second stage coding is somewhat arbitrary. It might be more useful to think of analysis of the data, as an ongoing process that continually switches between first and second stage coding in order to remain close to the data, and to challenge any a priori assumptions.

History of knee problem	Please can you tell me the history of your knee problem?
	Can you tell me what made you go to the doctor for the first time?
Cause/origin/progression	What do you think is wrong with your knee?
of knee problem	What do you think might have caused it?
1	What do you think would happen if you decided not have surgery
	for some reason?
Self-perceived need for	What made you decide that you needed a knee replacement?
TKR /Expectations from	When did you decide you needed one?
surgery	Did you think you need a TKR before you went to the doctor for
	the first time?
	How do you feel you will benefit from surgery?
Prioritisation for TKR	Do you have any idea how the doctor decides who should have surgery/have priority for surgery?
	If you were helping a doctor to decide who should have a knee replacement – What would you say to help this decision?
Important sources of	Can you tell me about anyone else who has the same problem?
information	Has anyone given you any other advice or information?
	What advice would you give to other people with the same
	problem as you?
Any other comments	Is there anything that we have not discussed that you think is
-	important?

Figure 2. Interview Schedule.

## **FINDINGS & DISCUSSION**

Results related to personal meanings in the construction of need for TKR from this study are reported elsewhere (Toye et al. 2006). This chapter interprets these results from a particular theoretical standpoint using Levanthal's Self regulatory Model. Results show that a person's perceptions of knee osteoarthritis, and the treatment available, affect a person's decision that they need a TKR. It supports findings that decisions to undergo TKR are not related to symptoms alone, but are related to certain beliefs.

Health Care Decision Making and Leventhal's Self-Regulatory Model

It is likely that faced with bodily symptoms we all behave differently (Mechanic 1978). Our own personal experience of symptoms may be very different to someone else's. This is supported by surveys showing that there is a significant group of people with severe knee pain who do not seek the doctor's help, or who would be unwilling to undergo TKR if offered (Tennant et al. 1995, Hawker et al. 2001).

Cognitive theories propose that behaviour is a result of our cognition – our thought processes (Rosenstock et al. 1988, Ogden 1996). It is likely that the *meaning* of symptoms for the individual affects their behaviour in the face of symptoms. Jones (1990) argues that people are inherently optimistic when trying to explain bodily symptoms, and that they are more likely to explain symptoms as natural events, than disease. People may only seek attention for symptoms that they cannot explain, or that seem unusual or different from normal. Even if we recognise a symptom as unusual, we may only seek help because the symptoms are interfering with our lives (Bury 1982, Williams 2000).

Leventhal's disease representation			Factor related to patients'	perception of need
Identity	What is its label and what are the symptoms of disease?	→	Medical Model of disease	X-ray findings and diagnosis by medical experts define need.
Consequences	What are the short and long-term effects of the disease?	<b>→</b>	Daily effects of arthritis	Pain, loss of function, "knee letting me down", "affecting my mood", fatigue.
Time-line	What is the course of the disease?	<b>→</b>	Disease progression	Osteoarthritis is a progressive disease.
Cure/control	How will I recover?	→	Expectations of TKR	TKR is the only cure.
Cause	What has caused my illness?	<b>→</b>	Not found to be related to need for TKR	

Figure 3. Relationship of Leventhal's Disease Representations to Factors Related to Patients Perceived Need For TKR in this Study.

Perceptions of need for TKR are therefore important, as they are likely to be related to a person's decision to undergo TKR. It is likely that a combination of the factors relate to need for TKR. Leventhal et al. (1980 and 1997) developed a model known as the 'Self Regulation Processing System'. The model is useful as it proposes that we have an active cognitive system for processing information, which results in 'representations' of illness (Baumann et al. 1989). These representations are formed as part of a person's own personal biography situated within a certain socio-cultural environment. Leventhal's model was useful in interpreting factors related to need for TKR. Following thematic analysis, Leventhal's model was modified and used to interpret findings related to patients' perceptions of need. The model did not provide a conceptual framework *prior* to analysis, but was used subsequently to frame perceptions of need for TKR. Leventhal et al. (1980) identified four disease representations through which a person gives meaning to their experience, which will affect a

person's action; Lau and Hartman (1983) included a 5th representation. These representations of disease are shown in Figure 3. This figure gives a representation of how Leventhal's model was used to frame patients' perceptions of need found in this study

Any combination of representations is possible, and is likely to result in different behaviour. Lau and Hartman (1983) found that the five disease representations were linked to each other, and that they affected illness behaviour. These representations have been consistently found to influence behaviour and outcome in those with chronic disorders such as arthritis (Sharloo and Kaptein 1997). However, most of the studies exploring representations of those with arthritis have focused on Rheumatoid Arthritis (Sharloo and Kaptein 1997), Barlow 1998), even though the population prevalence of this disease is considerably lower than that of osteoarthritis.

There are certain conceptual difficulties with using these 5 representations, and Leventhal's model does not fit perfectly with the findings in this study. It can be difficult to separate and clearly identify all five representations as they often overlap (Hampson 1997, Heijmans and de Ridder 1998), particularly in relation to chronic musculoskeletal diseases where pain is often an important factor. In addition, the 5 representations are not necessarily consistent, and may not be relevant in certain types of disease (Hampson 1997, Heijmans and de Ridder 1998). For example, Hampson (1997) found that for diabetes and arthritis, the 5 representations did not form independent constructs. Certain representations may overlap and be highly related to each other. The model is a useful way of looking and disease, but in using it, one must retain some flexibility.

Three other quantitative studies have investigated representations in those with osteoarthritis (Hampson et al. 1994, Hampson 1997, Orbell et al. 1998). These studies have included different components within representations, and it is therefore difficult to compare results. Hampson et al. (1994) developed a 'Personal Model of Arthritis Interview' (PMAI) to investigate representations of osteoarthritis patients over the age of 60 recruited from a rheumatology clinic (n = 61). The PMAI was originally devised to investigate the 5 representations of identity, consequences, cause, control and time-line. However, Hampson et al. (1994) found that several representations were highly correlated, and therefore combined representations. Identity, consequences and time-line were combined to form one representation, whilst both cause and control/cure were retained as separate representations. Orbell et al. (1998) used a qualitative study to elicit illness representations in a group of 72 consecutive patients for total joint replacement. In their study, the representations found differed from those in Hampson et al. (1994).

The following sections explore the results of this study in relation to Levanthal's disease representations.

### Identity - The Medical Model of Illness

Disease identity as defined in Levanthal's model usually refers to the disease *symptoms*, and the disease *label*. However, in the current study, the symptoms of osteoarthritis (e.g. pain) were not distinguished from the consequences on a person's life. For example, pain was often qualified by its effect on function. Symptoms were therefore included with the representation of consequences, rather than with identity. Hampson et al. (1994, 1997) also found that it is difficult to separate the representations of identity and consequences in patients with

osteoarthritis. An important aspect of identity (disease label) that emerged was the use of the medical model in labelling knee osteoarthritis. X-ray findings and medical diagnosis were important factors used by patients to identify a *disease label* and determine appropriate treatment (i.e. TKR).

## Importance of X-Ray Findings

Most patients adopted a medical model of osteoarthritis. That is - they felt that a specific medical diagnosis determined the need for a TKR. They described how an x-ray objectively 'showed' osteoarthritis, and this meant that they needed a TKR. If a person had osteoarthritis in their knee *confirmed* by x-ray, most felt that this meant that they needed a TKR, irrespective of symptoms. For example, some firmly believed that by looking at the x-ray the doctor can determine what is wrong with the knee, and how it should be treated. Following his consultation, one male patient had the right TKR done first, even though he did not have pain in this knee.

X-rays define [the surgeon] x-rayed them, and he said "they are both the same, which one do you want done first?" .....the main reason why I went for the right knee first was because of those pieces of bone. I couldn't *feel* anything in there. I didn't know which was the worse one......

This biomedical perspective was reinforced by contact with health professionals.

Importance of	The surgeon couldn't understand why I wasn't in paincos he said, normally
x-ray	people are in pain, I have never been in pain.
reinforced by	
doctor	

However, the majority also recognised that there were difficulties with basing decisions regarding need for TKR, purely on x-ray diagnosis. For example, they were surprised that their symptoms did not match the x-ray findings, or they said that the surgeon was surprised about how 'bad' the x-rays were.

My x-ray does<br/>not match my<br/>symptomsHe [surgeon] said.....which knee is it!? And I thought, It's the right knee that is<br/>hurting! He said, from the x-ray I can't tell the difference, one is as bad as the<br/>other......I don't know whether it is<br/>subconscious! (laugh) it probably started to hurt then!Q: Did that surprise you then that the x-rays were the same?<br/>Yes....it hadn't occurred to me....cos the right one was the one that was swelling up<br/>and making it difficult to get out of bed in the morning.

Other studies have also stressed the importance of x-ray results for the patient in defining disease. Lambert et al. (2000) found that doctors tended to focus on the medical markers of osteoarthritis, whereas patients focused on the symptoms of daily life. However, Lambert et al. (2000) also found that in spite of this, most patients stressed the importance of x-rays, even though doctors did not feel that tests had any effect on treatment. Busby et al. (1997) also

found that patients saw medical tests as very important: "*tests recurred like a litany in accounts*" (Busby et al. 1997: 93). Patients did not feel that the doctor would understand what was wrong with them until they had seen an x-ray. Woolhead et al. (2002) also found that most believed that the surgeon should decide on the need for TKR, based on medical diagnosis. This supports the finding that objective medical findings (x-rays) seemed to take preference over symptom burden in defining need for TKR. Bury (1982 and 1991) suggests that being able to objectify a disease legitimises the behaviour of those who are ill. It may be that as diagnosis by x-ray is necessary to legitimise the health care decisions of those with osteoarthritis, it would therefore be important for patients in the assessment of need for TKR.

### Importance of Diagnosis by an Expert

All those interviewed felt that through an objective diagnostic process, the doctor defined need for TKR. Contrary to expectations, only half of those interviewed remembered thinking that a TKR was a possible treatment prior to visiting the doctor (i.e. we *had* believed that the patient would have felt they needed a TKR before they saw the doctor). The other half said that they had definitely not expected a TKR to be suggested, but had expected the doctor to suggest pills or injections. Although this information was remembered retrospectively, and therefore some may have forgotten about their thoughts prior to seeing the doctor, some were adamant that they had *not* expected a TKR.

I did not	Q: Had you any thought in your mind [about TKR] before you saw the doctor?
expect a	Oh no! I didn't know what to expect reallyno, no, I thought perhaps a course of
TKR	physio and a, perhaps and injection or something.
	Oh, no! Never! I just thought I was gonna' get some painkillers or something!

This may of course reflect an inherent dislike for operations, but it may also reflect a cultural norm that emphasizes the doctor's diagnosis as pivotal to decisions of health need. People described great faith in the doctor as expert and felt that there was no point in seeing 'an expert' if you were not going to take his or her advice.

Duty to take<br/>the advice of<br/>expertsYou must take your doctors advice ..... if you are not going to, don't waste their<br/>time...... I am a firm believer that you don't go to your doctor and be given advice<br/>and then turn your back on it. I mean there is absolutely no point.... Do I go to a<br/>doctor if I have got something wrong with my leg, or do I go to a horse dealer?

Some specifically said that the surgeon would 'know' what was needed and would never do anything that was not needed.

Doctors won'tI mean if you didn't need it, they wouldn't give it. They won't do that, they have gotdo things thatplenty of work on; they don't need to have more work (laugh)....cos surgeons they arearen'tnot daft! They know if you need it or if you are just neurotic.necessary

There were only three cases where the patient remembered being asked to take part in the decision to have a TKR. However, in all three cases, although the doctor overtly handed the

decision over to the patient, the situation described did not seem to suggest that any real choice was available to the patient. The choice was a rhetorical one.

Rhetorical<br/>choice givenDr x [GP] said " I can't force you", he said, " but, in the end you will have to [have a<br/>TKR], else you will end up sitting in a wheelchair", and that sort of helped me make up<br/>my mind....... when I read between the lines I thought, he is more or less telling me<br/>that it is a thing that I have got to do, but I have got to make up me own mind.

However, although all agreed that the doctor was the expert and should define need, accounts were not consistent. For example, at the same time as seeing the doctor as an expert whose decision must be respected, more men than women talked about playing a personal part in the decision making, and stressed the importance of asking pertinent questions. Some felt that it was possible for the doctor to make mistakes and described situations where they had not agreed with the doctor.

Patient should<br/>play a part in<br/>theSome [people] are a bit more forceful I think .....some who are a bit more timid<br/>don't really question .....they just accept the first thing that comes along ......I<br/>think you should be a bit more....aggressive with the doctor.consultation

One man even questioned the financial motives of the surgeons.

The doctor is<br/>not always the<br/>patient'sThe doctor didn't ask me any questions. I know why they chop you up, because they<br/>are getting paid. Well, that's my opinion. If he does 2 of them a day and sits back on<br/>the rest of the month....like any business...if you make a lot of money you tend to<br/>do it....if I was in that business I would probably say, well have this knee, and make<br/>more money out of it.

## Importance of the Social Network

Although belief in the doctor as medical expert was an important factor in the patients' construction of need for TKR, the social network was also important. All the men and women interviewed described how social encounters influenced their decision that they needed a TKR. The majority described at least one positive account of TKR heard socially.

Positive	Mm, and I have got a friend I know in our village, he has had both his done, and he
accounts of	is pretty successful with hishe is back to nearly normal, he is walking normal
TKR from social	everybody you see is quite successful.
network	
	It is unbelievable, it's as if he never had a sore knee at all!

Although women also described positive accounts of TKR, they were more likely to describe negative accounts. However, in all cases where a negative case was described, a reason why this would not apply in their circumstances was always given. Failure of TKR was often put down to personal limitations. For example – "it is also in the head". Other reasons such as being overweight, being old, and not being in good health were also given.

Negative	She has never made any progress at all, I don't think it is just the knee, I think it is
examples of	also in the head!
TKR justified	That lady struggled afterwards, she was a little bit older than me and she still
	needs a walking stick, and I think that it is possibly her, you knowshe has
	always been, I think not really in the best of health.

Social comparison affected a person's decision that they needed a TKR. If a person felt that they were worse than other people that they had met with knee osteoarthritis, they were more likely to consider that they needed a TKR. If they felt that they were much better than others, then they were less likely to consider a TKR.

Social.....I think one of the things coming to [physiotherapy] I realised that I was muchcomparisonmore handicapped by pain than quite a few of them that were there. Because we<br/>used to go off and walk back to the cars, and I couldn't! and some of them....well<br/>my motorcycling friend...were striding right out.I came to [physiotherapy] ......I looked around and they had had their knees done<br/>and were walking with sticks, I was the fittest one there! You know I looked round<br/>there was a dozen people there, and they was in a lot worse condition than I<br/>was..... I was in front of them anyway as far as I could see.......way in front.

Patients also described how they had felt social pressure to have a TKR.

Social pressure<br/>to go to have aEverybody is saying go for it, and I daren't not now, I am on the track and I have<br/>got to go! (laugh).....I feel I have got to have it, and I have got to try.....<br/>everybody says "Mr So-and-so has had it – so they are marvelous!" and I thought<br/>well – go for it girl, you know, try.<br/>The boss's husband .... he was always drumming it in to me, you get and have it<br/>done, my girl he said, you will feel 10 times [better]....and another one was my<br/>daughter. She used to look after this little old lady, and she had her's done and she<br/>was 75 I think when she had it done. And she said, "Mum, she is great, I am telling<br/>you, you will be alright".

For those men whose wives sat in on the interview, it was clear that wives had an influence on their husband's decision. Some wives encouraged their husbands to have a TKR, even though the men were uncertain that they needed one.

Family	Wife:
influence	We have to get friends to take us out and they are very supportive. We couldn't
decision	have functioned this last year could we, not without them but you become
	embarrassed in the end, when you have got to keep asking and that all adds to
	the worry with his kneewhen you have been independent, all your life, you know,
	it is very very difficult.

Men also talked about the effect of the Media on their perception of need for TKR.

Positive effectI think they [newspapers] are all up for it I think.....I haven't seen any badof mediareports......they are all pretty positive about the joint replacements .....I think<br/>they are generally trying to promote it, aren't they. "It's out there if you want it!"<br/>.... I think that is why the waiting lists are going up, people are getting more<br/>information about it.

No women said that the Media had influenced their perception of need for TKR, and in fact several discussed how they actively *avoided* Media information once they had made the decision that they needed a TKR.

I avoid media I try not to watch [programmes on surgery]. I get a bit squeamish, you know, I think information oh.....I am not watching that, cos I think if I'd have watched it, I don't think I'd have had it done

Consequences - Effects of Osteoarthritis on Daily Life

Consequences of disease usually refer to the short and long-term effects of the disease. Although the patients were purposively sampled to represent those with a lower than average disease burden, patients described major effects of osteoarthritis on their lives, and this was directly related to perceptions of need for TKR. Effects included pain, loss of function, a feeling that the knee was 'letting you down' or affecting other parts of the body, and that the knee was negatively affecting mood of causing fatigue.

#### Pain

All mentioned pain as an important factor and related this to need for TKR. However, although pain was important, most found it very difficult to describe pain per se, rather, the description of pain was qualified by referencing it with the effects on a person's life, such as loss of function or low mood, or by pointing to or touching the body part. Some felt pain made life intolerable.

Pain makes life intolerable	I can understand people taking an overdose. I mean the value of life's gone, if you are knackered, and you are racked with pain 24 hours a day and you can't do anything, what is there to go on forI think there is a lot to be said for this euthanasia.
Pain qualified by gesture	I was sleeping at night with my leg hanging out of bedresting on a cushionI could not touch itI could not lay a sheet [placed sheet of paper on knee to demonstrate] on it at night it hurt so muchthe painand I don't mind paincos I have had a lot in my day.

Pain descriptors used by patients in this study suggest that the pain experience is complex. The complexity of any pain experience is well recognised (Melzack and Wall 1988). Different words may describe different qualities of pain. In this study, the intensity of pain ranged widely from being a 'nuisance' or 'inconvenient' to 'unbearable' or 'excruciating'. The McGill Pain Questionnaire (Melzack 1975) distinguishes sensory, affective and evaluative words for pain (Melzack and Wall 1988), as well as a group of

miscellaneous words. It also grades the intensity of different types of pain. Patients in this study used nine of the words included in the McGill Pain Questionnaire along with a further ten words (Table 1). This supports findings that disease burden may differ widely in patients listed for TKR.

		McGill classification
1.	Sharp	
2.	Shooting	Sensory
3.	Stabbing	
4.	Exhausting	
5.	Frightening	Affective
6.	Terrifying	
7.	Unbearable	Evaluative
8.	Annoying	Lvaluative
9.	Agonizing	Miscellaneous
10.	Discomfort	Intensity level 2
	Excruciating	Intensity level 5
	Ache	
13.	Awful	Other descriptors used in the current study
	Awkward	not found in Mcgill
	Continual	
	Frustrating	
	Maddening	
	Nuisance	
	Racked with pain	
	Ridiculous	
	Spasmodic	
22.	Terrible	

Tabla 1	Words in	the Megill De	in Augstianna	ire Used to De	sariha Pain ii	n this Study
Table 1.	worus m	the wieght i a	am Questionna	ine Useu io De	SCHUC I am n	i tins Study

## Loss of Function

Loss of functional activities was described as being related to need for TKR. There were noticeable gender differences in the activities described. Women were more likely to mention difficulty with walking or shopping. Men were more likely to the effect on social and leisure activities.

Walking and shopping	If I go shopping with my daughter and we are in a shopI can't stand for long I can walk to the shop at a distance, but then I have to stop or sit down,you can't carry on and walk around the shops for a few hours cos I wouldn't be able to do it.
Social and leisure activities	Quite a few of my mates have taken up golf, and I didn't start with them because I thought I couldn't walk 18 holes, and they have gone off and it is unlikely that I shall get in with them now because they can play 18 holes or so, and it too late for me to join them, they have got a bit of a system going.

Social and	I used to play a lot of golf, I can't do that at all now.
leisure activities	Q: So you would want to get back and play?
(Continued)	
	Oh yeah! Certainly! It's my main reason for doing it [the TKR] I think (laugh). To get back on the golf coursegot me down not playing golf again.

Baby-sitting, gardening and housework were examples of significant activities that had been affected, as well as difficulty with aspects of personal care. Getting up and down the stairs, or on and off the floor, were also mentioned as significant activities that had become difficult.

Baby-sitting	I baby-sit when [my daughter] goes to work. I can carry the little one from here say to about there [5 feet], or put him in his pushchair, but I would like to be able to carry him out to the carI would like to be able to nurse him more, I feel that if I carry him too far, then me knee might give way and, I am going to drop him.
Housework and DIY	I do sit down and do my ironing, butI get to the point when I think I can't be bothered to do that, whereas before I would be up, and I would do me gardening. I mean I would love to go out and do me garden, but, I have to wait for my grandson to do that, cos, I go out there for say about 5, 10 minutes, and then that is it I have to come in.
Difficulties with personal	When I get down to put my socks on, I have a struggle to get them on.
care	I am not incontinent or anything like that, but chances are I have to get up in the night, and of course when you have been asleep for 3 $\frac{1}{2}$ hours or something, it really is an effort to get up.

Men were concerned that their wives were doing more and more of their usual activities for them. Women were concerned about becoming reliant on their families or friends. This may be related to the fact that most women were widows and lived alone, whereas the men tended to be living with their wives.

Having to rely<br/>on othersI can't get in and out of the bath, I now have to sit on a seat, rather than get in the<br/>bath as such, I sit on the side of the bath and [my wife] had to help me out of there<br/>even!There are some jobs I can't do now like..... that [my wife] has to do that I would to<br/>take back to do myself! .....for instance if I hoover, I can't take the hoover, up and<br/>down the stairs, [my wife] has to do that.

Loss of transport or mobility was related to social isolation, or reliance on others.

TransportI don't do any driving now. Got rid of my car, the reason being is the fact that I have a<br/>job to get in the car.....I sort of sit on the side and just literally swing my leg in,<br/>and the struggle, and that.....just marooned in this house.

In most cases, employment status was not related to need for TKR, as only 4 people were still in paid employment. Only two people had given up work specifically because of their knee

Work I found I was having the devil's own job to get in the tractor. Once I was in it I was alright, but it is getting up into them.Sometimes, with my work, as I say, it becomes a bit dangerous, because of the aircraft work, the last thing you want to do is fall between propellers (laugh)

### Feelings of Knee 'Letting Me Down' or Affecting Other Parts of Body

Most discussed the fact that their knee was 'letting me down' in some way: giving way; feeling weak; having to take care or do things slowly; and not being able to rely on the knee to do what they wanted, whenever they wanted to do it. There was a sense of vulnerability. They could no longer take it for granted that they could do things.

Feeling I think if someone attacked me I couldn't chase after them, it all built in my mind I wasn't quite right, I was thinking I needed about twice the amount of time to cross the road.

My knee was starting to give up a little bit. At times at least it used to just give up, the knee, and I would lurch forward, and other times it would lock. Now I am extremely *gentle* with my knee when I go up on stairs, or steps of anything like that, in case, its not 100% *secure* if you wish, and I don't know when it is going to give in, so I have to just be aware.

People described a sense of urgency because their knee osteoarthritis was affecting other parts of the body. For example, believing that limping was giving them back pain, or 'wearing out' the other knee.

Affecting I had sciatica ever so bad, and it was caused by [the knee], I went to see [the surgeon] other parts and he said it was my knees that caused my bad back, got a disc out at the bottom.....and hopefully once I start walking properly when my knee is done, my back might be a lot better...

## Effect on Mood and Fatigue

Women were more likely to mention that osteoarthritis was making them 'grumpy' or that they lost their temper more often than usual. Others felt depressed or frustrated by their condition.

I am more grumpy I am going to be sat indoors like a cabbage, and sat about being miserable, and might get snappy and that towards the family, which is what I didn't want to do, I mean I am not that type, I am usually happy go lucky and like a good joke. I used to like to go out.....I mean we have quite a lot of parties, but I feel guilty that they say "come on!" but if I get up on the floor and have a dance, then I am going to suffer, and I think, "no I am not going to go through that"; but I sit there and think, "oh I wish I could get up there." Women were also more likely to mention fatigue as having a major effect on their lives.

 Fatigue
 My life was being so circumscribed.....and I was getting so tired all the time.....I hadn't realised how tiring pain was......I still get tired but not as tired.....not the waves of crippling tiredness.

Other studies have also found that pain could make people 'cranky', 'annoyed' or 'moody' which might lead to social withdrawal (Lambert et al. 2000). In many studies, pain and loss of function appear to be the most important consequences facing those with arthritis (Lambert et al. 2000, Tallon et al. 2000). Patient-derived outcome measures support this finding (Bellamy et al. 1988a, Dawson et al. 1998). However, the consequences seem to be more far reaching than pain and loss of function. Tallon et al. (2000) suggest that a person's concern that the knee will 'give out', as well as the psychological impact of knee osteoarthritis were both important consequences that are often neglected in knee osteoarthritis studies.

#### Modifying Effect of Age on Perceptions of Need

Advancing age seemed to modify perceptions of need for TKR, in that there were lower (or less commonly, higher) functional expectation. It was more usual for people to describe how getting older decreased their functional expectations, and therefore decreased their need for TKR. One man refers to this as learning to "grow old gracefully". Some felt it was not worth having a TKR once you got past a certain age. They described a window of opportunity in their lives where TKR was appropriate.

People	I don't think that I am very likely at my age, and I am 74 this coming week, I don't
shouldn't	think I am going to be walking much more than 3 or 4 miles at a time. Whereas at one
expect as	time I was doing 10 or 12 milesand as an aunt of mine once said to me with great
much when	feeling, "You must learn to grow old gracefully".
they are older	
	Perhaps if you are older maybe, you might not want it [TKR] so bad, I think maybe when you are older you just want it just to get rid of the pain, when you are a bit younger maybe just want to, well get rid of the pain obviously, and to get out and about more.

However, increase in age need not always decrease perceived need for TKR. One person described how an improvement in quality of life since his retirement has made it more important for him to stay active and therefore had increased his perceived need for TKR.

 My life has
 I used to work to 7 days a week. I never used to have a holiday ... I have got more time now to see more.....I go on the coach to the seaside .....had 3 trips last year.

 now I am
 older

Leventhal and Crouch (1997) also found that representations could change across a person's lifespan, meaning that need for treatment may therefore be underestimated in an older group of patients with knee osteoarthritis. Sanders et al. (2002) found that people with

osteoarthritis described the consequences of osteoarthritis as 'disruptive', at the same time as being a 'normal' part of ageing. Both Sanders et al. (2002) and Charles and Walters (1998) found that those with osteoarthritis tended to minimise its effects, in spite of the fact that it could have very serious consequences in their life. Disease categories can acquire social meaning and have an impact on a person's decisions (Good 1994). For example, age stereotypes may modify a person's behaviour. Modification of expectations as a result of age can make it difficult for a person to decide if they need a TKR. A person may ask themselves - am I a candidate for TKR, or is this normal for someone of my age? Due to the social association of advancing age with disability, this can mean that a lot of people do not consult a doctor when they have knee pain. Sanders et al. (2002) found that some patients with severe symptoms of hip and knee osteoarthritis were embarrassed about their condition, possibly due to its associations with ageing, and the negative social stereotypes related to this.

#### Time-Line - Beliefs Regarding Disease Progression

Time-line usually refers to the perceived course of the disease. In this study, beliefs about likely progression of osteoarthritis were important in patients' perception of need for TKR. Most of those interviewed believed that their knee osteoarthritis would progress, and had vivid pictures of what would happen to them if they did not have a TKR. Patients felt it was important to get something done as quickly as possible to halt the progression.

GetGet something done as quick as possible! Oh yes! not half. Because it can build.... I cansomethingimagine myself a year from now in a shocking state! ....whereby my quality of life willdone quickbe totally dictated by what I can and can't do.

I don't think I am being you know at all ... *pessimistic*, but I think that if I don't have it done within the next 3 months I doubt if I will be able to live by myself. You know just temporarily.

People had vivid pictures of what would be the natural course of their osteoarthritis. It was anticipation of the future that strongly shaped the decision to undergo surgery. Several said that they could live with their knee osteoarthritis if it did not get any worse than it was now. Being 'crippled' or being in a wheelchair was mentioned by half of the men and women as their worst fear. None were in this condition already. One man believed that the pain would eventually kill you.

I will be in a<br/>wheelchairIt must get worse and worse .....its frightening..... I will just become totally immobile,<br/>chair-bound, and whatever, that is all I can assume.....The surgeon, specialist over there<br/>said to me, "I am afraid you will have to get flat [live on a flat road], you can't live on a<br/>hill", like we are living here, "otherwise you will be a total cripple".<br/>I am only 65 I don't want to be in a wheelchair, a cripple for the rest of you life, I have<br/>got a few more years to go I hope! See things.

Patients felt that if their condition progressed in the future, they would not be able to live with it, and some even said that they would rather be dead than get any worse.

 I couldn't
 Well I was getting to the point that .....I don't want to go on like this......

 handle it if it
 this......and if it gets much worse, I shan't go on like this......

 got worse
 Well.....I am not going to be any bloomin good if I don't have it done....and I don't want to sit indoors......If I am going to sit indoors ...... I might as well be in my box.......I can understand people taking an overdose. I mean the value of life's gone.

The *anticipation* that they would not be able to do important activities in the future was an important factor in the decision to undergo TKR, for example, not being able to work, care for family, or be independent.

Anticipated<br/>disruption to<br/>important<br/>activitiesI could see a point coming where I couldn't...I wouldn't be able to drive... I had put<br/>[my husband] in the best nursing home I could find.....which was 20 miles<br/>away...there was no public transport connecting directly...I thought this is ridiculous!<br/>.....I decided that I must do something about it .....so I elected to go privately [for my<br/>TKR].

Beliefs about disease progression were reinforced by health care professionals, particularly during interactions described by the men. Most men had been told by a doctor that they could expect to need a TKR sometime in the future.

Doctor told[the surgeon] did the washout on the knee and he said he didn't think it would beme I willsuccessful, so I would have to come back in for the knee, to have a joint.need a TKRWhen I went in for the final check up, [the surgeon] said we can't do any more now,in the futurethe next step is a replacement joint. When you think you are ready for it, go and see<br/>your GP and he will get it lined up.

The representation of time-line is rarely discussed in the literature. In a review of illness perceptions in patients with chronic somatic illness (Sharloo and Kaptein 1997) was the most commonly explored representation (58%), and time-line was the least frequent (17%). Orbell et al. (1998) refer to time-line as how many weeks a person felt that it would take to make a full recovery from surgery. However, this does not relate to beliefs about disease progression. Hampson et al. (1994) do consider the importance of time-line, but combine this with consequences. However, the current study suggests that time-line is an important representation of knee osteoarthritis that relates to the decision to undergo TKR.

To support this finding, in a survey exploring the concerns of 197 patients with differing types of arthritis, Neville et al. (1999) found that the most frequently raised concerns, expressed by 88% of the people, was that their arthritis would worsen. Progression of the disease was a greater concern than pain or disability. As in this study, the patients in Neville et al. (1999) had a relatively low pain score (3.9 on a pain-VAS of 0-10). It may be that those with lower pain levels are more concerned with future expectations of the disease.

Despite the finding that patients saw osteoarthritis as a progressive disabling disease, there is limited evidence regarding the progression of osteoarthritis, which seems to vary substantially between individuals (Felson et al. 1995, Hart et al. 1999, Wolfe and Lane 2001). There are difficulties with estimating progression rates. Firstly, many studies investigate the radiological *incidence* (number of new cases occurring with positive x-ray findings) of osteoarthritis, rather than *progression* over a certain time period. Secondly, many studies base progression on x-ray findings, and few studies have looked at progression of knee

osteoarthritis from a patient's perspective. It seems likely that a significant number of people with knee osteoarthritis improve symptomatically (Cooper et al. 2000, Dieppe et al. 2000, McAlindon 1999). In an eight-year prospective study (n=349), although 63% of those with knee pain described themselves as 'worse' over 8 years, 17% described themselves as 'the same', and 20% were 'better' (Dieppe et al. 2000). In a similar patient group (Ledingham et al. 1995), although radiographic improvement was seen in only 5%, when patients were asked whether their symptoms had 'got better', 'stayed the same', or 'got worse', 23% felt they had 'got better' during the follow-up (mean 2 years).

## Cure/Controllability - Expectations of Treatment Outcome

Cure or controllability of disease usually relates to how a person expects to recover or to control their disease. Differences in expectation of cure/controllability may affect a person's decision to undergo TKR. The majority of patients in the current study had positive expections of TKR, and believed it was the only cure.

Patients in the current study had a range of expectations from TKR. This is supported by Mancuso et al. (2001) who developed a questionnaire to measure the expectations of TKR in a group of 377 patients undergoing TKR. Beliefs related to other possible treatment options and expectations of a possible cure were also related to the need for TKR. Although people talked about the positive results of using at least one complementary therapy (acupuncture, cod-liver oil, glucosamine, green oils and heat), they did not see them as a cure. Although most had tried other treatments (physiotherapy, medication, steroid injections, arthroscopy), TKR was seen as the only *cure* for knee. There were only two men who felt that they had 'cured' their arthritis themselves, the first by taking tablets, and the second by doing exercise. Both of these men had decided against TKR.

I can cure Sulphur rebuilds that skin [cartilage] between the two [bones] .... I can feel it, I can sense it happening over these weeks, it was getting less and less, you know ....I don't think it can turn the other way, I have got utter faith in that tablet, I don't think it will turn back, in fact it helps me all round.

Osteoarthritis of the knee was seen as a localised and curable problem, and was explicitly compared positively to other diseases, such as Rheumatoid Arthritis and Alzheimer's disease, that could not be cured.

The knee can beWith Rheumatoid Arthritis you can't chop it out and stick a new bit in. You are stuck<br/>with it .......[and] Alzheimer's, that's a one way ticket innit...... we are bloomin'<br/>lucky, you go in, they cut it out, stick another bit in, and off you go again until the<br/>next time, but not with them, you can't cut the brain out and put another one in!

However, despite the belief that the TKR was the only cure, it was not described as an easy option. Prior to the interviews it was expected that those people purposively sampled for interview might see TKR as a quick and easy fix that would return them to, or maintain normal function. This expectation was not supported by comments made by participants. The

decision to undergo TKR was not taken lightly, or seen as an 'easy fix' with no risks. Most said that they did not expect a 100% improvement from TKR.

**TKR offers no** They say, when you have had this one done you will be able to go off disco-ing and things like that, I mean I won't be but, I shall be able to get about more, more freedom sort of thing.

I am not....going to ask for the moon, cos I know I shan't get it, but just to be able to get up from a chair and walk, without grabbing grab rails. Just to be able to walk, I am not going to ask to run, and I am not going to ask to dance, although I love doing both. If I can just walk around the supermarket, or even just a small distance I would be happy.

There is no way I am going to get back to 100% whatever happens... I could have a TKR and come back and say look what have you done! Its true! There is no guarantee. Are you going to tell me I am going to walk perfect when I have got this TKR. You won't be able to say yes will you, because there is no guarantee with it!

Patients also discussed fear of surgery itself.

**Fear of surgery** I used to keep putting it off ..... I told them when I was in there that I was petrified of having it done. I always had that fear that when they were going to do it I would come round, while they were doing it.

You have got to watch that thrombosis don't set in, and that did frighten me, but years ago I did have thrombosis of the knee. They said about if a bit breaks off it could travel up [points to heart] and I thought oooh no!

I said, "oh what are you in for then" and she said, "oh I had a knee replacement" she said, "have got an infection in it", and her knee it was right out here!

I said [to surgeon] "I hear you have a 90% success rate", he said, "oh we have about that, yes, but in fact you are one of the 10%". So I said "it doesn't mean that I am going to die though does it?" He said, "oh, most certainly". He frightened me, so I wouldn't have had it done anyway, I would have suffered the pain.

Beliefs about cure/controllability of knee osteoarthritis in other studies seem to be mixed. Tallon et al. (2000) also found that TKR was seen as the 'only cure' for osteoarthritis, whereas Lambert et al. (2000) found that patients saw TKR as both frightening and impractical. In other studies, many people perceive arthritis to be a disease with no cure, but at the same time hope for effective treatment (Hampson et al. 1994, Busby et al. 1997), where others think that very little can be done (Rao et al. 1998, Lambert et al. 2000).

Patient expectations of outcome following TKR may not be realistic. Mahomed et al. (2002) found that 40% of people undergoing hip and knee replacement expected to return to their normal activities. However, a study by Heck et al. (1998) has shown that the average patient does not return to normal functioning after TKR, and continues to have a functional status, as measured by SF36, slightly below one standard deviation from the normal population in the USA. In a small study of 29 consecutive patients undergoing TKR, Walsh et al. (1998) found that 1 year following TKR, walking speeds were around 17% slower than age matched controls. Men with TKR were also 37-39% weaker and women were 28-29% weaker than age matched controls.

However, Moran et al. (2003) suggest that patients' expectations from TKR *are* realistic, and in fact are lower than the expectations of surgeons. Moran et al. asked both a group of

patients with hip and knee osteoarthritis (n = 217), and a group of orthopaedic surgeons (n = 70) to indicate on the Oxford Knee/Hip Score, the level of symptoms that they would expect 6 months following surgery. They found that the surgeons expected a significantly higher level of function following joint replacement than the patients. Of particular interest is that, patients who had *not* previously had a joint replacement had similar expectations to those who *had*, suggesting that patient expectations were realistic. This supports the finding that most patients did not expect 100% cure from surgery, and seems to be incongruent with the finding that TKR was seen as the only 'cure'.

Despite the high expectations from TKR found in this study, the amount of benefit from TKR is unclear. Two meta-analyses investigating outcomes of knee replacement (Callahan et al. 1994, Callahan et al. 1995) showed that knee replacement is an effective treatment for osteoarthritis of the knee. However, there are no trials that compare TKR with any other intervention, and only three randomised controlled trials that compare different surgical methods, rather than comparing TKR to other interventions (Dieppe et al. 1999). Also, in a recent systematic review of treatment for osteoarthritis carried out by a task force of the European League against Rheumatism (Pendleton et al. 2000), it was found that TKR was supported only by descriptive studies (Doherty and Dougados 2001). One of the difficulties with interpreting data from TKR outcome studies, is that a variety of outcome measures are used. In a critical review of studies reporting patient-related outcome from TKR (Dieppe et al. 1999), only seven studies reported patient-related outcomes, and only four of these used patient-completed outcome measures. This is important, as there may only be a low correlation between surgeon-completed outcome measure and patient satisfaction following TKR (Bullens et al. 2001). It is therefore difficult to compare or evaluate the results of various studies.

## Cause - Beliefs About Cause of Osteoarthritis

As beliefs about cause of illness have been related to illness behaviour in other studies (Hampson et al. 1994, Orbell et al. 1998), it was included as an item in the interview schedule, and explored during the interviews. Beliefs regarding cause of osteoarthritis are listed in Table 2. However in the current study, cause was not related to patients' perceived need for TKR. This may be because this study focuses on patients with localised knee arthritis, rather than generalised arthritis. This sample also differs from other studies in that it looks specifically at patients who have lower than average pain and higher function. Only one person directly related cause of osteoarthritis to need for TKR, and possible cure. However, he still wanted to stay on the waiting list 'just in case'.

Lack of<br/>sulphurNow the way I look at it, if animals naturally eat grass, tree leaves and bark out of a<br/>tree, or whatever they eat, it has all got this sulphur in it and I think basically it is<br/>sulphur that is processed out of all food. I know if I stop taking the [sulphur] for<br/>probably 3 months it [the arthritis] would probably grow back again.

In other studies, causal attributions have been shown to affect patients' decisions regarding treatment (Hampson et al. 1994, Orbell et al. 1998). As in this study, osteoarthritis

was *not* usually seen as something that is caused by an external agent. Work, the hardships of life, and repetitive actions are common attributions of cause in arthritis (Busby et al. 1997). In interviews with patients with knee osteoarthritis, Campbell et al. (2001) found that those who attributed osteoarthritis to factors such as "wear 'n' tear" (Campbell et al. 2001: 135), had a more resigned attitude to their condition, and were less likely to believe that exercise intervention would work. In-depth interviews suggest that those with osteoarthritis see it as a normal and unavoidable part of aging (Charles and Walters 1998, Lambert et al. 2000, Hudak et al. 2002, Sanders et al. 2002).

	Number of patients
Wear 'n' tear/activity	12
Age	4
Shock	2
Genetic	6
Damp/rain	2
Accident	1
Complications of previous surgery	3
Chance/no reason	5
Sport	10
Army	1
Work	2
Lack of sulphur in diet	1
Body weight	3

Table 2. Causes of Osteoarthritis Reported

Moral Narratives – Patient Accounts of Priority for TKR

Patients were also asked what factors gave a person priority for TKR. In accord with Woolhead et al. (2002), participants reported that pain, functional loss, employment status, social circumstances, fitness, and age were priority factors. Woolhead et al (2002) also found that if a person needed bilateral replacement, or if they had paid national insurance contributions they were perceived as having higher priority. Perhaps the most surprising finding was that almost all felt that certain personal factors should affect their priority. This finding challenges the assumptions of a biomedical explanatory model. For example, a person's attitude to the doctor should be considered. A person was expected to take the doctor's advice and to play a part in getting well. Others felt that those who did not 'give up' should have priority. Only one person specifically said that a person's characteristics should not affect the decision.

Attitude should affect	I think the first thing you look for is the attitude of the person.
priority	Q: What would be a bad attitude?
	Well big demands, and rudeness. "I want this." He might not need the
	surgery. Because a lot of people read medical books which is a bit
	dangerous sometimes.
	Q: So what might be a good sort of attitude?
	Just listen to what the surgeon has to say, he might ask you questions, you
	just give him the honest answer and let him decide I didn't go and insist
	to the doctor, I need to do this, he knows his job, I don't want to make their
	job any worse by going out and making a fuss about things.

One person illustrates 'the right way to be ill'

'the right way to be	The motorcyclistwas determined to beat it. I thought his attitude was
ill'	greathe was looking forwardhe wasn't dispirited about the whole
	thingMy sister who has had a knee replacementshe took a long time
	to get over it she wasn't cooperative her attitude was - "the
	doctor has got to do this for me and then it will go away"she
	has enjoyed ill health.

Most felt that if a person still had the desire to lead full and active lives in spite of their age, then they should have priority. Their comments suggested that being active and strong, having aims and ambitions despite advancing age were commendable attributes that should give priority. One patient discusses how important it was to make the doctor aware of what type of person that you are. Having been active is described as something that is virtuous and that gives priority for TKR.

Right age attitude	Well (sigh), their mental attitude, their age attitude, well you get some people who are 60 going on 80 don't you(laugh) and some people 80 going on 45 I think someone who would give the impression that he would want to be mobile, that you are active, you have perhaps got an active mind, you know, your attitude someone who had got a bit more go. I was trying to work out how best to convey to him [the surgeon], that it was incredibly important to me that I had it done, I really didn't know what tack to take (laugh) I thought the thing to do was to show my worth as an individual that was being hamperedI was a lecturerand I wrote books and thingsthat my condition was making it very difficult to carry on as a useful member of society.
	as a useful member of society.

In summary, certain personal characteristics were described as giving a person priority for TKR. That is, patients described undesirable traits as: letting the disease 'beat you', having a 'bad attitude' towards the doctor, 'letting yourself' get worse, not having a 'positive age attitude'. The dimension supports the idea that health care need is socially constructed (Turner 1992). It is perhaps also notable that all the other priority criteria mentioned were related to a description of good moral character. This is illustrated in Table 3. Discussion about pain giving priority for TKR was related to judgments about whether someone *really* had pain. Although pain was seen as an important determinant of priority, it was regarded as subjective and difficult to measure, presupposing that there was a certain right way to behave

when you have pain. Several observed that it was difficult to know if someone really had pain. Some suggested that certain people were less able to cope with pain. The subjective nature of pain may explain why patients with osteoarthritis place such an importance on x-ray findings. The legitimising purpose of objective tests has been suggested by Bury (1982, 1991).

PAIN	Does a person really have pain?	How do they know who has got more pain than another? It is difficult. I mean you could say oh, oh, its agony. and maybe it is not, I mean some people are quite dramatic aren't
		they, let's face it [laugh].
WORK	Does a person do good work?	The health visitor from up here she was in real agony. We used to say to her, "Won't they do it any quicker because you are working for the health". We did feel sorry for her because she provides a good service.
SOCIAL CIRCUMSTANCES	Does a person look after others?	I think people with a young family [should have priority], cos I suppose that they find it hard to take them to school and all that sort of thing, I should think they should get a bit of help, quicker.
BEING FIT	Described as morally superior	If two people came to the doctor and one has got a bad heart and isn't going to last very long, and the other one is a fit man, I think the fit man should go first. Perhaps I shouldn't say this, but, er (laugh) I am going in next Tuesdaywell, I have got to <i>prove</i> that I am fit when I am in there.

Table 3. Moral Character Related to Priority Criteria for TKR

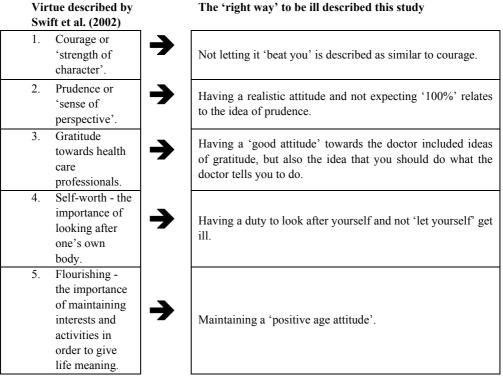
The concept of 'virtue' has been explored in semi-structured interviews of older women with osteoarthritis (Swift et al. 2002). Swift et al. describe 5 virtues or 'strengths' that may help a person to cope with disease. In the current study, similar virtues are described as the 'right way' to be ill, and are seen as traits that should affect a person's priority for TKR. Figure 4 shows a comparison between the virtues found in Swift et al. (2002), and the 'right way' to be ill described in by patients in the current study. Some might argue that if possessing these virtues gives a person priority for treatment, it is a mechanism for maintaining social structure that disadvantages certain groups. For example older patients who do not possess traits associated with a 'positive age attitude' might not be considered a priority for TKR. In other words, if a person is 'virtuous' they will benefit in terms of health care; those who are not virtuous will not benefit. This is an important issue, as it seems that a person needs to appear legitimately ill if they are to seek, and gain access to, treatment.

Friedson (1970) suggests that legitimate illness can either be *conditional* (a patient has a duty to get well), or *unconditional* (a person is not expected to get well). In theory, knee osteoarthritis is considered *unconditionally* legitimate, because although osteoarthritis was attributed to a person's own activity (wear 'n' tear), people did not feel responsible for having it. However, it could be argued that unconditional legitimacy is rare, and that legitimacy

depends on behaving in a certain way (Friedson 1970). In this study legitimacy of disease was related to two factors:

- 1. A medical model of disease, in particular the importance of x-ray findings and expert diagnosis,
- 2. Demonstrating the 'right way' of being ill.

Due to the finding that health care decision-making is related to a person's virtue, a person with osteoarthritis must negotiate being ill with being a good person (Annandale 1998). Bury (2000) describes this as a "moral narrative", where a person tells someone how they are "successfully ill" (Bury 2000:272), and maintains their "moral character" (Riessman 1990: 1195). Within a moral narrative, people present a certain picture of the self in order to pursue "virtue" in the face of illness (Williams 1993:96).



## The 'right way' to be ill described this study

Figure 4. Comparison of 'virtues', and descriptions of the 'right way' to be ill in this study.

## Gender – Perceptions of Need for TKR

Several authors have found gender differences in utilisation of TKR. It may be that gender differences, explain the finding that TKR is more underused in women (Juni et al. 2003) and that once listed, women tend to have a higher disease burden than men (Kennedy et al. 2003).

There were several factors in the current study that may help to explain why women are listed later in the course of their disease:

## 1. Approach to the TKR Waiting-List

Men talked about the waiting-list being a 'safety-net', and although some felt they did not need a TKR at that time, they all wanted to stay on the list. Most of the men said that if they were offered a TKR now, that they would decline, whereas no women said this.

Waiting list is a safety net	[the surgeon] said, "you're OK we will knock you off the list", I said "no I want to stay on the list, I have waited 18 months, if I suddenly go downhill I don't want to wait another 18 months so I am quite happy if you see me 12
	months down the line, but don't take my name off the list (laugh)!"while I am on the list I have got a safety-net.

Men tended to say that the only fair way to order the list was on a first come first served basis. No women mentioned this way of organising the waiting-list.

First come first served	Either the problem is there or it is not! And it is jolly well thereArthritis is
	not going to get betterif I have taken my position on the waiting-list, and I
	have waited that long, without being awkward about it, I think , when my time
	comes, per the rota, I am entitled to this operation!

## 2. Biomedical Perspective

Men tended to describe need for TKR in terms of a biomedical model, as 'all or nothing', and based on x-ray and expert diagnosis. Women described need in more relative terms, and tended to consider the relative need for TKR with respect to others. Most women believed priority should depend on the symptom burden and said that they would let those with higher burden have priority.

<b>Relative need for TKR</b>	If I thought the next person next door to me was worse off than me then I
	would say, well you do theirsand I will step back. If I thought that they
	needed it more than I did.

Interaction with healthcare professionals may have reinforced this gender difference. Most men had been told by a doctor that they would definitely need a TKR sometime in the future, whereas only one woman had been told this.

Importance of expert	When I went in for the final check up [15 years ago], [the surgeon] said we
diagnosis	can't do any more now, the next step is a replacement joint. When you think
	you are ready for it, go and see your GP and he will get it lined up.

## 3. Patient As Expert

Men appeared more pro-active than women in discussing and questioning treatment options with the doctor. For example, relatively more men than women stressed the importance of asking pertinent questions. Men felt that it was possible for the doctor to make mistakes and described situations where they had not agreed with the doctor. Only one woman discussed the importance of asking the doctor questions and being involved in health care decisions. Hawker et al. (2000) suggest that one of the possible mechanisms for differential utilisation of TKR is whether or not a patient had discussed TKR with their doctor. They found that this had the greatest impact on a person's willingness to consider TKR. It is difficult to know whether this is a result of patient or clinician factors, or a combination of both.

Patient is expert	My uncle Frank always said that if the patient has an idea of what is wrong with
	them then it is a tremendous help! (laugh)'cos they have got it all the time
	and you only see them for 10 minutes!

Men were also more willing to seek out information via the media, and described how reading articles about TKR in the newspaper, magazines or watching television were influential. No women said that they had been influenced by the media. In fact, most women discussed how they avoided media information once they had made the decision that they needed a TKR.

Accessing media	I think they [newspapers] are all up for it I thinkI haven't seen any bad
Accessing meura	
information	reportsthey are all pretty positive about the joint replacementsI
	think they are generally trying to promote it, aren't they. "It's out there if you
	want it!" I think that is why the waiting-lists are going up, people are
	getting more information about it.
	I try not to watch [programmes on surgery]. I get a bit squeamish, you know, I
	think ohI am not watching that, cos I think if I'd have watched it, I don't
	think I'd have had it done

It is also interesting that more women described a negative account of TKR that they had heard socially. It may be that women access different types of information when making decisions about health care needs. It is therefore important for clinicians, to create an environment where women can talk openly about the possibility of TKR, if that is the appropriate treatment option.

## 4. Expectations of TKR

There were also gender differences in expectations of TKR. All of the women with pain specifically mentioned pain relief as an expectation of TKR, whereas only 2 men mentioned pain relief as a specific expectation. It may be that pain relief was not such an important outcome for most men, or it may also reflect different gender stereotypes in relation to the expression of pain. In terms of functional outcome, Half of the men suggested that they would expect to be '*back to normal*' after a TKR. None of the women expected this. Men were also more likely to expect a return to sport of leisure activities. Women talked more of a return to a basic level of walking.

I want to be back to normal	Q: what do you expect from your TKR?
	Wellto be back to normal! to put it briefly [laugh] yeah. I just want to be normal I will try and live into my 100s if can.

Karlson et al. (1997) also found that men were more likely to choose surgery earlier in the course of their disease than women. Men talked more about giving up activities requiring high function, such as sport and travel, whereas women talked about mobility issues and simple functional activities such as combing their hair. Men spoke less about distrusting doctors, and seemed to have more faith in doctors. Women were more concerned about finding the more experienced doctor, or the 'best hospital'. Men did not talk about the correct timing of surgery, whereas women did. Women gave reasons for waiting such as: waiting for a certain 'threshold'; waiting for technology to improve; and also that surgery was irrevocable. Men did not discuss reasons for waiting.

Mancuso et al. (1996) found that more women expected improvements in walking (57% compared to 29%), and more men expected improvements in sport (51% compared to 40%). It may be that women have a more realistic approach to TKR, and therefore wait until their symptom burden is more severe before discussing it with their doctor. It may be that women have a more realistic approach to TKR, and therefore wait until their symptom burden is more severe before discussing it with their doctor.

#### 5. Effect on Mood

Finally, women were also more likely to describe reduced mood as an effect of osteoarthritis. Half of the women mentioned that osteoarthritis was making them 'grumpy' or that they lost their temper more often than usual.

Effect on mood	I am going to be sat indoors like a cabbage, and sat about being miserable, and
	might get snappy and that towards the family, which is what I didn't want to do, I
	mean I am not that type, I am usually happy go lucky and like a good joke.

Unruh (1996) suggests that women may be vulnerable to emotional attributions for pain, and therefore may be less likely to be listed for surgery or be prescribed painkillers. These aspects to osteoarthritis may be missed, or not taken seriously during a consultation, and this may result in an underestimation of the symptom burden in women. On the other hand, the psychological burden involved when experiencing chronic disease may be underestimated in men, which may mean that men are more likely to be referred for invasive procedures when other approaches might be effective.

It is also possible that the social construction of the 'appropriate' way to behave when faced with disease is different in men and women, meaning that not only do men and women behave differently, but also may be treated differently by others (including health care professionals). Using focus groups, Karlson et al. (1997) found that most men thought that: women would delay having surgery; women would have more responsibilities; men would be the risk takers. Women thought that: men would have surgery sooner because men dealt poorly with pain; men were afraid of pain; men were impatient with decreased function; men would place their own interests first.

#### Study Limitations

This study confirms the idea that personal meanings are important in the construction of need for TKR. However, there are some considerations that may influence interpretation of the results.

Firstly, the sample was purposive and only included patients with a lower than average disease burden as measured by WOMAC pain and function. Those with a lower disease burden on the list for TKR surgery may have a different perception of need, than those with a high disease burden, and findings from this study cannot be extrapolated to all on the waiting list for TKR. However, there were several reasons (given above) why this particular group was chosen:

Secondly, more men than women agreed to participate in qualitative interviews. The reasons for this are not known. It may be explained by the finding that more men then women had decided that they did not want a TKR at the time of the interview, and therefore that men wanted to discuss the reasons behind their decision. It may also be explained by the finding that women were more likely to avoid information related to TKR once they had made their decision to have surgery. Having made the decision that they needed a TKR, they may not have wanted to discuss this. However, despite this limitation, the sample of women was large enough to allow several important themes to emerge.

Thirdly, a criticism that can be made of qualitative research, are that its findings are anecdotal. That is, the researcher uses data to support a priori assumptions. In this study, recognised research methods were used to avoid anecdotal use of data and no factor was excluded from the analysis because it did not fit the developing theory.

Finally, the findings present only one interpretation of patients' perceptions of need for TKR. The philosophical underpinnings of the research method – Interpretative Phenomenological Analysis – recognise that interpretation is likely to be context- specific. Some argue that meaning is constructed within the interaction, and that therefore a different context would provide different findings, although it could be argued that this criticism could be made of all research. However, although the results cannot be extrapolated to another setting, several themes arose that may give a deeper understanding of the beliefs and perceptions of patients waiting for TKR surgery. A strength of this study, is the apparent candidness in the responses of all participants.

## CONCLUSION

This chapter has discussed the findings from 18 qualitative interviews exploring patients' perceptions of need and priority for TKR. Four of the five representations suggested by Leventhal et al. (1980), and Lau and Hartman (1983) appear useful in interpreting the findings related to patients' perceptions of need for TKR. These representations are important because they are likely to affect a person's decision to be listed for TKR, and to undergo surgery. Certain representations of knee osteoarthritis are likely to augment patients' perceptions of need for TKR, irrespective of the burden of symptoms, and this may partly explain the variability of symptoms in those listed for surgery. For example, a person may be more likely to decide to go ahead with surgery if they believe that: Osteoarthritis is identified

by the x-ray (identity); it is causing pain, loss of function, fatigue and low mood (consequences); it is progressive (time-line); and TKR is the only cure (cure/control).

Differences in disease representations are important as they may explain some of the variability in decision-making regarding need for TKR. Hampson et al. (1996) suggest that interventions targeting osteoarthritis should attempt to modify patients' appraisal of the disease. They found that negative appraisal resulted in passive coping strategies and low mood. Some of the representations found in the current study might not hold true, and may augment certain patients' perceptions of need for TKR. Representations may be modifiable through interventions such as Cognitive Behavioural approaches (Keefe et al. 1999, Keefe and Caldwell 1997), or through self-management (Barlow et al. 1999, Barlow et al. 2002). In view of the findings of this study, it may be useful to target the following representations during a consultation:

- 1. *Identity* This study showed that patients, particularly men, tend to adopt a medical model of disease. The reliance of the patient on the doctor's diagnosis is supported by the findings of Woolhead et al. (2002). One way to address this problem may be to tackle the model of disease used. For example, if one takes the view that need for TKR is defined by objective x-ray findings, then it is likely that the decision will be left to the doctor. If however, a patient knows that their own symptom burden is a determinant of need for TKR, then they can take their place as co-expert in decision-making.
- Time-line A more realistic view of the likelihood of disease progression would be useful. It may be that that a significant number of people with knee osteoarthritis improve symptomatically and that the course of osteoarthritis is heterogeneous with a significant number remaining symptomatically stable (Cooper et al. 2000, Dieppe et al. 2000).
- 3. *Control/Cure* It might be useful to explore other treatment options, rather than encouraging a view that TKR is the only cure for the symptoms of knee osteoarthritis. Research suggests that after TKR, patients are still below normal levels of function and do not return to 'normal' (Heck et al. 1998, Walsh et al. 1998). Kennedy et al. (2003) suggest that 10% of patients have a poor outcome following TKR, and Heck et al., 1998 suggest that 9% are dissatisfied with the outcome. However, it is important to consider the impact of any negative messages regarding outcome from TKR. Mahomed et al. (2002) have shown that an expectation of complete pain relief was an independent predictor of better function following TKR. This issue is of particular importance in the context of recent attempts to fully inform patients regarding the risks of invasive technologies (Bowling and Ebrahim 2001). However, as there is evidence to show that interventions such as: exercise (Ettinger et el. 1997, Van Baar et al. 1998b, Van Baar et al. 1999), Cognitive Behavioural approaches (Keefe et al. 1990a, Keefe et al. 1990b, Keefe et al. 1999, Keefe and Caldwell 1997), self management (Barlow et al. 1998, Barlow et al. 1999), can modify the effect of knee osteoarthritis. It might therefore be more useful to direct patients to these non-invasive treatments, prior to considering TKR. No randomised controlled trial has compared the effect of these interventions with TKR (Dieppe et al. 1999).

It would be useful to elicit patients' representations of disease during medical consultations, and to target any representations that might adversely affect a person's assessment of need for TKR. Simple questions during the consultation might be useful, such as: What do you think would happen if you did not have a TKR; Do you think that any other treatment would be effective; What do you expect from your TKR?

The results discussed in this chapter show evidence for the use of moral narratives in health care decision making. It is likely that social criteria are used in medical decision-making, and therefore that social constructs are likely to affect all treatment allocation decisions. If 'good' people are felt to have priority for health care, this may make it difficult to incorporate patients' or professionals' views on priority, into health policy. For example, although the majority may feel that a person with a 'good attitude' should have priority – does this mean that they *should* have priority in practice? Although it is important to incorporate the patient as expert, what if perceptions of need and priority are related to factors such as age, gender or race? As these criteria may implicitly shape the decision-making of professionals, this may be a strong argument for the use of explicit criteria to determine need for TKR. This raises an important issue related to the development of criteria of need, and is an exceptionally difficult area. How do we determine which criteria should be used, and what are the legal, moral and ethical implications of incorporating these criteria?

Finally, there seem to be important gender differences in representations that may explain why women tend to be listed later in the course of their disease. Gender issues are important as they may result in inequalities in health provision. However, consideration needs to be given as to whether TKR is under-utilised in women, or over-utilised in men.

## REFERENCES

Annandale, E. (1998) The Sociology of Health and Illness Polity press.

- Barlow, JH. Turner, AP. Wright, CC. (1998) Long-term outcomes of an arthritis selfmanagement programme, *British Journal of Rheumatology*, vol 37, pp1315-1319.
- Barlow, JH. Williams, B. Wright, C. (1999) 'Instilling the strength to fight the pain and get on with life': learning to become an arthritis self-manager through an adult education programme, *Health Education Research*, vol 14, pp533-544.
- Barlow, JH. Wright, C. Sheasby, J. Turner, A. Hainsworth, J. (2002) Self Management approaches for people with chronic conditions: a review, *Patient Education and Counseling*, vol 4, pp177-187.
- Baumann, LJ. Cameron, LD. Zimmerman, RS. Leventhal, H. (1989) Illness representations and matching labels with symptoms, *Health Psychology*, vol 8(4), pp449-49.
- Bellamy, N. (2000) WOMAC osteoarthritis index. User Guide IV, available: (www.womac.com).
- Bellamy, N. Buchanan, WW. Goldsmith, CH. Campbell, J. Stitt, L. (1988b) Validation Study of WOMAC; A health status instrument for measuring clinically important patient relevant outcomes following total hip or knee arthroplasty in osteoarthritis, *Journal of Orthopaedic Rheumatology*, vol 1, pp95-108.
- Blumer, H. (1969) Methodological position of Symbolic interactionism. In Blumer, H. (ed) *Symbolic Interactionism*, University of California Press.

Bowling, A. (1997), 1<sup>st</sup> Edition, Research Methods in Health, Oxford University Press.

- Bowling, A. Ebrahim, S. (2001) Measuring patients' preferences for treatment and perception of risk, *Quality in Health Care*, vol 10(supplement 1), i2-i8.
- British Medical Association (1998) Waiting List Prioritisation Scoring Systems A discussion paper, *Health Policy and Economic Research Unit, discussion Paper* no. 6 December 1998.
- Bruner, J. (1986) Actual Minds, Possible Worlds. Harvard University Press.
- Bullens, PH. van Loon, CJ. de Waal Malefijt, MC. Laan, RF. Veth RP. (2001) Patient Satisfaction after total knee arthroplasty, *The Journal of Arthoplasty*, vol 16(6), pp740-747.
- Bury, MR (1982) Chronic illness as biographical disruption, *Sociology of Health and Illness*, vol 4, pp167-72.
- Bury, MR. (1991) The Sociology of chronic illness: a review of research and prospects, *Sociology of Health and Illness*, vol 13, pp451-468.
- Bury, MR. (2000) Illness narratives: Fact or fiction? *Sociology of Health and Illness*, vol 23, pp263-285.
- Busby, H. Williams, G. Rogers, A. (1997) Bodies of knowledge: lay and biomedical understandings of musculoskeletal disorders. In, Elston, MA. (ed.), *The Sociology of Medical Science and Technology*, Blackwell.
- Callaghan, D. (1991) Ethical Considerations. In Hadorn, DC. (ed.), *Basic Benefits and Clinical Guidelines*, Westview Press.
- Callahan, CM. Drake, BG. Heck, DA. Dittus, RS. (1994) Patient outcomes following tricompartmental knee arthroplasty. A meta-analysis, *Journal of the American Medical Association*, vol 271(17), pp1349-1357.
- Callahan, CM. Drake, BG. Heck, DA. Dittus, RS. (1995) Patient outcomes following unicompartmental or bicompartmental knee arthroplasty. A meta-analysis, *Journal of Arthroplasty*, vol 10(2), pp141-150.
- Campbell, R. Evans, M. Tucker, M. Quilty, B. Dieppe, P. Donovan, JL. (2001) Why don't patients do their exercises? Understanding non-compliance with physiotherapy in patient with osteoarthritis of the knee, *Journal of Epidemiological Community Health*, vol 55, pp132-138.
- Campbell, R. Quilty, B. Dieppe, P (2003) Discrepancies between patients' assessments of outcome: qualitative study nested within a randomised controlled trial, *British Medical Journal*, vol 326, pp252-253.
- Charles, N. Walters, N. (1998) Age and gender in women's accounts of their health: interviews with women in South Wales, *Sociology of Health and Illness*, vol 20, pp331-50.
- Charmaz, K. (2000) Grounded Theory. In, Denzin, NK. Lincolm, YS. (eds), 2<sup>nd</sup> Edition, *Handbook of Qualitative Research*, Sage publications.
- Cooper, C. Snow, A. McAlindon, TE. Kellingray, S. Coggon, A. Dieppe, PA. (2000) Risk Factors for the incidence and progression of radiographic knee osteoarthritis, *Arthritis* and Rheumatism, vol 43(5), pp995-1000.
- Coyte, P. Wang, PP. Hawker, G. Wright, JG. (1997) The relationship between variations in knee replacement utilization rates and the reported prevalence of arthritis in Ontario, Canada, *The Journal of Rheumatology*, vol 24(12), pp2403-2412.

- Dawson, J. Fitzpatirck, R. Murray, D. Carr, A. (1998). Questionnaire on the perceptions of patients about TKR, *Journal of Bone and Joint Surgery*, vol80B(1), pp63-69.
- Denzin, NK. Lincoln, YS. (2000), The discipline and practice of qualitative Research. In, 2<sup>nd</sup> Edition, Denzin, NK. Lincoln, YS, *Handbook of Qualitative Research*, Sage publications.
- Dexter, P. Brandt, K. (1993) Relationships between social background and medical care in osteoarthritis, *Journal of Rheumatology*, vol 20, pp698-703.
- Dieppe, P. Basler, HD. Chard, J. Croft, P. Dixon, J. Hurley, M. Lohmander, S. Raspe, H. (1999) Knee replacement surgery for osteoarthritis; effectiveness, practice variations, indications and possible determinants of utilization, *Rheumatology*, vol 38, pp73-83.
- Dieppe, PA. Cushnagan, J. Tucker, M. Browning, S. Shepstone, L. (2000) The Bristol 'OA500' Study: progression and impact of the disease after 8 years, *Osteoarthritis and Cartilage*, vol 8, 63-68.
- Doherty, M. Dougados, M (2001) Evidence-based management of osteoarthritis: practical issues relating to the data, *Best Practice and Research Clinical Rheumatology*, vol 15 (4), pp517-525.
- Doyal, L. Gough, I. (1991) A theory of human need, Macmillan.
- Ettinger, WH. Burns, R. Messiers, SP. Applegate, W. Rejeski, WJ. Morgan, T. Shumaker, S. Berry, MJ. O'Toole, M. Craven, T. (1997) A Randomised trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis, *Journal of the American Medical Association*, vol 277(1), pp 25-31.
- Felson, DT. Zhang, Y. Hannan, MT. Naimark, A. Weissman, BN. Aliabadi, P. Levy, D. (1995) The incidence and natural history of knee osteoarthritis in the elderly. *The Framingham osteoarthritis study, Arthritis and Rheumatism*, vol 38(10), pp1500-1505.
- Fitzpatrick, R. Davey, C. Buxton, MH. Jones, DR. (1998) Criteria for selecting a patientbased outcome measure, *Health Technology Assessment*, vol 2, p14.
- Fortin, PR. Clarke, AE. Joseph, L. Liang, MH. Tanzer, M. Ferland, D. Phillips, C. Partridge, AJ. Delisle, P. Fossel, AH. Mahomed, N. Sledge, CB. Katz, JN. (1999) Outcomes of total hip and knee replacement, *Arthritis and Rheumatism*, vol 42(8), pp1722-1728.
- Friedson, E (1970) *Profession of medicine. A study of the sociology of applied knowledge.* Harper and Row.
- Garro, LC. Mattingly, C. (2000). Narrative as construct and construction. In, Mattingly, C. Garro, LC. (eds) *Narrative and the cultural construction of illness and healing*, University of California press.
- Good, BJ. (1994) Medicine, Rationality and Experience, Cambridge University Press.
- Gubrium, JF. Holstein, JA. (2000) Analysing Interpretive Practice. In, 2<sup>nd</sup> Edition, Denzin, NK. Lincoln, YS. (eds) *Handbook of Qualitative Research*, Sage Publications.
- Hadorn, DC. (1991a) Necessary Care Guidelines. In, Hadorn, DC. (ed.) Basic Benefits and Clinical Guidelines, Westview Press.
- Hadorn, DC. (1991b) The Model Proposal. In, Hadorn, DC. (ed.) Basic Benefits and Clinical Guidelines, Westview Press.
- Hadorn, DC. (1991c) Response to Kaplan. In, Hadorn, DC. (ed.) Basic Benefits and Clinical Guidelines, Westview Press.
- Hadorn, DC. Brook, RH. (1991) Defining our terms. In, Hadorn, DC. (ed.) Basic Benefits and Clinical Guidelines, Westview Press.
- Hampson, SE. (1997) Personal Models and the management of chronic illness: A comparison of diabetes and osteoarthritis, *European Journal of Personality*, vol 11, pp401-414.

- Hampson, SE. Glasgow, RE. Zeiss, AM. (1994) Personal models of osteoarthritis and their relation to self management activities and quality of life, *Journal of Behavioural medicine*, vol 17(2), pp143-158.
- Harry, LE. Nolan, JF. Elender, F. Lewis, JC. (2000) Who gets priority? Waiting list assessment using a scoring system, *Annals of the Royal College of Surgery England*, vol 82 (Supplement), pp186-188.
- Hart, DJ. Doyle, DV. Spector, TD. (1999) Incidence and risk factor for radiographic knee osteoarthritis in middle aged women, *Arthritis and Rheumatism*, vol 42(1), pp17-24.
- Hawker, GS. Wright, JG. Coyte, PC. Williams, I. Harvey, B. Glazier, R. Badley, EM. (2000) Differences between men and women in the rate of use of hip and knee arthroplasty, *The New England Medical Journal*, vol 342(14), pp1016-1022.
- Hawker, GS. Wright, G. Coyte, PC. Williams, JI. Harvey, B. Glazier, R. Wilkins, A. Badley, EM. (2001) Determining the need for hip and knee arthroplasty; the role of clinical severity and patient preferences, *Medical Care*, vol 39(3), pp206-16.
- Hawker, GA. Wright JG. Glazier, RH. Coyte, PC. Harvey, B, Williams, JI. Badley, EM (2002) The effect of education and income on need and willingness to undergo total joint arthroplasty, *Arthritis and Rheumatism*, vol 46(12), pp3331-3339.
- Heck, DA. Robinson, RL. Partridge, CM. Lubitz, RM. Freund, DA. (1998) Patient outcomes after knee replacement, *Clinical Orthopaedics*, vol 356, pp93-110.
- Heijmans, M. de Ridder, D. (1998). Assessing illness representations of chronic illness: explorations of their disease-specific nature, *Journal of Behavioral Medicine*, vol 21(5), pp485-503.
- Helman, CG. (2000), 4th Edition, Culture Health and Illness, Butterworth Heinemann.
- Hospital episode statistics (2007) London: Department of Health [online] available: www.hesonline.nhs.uk
- Hudak, PL. Clark, JP. Hawker, GA. Coyte, PC. Mahomed, NN. Kreder, HJ. Wright, JG. (2002) "You're perfect for the procedure! why don't you want it?" Elderly arthritis patients' unwillingness to consider total joint arthroplasty surgery; a qualitative study, *Medical Decision Making*, vol 22, 272-78.
- Husserl, E. (1970) Logical Investigation. Humanities press.
- Ibrahim, SA. Siminoff, LA. Burant, CJ. Kwoh, CK. (2002) Understanding ethnic differences in the utilisation of joint replacement for osteoarthritis. The role of patient level factors, *Medical Care*, vol 40(1), Supplement I44 – I55.
- Jones, R. (1990) Expectations and delay in seeking medical care: *Journal of Social issues*, vol 46(2), pp81-95.
- Juni, P. Dieppe, P. Donovan, J. Peters, T, Eachus, J. Pearson, N. Greenwood, R. Frankel, S (2003) Population requirements for primary knee replacement surgery: a cross sectional study, *Rheumatology*, vol 42, pp516-521.
- Kaplan, A. (1991) Technical Considerations. In, Hadorn, DC. (ed.) Basic Benefits and Clinical Guidelines. Westview Press.
- Karlson, EW. Daltroy, LH. Liang, MH. Eaton, HE. Katz, JN. (1997) Gender Differences in patient preferences may underlie differential utilisation of elective surgery, *The American Journal of Medicine*, vol 102, pp524-530.
- Keefe, FJ. Caldwell, DS. (1997) Cognitive Behavioral control of arthritis pain, Medical Clinics of North America, vol 81(1), pp277-291.

- Keefe, FJ. Caldwell, DS.Williams, DA. Gil, KM. Mitchell, D. Robertson, C. Martinez, S. Nunley, J, Beckham, JC. Crissen, JE. Helms, M. (1990a) Pain coping skills training in the management of osteoarthritic knee pain. A comparative study, *Behavior Therapy*, vol 21, pp49-62.
- Keefe, FJ. Caldwell, DS. Williams, DA. Gil, KM. Mitchell, D. Robertson, C. Martinez, S. Nunley, J. Beckham, JC. Helms, M. (1990b) Pain coping skills training in the management of osteoarthritic knee pain-II: Follow up results. *Behavior Therapy*, vol 21, pp435-447.
- Keefe, FJ. Caldwell, DS. Baucom, D. Salley, A. Robinson, E. Timmons, K. Beaupre, P. Weisberg, J. Helms, M. (1999) Spouse assisted coping skills training in the management of knee pain in osteoarthritis: Long term follow up results, *Arthritis Care and Research* vol 12(2), pp101-111.
- Kelly, KD. Voaklander, D. Kramer, G. Johnston, DW. Redfern, L. Suarez-Almazor, ME. (2000) The Impact of Health Status on Waiting Time for major joint arthroplasty, *The Journal of Arthroplasty*, vol 15(7), pp877-883.
- Kennedy, LG. Newman, JH. Ackroyd, CE. Dieppe, PA (2003) When should we do knee replacements, *The Knee*, vol 10, pp161-166.
- Kleinman, A. (1988) *The illness narratives. Suffering healing and the human condition.* Basic books.
- Lachance, L. Sowers, MF. Jamadar, D. Hochberg, M. (2002) The natural history of emergent osteoarthritis of the knee in women, *Osteoarthritis and Cartilage*, vol 10, pp849-854.
- Lambert, BL. Butin, DN. Moran, D. Zhao, SZ. Carr, BC. Chen, C. Kizis, A (2000) Arthritis Care: Comparison of Physicians' and Patients' Views, *Seminars in Arthritis and Rheumatism*, vol 30(2), pp100-110.
- Lau, RR. Hartman, KA. (1983) Common Sense Representations of Common illnesses, *Health Psychology*, vol 2(2), pp167-185.
- Ledingham, J. Regan, M. Jones, A. Doherty, M. (1995) Factors affecting radiographic progression of knee osteoarthritis, *Annals of the Rheumatic Diseases*, vol 54, pp53-58.
- Leventhal, EA. Crouch, M. (1997). Are there differences in perceptions of illness across life span. In, Petrie, KJ. Weinman, JA (ed), *Perceptions of Health and Illness*, Harwood academic publishers.
- Leventhal, H. Meyer, D. Nerenz, D. (1980) The common sense representation of illness danger. In Rachman, S. (ed) *Medical Psychology*, volume 2, Elmsford, NY: Pergamon.
- Leventhal, H. Benyamini, Y. Brownlee, S. (1997) Illness Representations. In, Petrie, KJ. Weinman, JA (ed), *Perceptions of Health and Illness*, Harwood academic publishers.
- Lewis, S. Barer, ML. Sanmartin, C. Sheps, S. Shortt, SED. McDonald, PW. (2000) Ending waiting list mismanagement: principles and practice, *Canadian Medical Association Journal*, vol 162, pp1297-1300.
- Mahomed, NN. Liang, MH. Cook, EF. Daltroy, LH. Fortin, PR. Fossel, AH, Katz. JN. (2002) The importance of patient expectations in predicting functional outcomes after total joint arthroplasty, *Journal of Rheumatology*, vol 29(6):1273-9
- Mancuso, CA. Ranawat, CS. Esdail, JM. Johanson, NA. Charlson, ME. (1996) Indications for total hip and total knee arthroplasties, *Journal of Arthroplasty* vol 11(1), pp34-46.
- Mancuso, CA. Sculco, TP. Wickiewicz, TL. Jones, EC. Robbins, L. Warren, RF. Williams-Russo, P. (2001) Patients Expectation of knee surgery, *Journal of Bone and Joint Surgery*, vol 83(7), pp1005-1012.

- McAlindon, TE. (1999) The knee. In, Croft, P. Brooks, PM. Regional Musculoskeletal Pain, *Clinical Rheumatology*, vol 13, p2.
- McAlindon, TE, Cooper, C. Kirwan, JR. Dieppe, PA. (1992) Knee pain and disability in the community, *British Journal of Rheumatology*, vol 31, pp189-92.
- McAlindon, TE. Cooper, C. Kirwan, JR. Dieppe, PA. (1993) Determinants of disability in osteoarthritis of the knee, *Annals of Rheumatic Diseases*, vol 52, pp258-262.
- McConnell, S. Kolopack, P. Davis, AM. (2001). The Western Ontario and McMaster Universities Osteoarthritis index (WOMAC): A Review of its utility and measurement properties, *Arthritis Care and Research*, vol 45, pp453-461.
- Mechanic, D. (1978) Illness Behaviour. In, Mechanic D, *Medical Sociology*. 2<sup>nd</sup> edition. Free press.
- Melzack, R. (1975) The McGill Pain Questionnaire: major properties and scoring methods, *Pain*, vol 1, pp277-99.
- Melzack, R.Wall, PD. (1988), 2<sup>nd</sup> edition, The Challenge of Pain. Penguin Books.
- Moran, M. Khan, D. Sochart, DH. Andrew, G (2003) Expect the best, prepare for the worst: surgeon and patient expectation of the outcome of primary total hip and knee replacement, *Annals of the Royal College of Surgeons of England*, vol 85, pp204-206.
- National Health Service Plan (2000), available: www.nhs.uk/nationalplan (September 2003).
- Neville, C. Fortin, PR. Fitzcharles, M-A. Baron, M. Abrahamowitz, M. Du Berger, R. Esdaile, JM. (1999). The needs of patients with arthritis: the patient's perspective, *Arthritis Care and Research*, vol 12(2), pp86-95.
- O'Reilly, SC. Muir, KR. Doherty, M. (1996) Screening for pain in knee osteoarthritis: which question? *Annals of the Rheumatic Diseases*, vol 55, pp931-3.
- Ogden, J. (1996) Health Psychology, Open University Press.
- Orbell,S. Johnston, M. Rowley, D. Espley, A. Davey, P. (1998) Cognitive representations of illness and functional and affective adjustment following surgery for osteoarthritis, *Social Science and Medicine*, vol 47(1), pp93-102.
- Peat,G. McCarney, R. Croft, P. (2001) Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care, *Annals of the Rheumatic Diseases*, vol 60, pp91- 97.
- Pendleton, A. Arden, N. Dougados, M. Doherty, M. (2000) EULAR recommendations for the management of knee osteoarthritis: report of a task force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT), Annals of Rheumatic Disease, vol 59, pp936-944.
- Pescosolido, BA. (1992) Beyond rational choices; the social dynamics of how peoples seek help, *American Journal of Sociology*, vol 97(4), pp1096-1138.
- Peterson, MGE. Hollenberg, JP. Szatrowki, TP. Johanson, HA. Mancuso, CA. Charlson, ME. (1992) Geographic variation in the rates of elective total hip and knee arthroplasties among Medicare beneficiaries in the United States, *The Journal of Bone and Joint Surgery*, vol 74A(10), pp1530-1539.
- Rao, JK. Arick, R. Mihaliak, K. Weinberger, M. (1998) Using focus groups to understand arthritis patients' perceptions about unconventional therapy, *Arthritis Care and Research* vol 11(4), pp253-60.
- Riessman, CK. (1990) Strategic uses of narrative in the presentation of self and illness, *Social Science and Medicine*, vol 30(11), pp1195-2000.
- Riessman, CK. (1993) Narrative Analysis, Sage publications.

- Rogers, A. Hassell, K. Nicolaas, G. (1999) *Demanding patients? Analysing the use of primary care*, Open University Press.
- Rosenstock, IM. Strecher, VJ. Becker, MH. (1988) Social Learning Theory and the Health Beliefs model, *Health Education Quarterly*, vol 15(2), pp175-183.
- Roy, CW. Hunter, J. (1996) What happens to patients awaiting arthritis surgery? *Disability and Rehabilitaion* vol 18(2), pp101-105.
- Sanders, C. Donovan, J. Dieppe, P. (2002) The significance and consequences of having painful and disabled joints in older age : co-existing accounts of normal and disrupted biographies, *Sociology of Health and Illness*, vol 24(2), pp227-253.
- Schwandt, TA. (2000) Three epistemological stances for qualitative inquiry. In, Denzin, NK. Lincolm, YS. 2<sup>nd</sup> edition, *Handbook of Qualitative Research*, Sage publications.
- Seale, C. (1999) The quality of qualitative research, Sage publications.
- Sharloo, M. Kaptein, A. (1997) Measurement of illness perceptions in patients with chronic somatic illness; A review. In Petrie, JK. Weinman, JA.(ed.) *Perceptions of Health and Illness*, Harwood Academic Publishers.
- Sheaff, R. (1996) The need for health care, Routledge.
- Silverman, D. (2000) Doing Qualitative Research. Sage publications.
- Silverman, D. (2001) Credible Qualitative Research. In, Silverman, D. 2<sup>nd</sup> edition, *Interpreting Qualitative Data. Methods for analysing talk, text and interaction.* Sage publications.
- Smith, JA. (1995) Semi-structured interviewing and qualitative analysis. In Smith, JA. Harre, R. Van Langenhove, L. (eds.) *Rethinking Methods in Psychology*, Sage publications.
- Smith, JA. (1996) Beyond the divide between cognition and discourse: using interpretative phenomenological analysis in health psychology, *Psychology and Health* vol 11, pp261-271.
- Smith, JA. Jarman, M. Osborn, M. (1999) Doing Interpretative Phenomenological Analysis. In, Murray, M. Chamberlain, K. *Qualitative Health Psychology*. Sage publications.
- Stevens, A. Raftery, J. (1994) Health Care Needs assessment, Radcliffe Medical Press Ltd.
- Strauss, A. Corbin, J. (1998) Grounded Theory Methodology. In, Denzin, NK. Lincoln, YS. *Strategies of Qualitative Inquiry*, Sage publications.
- Swift, TL. Ashcroft, RE. Tadd, W. Campbell, AV. Dieppe, PA. (2002) Living well through chronic illness: the relevance of Virtue Theory to Patients with Chronic Arthritis, *Arthritis Care and Research*, vol 47(5), pp474-478.
- Tallon, D. Chard, J. Dieppe, P. (2000) Exploring the priorities of patients with osteoarthritis of the knee, *Arthritis Care and Research*, vol 13(5), pp312-319.
- Tennant, A. Fear, J. Pickering, A. Hillman, M. Cutts, A. Chamberlain, MA. (1995) Prevalence of knee problems in the population aged 55 years and over: identifying the need for TKR, *British Medical Journal*, vol 310, pp1291-1293.
- Thomson, R. Bowling, A. Moss, F (2001) Engaging patients in decisions: a challenge to health care delivery and public health, *Quality in Health Care*, vol 10 (Supp 1) p i1.
- Toye, F (2004) Assessing the need for total knee replacement. Phd thesis. Coventry University
- Toye, FM. Barlow, J. Wright, C. Lamb, SE (2006) Personal Meanings in the construction of need for total knee replacement surgery. *Social Science and Medicine*, vol 63 (1): 43-53
- Turner, BS. (1992) Regulating Bodies, Routledge.
- Unruh AM. Gender variations in clinical pain experience. Pain. 1996, vol 65, pp 123-167

- Urwin, M. Symmons, D. Allison, T. Brammah, T. Busby, H. Roxby, M. (1998) Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation, *Annals of the Rheumatic Diseases*, vol 57, pp649-55.
- Van Baar, ME. Dekker, J. Oostendorp, RA. Bijl, D. Voom, TB. Lemmens, JA. Bijlsma, JW. (1998) The effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee; a randomised clinical trial, *The Journal of Rheumatology*, vol 24(12), pp2432-2439.
- Van Baar, ME. Assendelft, WJ. Dekker, J. Oostendorp, RA. Bijlsma, JW (1999) Effectiveness of exercise therapy in patients with osteoarthritis of hip or knee. A systematic review of randomised clinical trials, *Arthritis and Rheumatism*, vol 42(7) pp1361-1369.
- Van Walraven, C. Paterson, JM. Kapral, M. Chan, B. Bell, M. Hawker, G. Gollish, J. Schatzker, J. Ivan Williams, J. Naylor, D. (1996) Appropriateness of primary total hip and knee replacements in regions of Ontario with high and low utilisation rates, *Canadian Medical Association Journal*, vol 155(6), pp697-706.
- Walsh, M. Woodhouse, LJ. Thomas, SG. Finch, E. (1998) Physical Impairments and functional limitations : A comparison of Individuals 1 year after total knee arthroplasty with control subjects, *Physical Therapy*, vol 78(3), pp248-258.
- Ware, JE. (1983) Measuring patient views: the optimum outcome measure, *British Medical Journal*, vol 306, pp1429-1430.
- Whitehead, M. Townsend, P. Davidson, N. Davidsen N. (1990) *Inequalities in health: Black Report*, Penguin books.
- Williams, G. (1993) Chronic illness and the pursuit of virtue in everyday life. In, Radley, A. (ed) *Worlds of Illness*. Sage publications.
- Williams, SJ. (1999) Is anyone there? Critical realism, chronic illness and the disability debate. *Sociology of Health and Illness*, vol 21(6), pp797-817.
- Williams, SJ. (2000) Chronic illness as biographical disruption or biographic disrupt as chronic illness? Reflection on a core concept, *Sociology of Health and Illness*, (22) 1, 40-67.
- Williams, MH. Frankel, SJ. Nanchahal, K. Coast, J. Donovan, JL. (1994) TKR. In, Stevens, A. Raftery, J. Health Care Needs Assessment, Radcliffe Medical Press Ltd.
- Wilson, MH. May, DS. Kelly, JJ. (1994) Racial differences in the use of total knee arthroplasty for osteoarthritis among older Americans, *Ethnicity and Disease*, vol 4, pp57-67.
- Wolfe, F. Lane, N. (2001) The longterm outcome of osteoarthritis: rates and predictors of joint space narrowing in symptomatic patients with knee osteoarthritis, *The Journal of Rheumatology*, vol 29(1), pp139-146.
- Woolhead, GM. Donovan, JL. Chard, JA. Dieppe, PA. (2002) Who should have priority for a knee joint replacement? *Rheumatology*, vol 41, pp390-394.
- Zola, IK. (1973) Pathways to the doctor- from person to patient, *Social Science and Medicine*, vol 7, 677-689.

Chapter 7

# A CONSTRUCTIVIST FRAMEWORK FOR EXPLORING INFLUENCES ON DECISION-MAKING ABOUT TREATMENTS FOR CHRONIC PAIN<sup>\*</sup>

## Cary A. Brown<sup>1,†</sup>

<sup>1</sup>University of Alberta Edmonton, Alberta Canada T6G 2G4

## ABSTRACT

This chapter will first present the theoretical underpinnings of a constructivist analysis of chronic pain and then discuss more specifically the relevance of the derivative Consciousness Model [1]. The current literature will be presented to illustrate the applicability of a constructivist analysis of people's chronic pain experience and the decisions they make about treatments for their pain. Particular attention will be placed on exploring in greater depth the four components of the Consciousness Model (coherence, sense of self, purposiveness and affect). The second part of the chapter will present the outcomes of a research study that used Delphi methods (iterative rounds of feedback and mixed methods questionnaires) [2] to explore the utility of the Consciousness Model in understanding how service users and providers construct meaning for chronic pain and how, in turn, that meaning influences decision-making about treatment. The findings presented in this chapter are extracted from a larger multi-stage study carried out in the United Kingdom with the multidisciplinary membership of the British Pain Society and three chronic pain sufferers' support groups in the North-West of England. The aim of this study was to examine how service users and providers made decisions about which treatment components they believed were important for chronic pain. Stage one of the large study identified, via postal survey, what treatment components service providers and service users identified as 'important'. In Stage two a constructivist framework was applied to develop an iterative Delphi questionnaire exploring service users' and

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Medicine and Health Care*, edited by Thomas E. Lynch published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: 2-64 Corbett Hall, University of Alberta Edmonton, Alberta Canada T6G 2G4, phone: (780) 492-9545, email: cary.brown@ualberta.ca

providers' rationale for their choices. Stage three also employed a Delphi process and asked participants to reflect on their previous comments and the researcher's analysis of these data. Stages two and three used postal questionnaires and summary reports from preceding rounds for background information.

The study findings revealed that there were statistically significant differences between what service users and service providers perceived as influencing their decision-making. While service user findings showed that the influences on decision-making were fairly evenly distributed across all four of Chapman's Consciousness Model domains this was not the case for service providers. Most service providers clearly felt that *affect* and *self-image* had little influence on their decision-making. Rather they more strongly endorsed the categories of *coherence* and *purposiveness*. An explanatory model has been derived from these findings and is presented in the latter section of the chapter. The model details how differential access to control and knowledge can play a strong mediating role on all aspects of consciousness and consequently on decision-making.

Lastly the chapter will discuss that although a constructivist paradigm advances our understanding of influences on decision-making in chronic pain it remains an insufficient explanatory tool. The chapter concludes by identifying how the emerging transtheoretical [3] and complex adaptive systems [4] paradigms hold promise for advancing our knowledge about decision-making yet further in relation to what is defined, by progressive thinkers grappling with the post-modern issues of health in the 21<sup>st</sup> century, as the 'fifth vital sign': pain [5].

Key words: Chronic pain, constructivism, decision-making, complexity science

#### INTRODUCTION

The volume of research generated around the issue of chronic pain demonstrates not only pain's significance as a healthcare concern but also its complexity and continued resistance to understanding. A number of theoretical models have been proposed to explain the dynamics of chronic pain. However, a clear evidence-base remains elusive. The call for theoretical frameworks that integrate the biological, psychosocial and environmental aspects of chronic pain is a consistent theme in current research. One such framework, informed by the disciplines of anthropology and sociology, is that of constructivism. The goal of a constructivist analysis is to gain an understanding of the multifaceted nature of a life experience from the perspective of the person living it. Basic assumptions of this model are that people create knowledge based on their own worldview. Reality is not an objective and universally shared entity but rather a construction based on the beliefs a person holds about an event and the meaning assigned to those beliefs.

...constructivism means that human beings do not find or discover knowledge so much as construct or make it. We invent concepts, models, and schemes to make sense of experience and, further, we continually test and modify these constructions in the light of new experiences' [6:125].

This chapter will first present the theoretical underpinnings of a constructivist analysis of chronic pain and then discuss more specifically the relevance of the derivative Consciousness Model [1]. The current literature will be presented to illustrate the applicability of a constructivist analysis of people's chronic pain experience and the decisions they make about

treatments for their pain. Particular attention will be placed on exploring in greater depth the four components of the Consciousness Model (coherence, sense of self, purposiveness and affect). The second part of the chapter will present the outcomes of a research study that used Delphi methods (iterative rounds of feedback and mixed methods questionnaires) [2] to explore the utility of the Consciousness Model in understanding how service users and providers construct meaning for chronic pain and how, in turn, that meaning influences decision-making about treatment. The findings presented in this chapter are extracted from a larger multi-stage study carried out in the United Kingdom with the multidisciplinary membership of the British Pain Society and three chronic pain sufferers' support groups in the North-West of England. The aim of this study was to examine how service users and providers made decisions about which treatment components they believed were important for chronic pain. Stage one of the large study identified, via postal survey, what treatment components service users (SU) and service providers' (SP) identified as 'important'. In Stage two a constructivist framework was applied to develop an iterative Delphi questionnaire exploring SUs' and SPs' rationale for their choices. Stage three also employed a Delphi process and asked participants to reflect on their previous comments and the researcher's analysis of these data. Stages two and three used postal questionnaires and summary reports from preceding rounds for background information.

The study findings revealed that there were statistically significant differences between what service users and service providers perceived as influencing their decision-making. While service user findings showed that the influences on decision-making were fairly evenly distributed across all four of Chapman's Consciousness Model domains this was not the case for service providers. Most service providers clearly felt that *affect* and *self-image* had little influence on their decision-making. Rather they more strongly endorsed the categories of *coherence* and *purposiveness*. An explanatory model has been derived from these findings and is presented in the latter section of the chapter. The model details how differential access to control and knowledge can play a strong mediating role on all aspects of consciousness and consequently on decision-making.

Lastly the chapter will discuss that although a constructivist paradigm advances our understanding of influences on decision-making in chronic pain it remains an insufficient explanatory tool. The chapter concludes by identifying how the emerging transiciplinary [3] and complex adaptive systems [4] paradigms hold promise for advancing our knowledge about decision-making yet further in relation to what is defined, by progressive thinkers grappling with the post-modern issues of health in the 21<sup>st</sup> century, as the 'fifth vital sign': pain [5].

#### Consciousness and the Construction of Pain

Many theorists construe chronic pain not as a sensation, dependent on intensity and duration of neural stimuli, but rather as a perception [1, 6-10]. This perception requires the brain's processing of sensory information, in conjunction with association and memory, to create a subjective reality. This reality is dynamic and modified by forces both internal and external to the individual. Chapman and colleagues [1], propose that consciousness is a key component in this perception formation and that without consciousness (for example. during coma or an anaesthetised state) pain does not exist. In their work, consciousness is

characterised by four key features: coherence, sense of self, purposiveness and personal nature and affect.

The first feature, *coherence*, is seen as the individual's inherent tendency to organise experiences into meaningful patterns with a temporal (sequence of time) component. *Sense of self* is what this coherence and organisation is structured around. This holds for the temporal aspect as well, people organise their meanings around temporal markers in their own lives – present, past, future. An example of this would be a person telling you that, "when I was a child I enjoyed mowing the lawn because it meant I was trusted by my parents. Now, as an adult, I find it a chore, more of an obligation and not pleasurable. In the future, when I am retired, I will look forward to using the lawn mower because I will enjoy the exercise and feeling that I am still capable of looking after my own home".

The third feature is *purposiveness*. This is seen as social, biological and psychological forces that drive an individual's attention to select certain events requiring consciousness and to remain unaware of others. (" All I focused on was winning the race, I didn't even realise we ran past that big fire until I saw the newspaper the next day"). The final feature proposed by Chapman [11] as required for consciousness is the uniquely personal nature of the event and resulting *affect*. 'My back pain is mine alone' as are the resultant emotions and the intensity to which one experiences them.

#### Cognition and the Construction of Pain

Sharp [9] proposes that the current cognitive-behavioural analysis of chronic pain remains too focused on behaviour (e.g. limping, grimacing, rubbing the painful area) and has failed to evolve to the level of cognitive theory in other disciplines, (such as education). This work maintains that cognitive-behavioural therapy needs to be reformulated with a stronger emphasis on the essential nature of the meaning people ascribe to their pain and behaviours subsequent to this meaning construction. Applying cognitive processes within a constructivist model does not deny the existence of physiological or psychological features in the development of chronic pain. Rather, it recognises that through the process of becoming conscious of these features, individuals must do one of two things. Either they are able to develop an acceptable understanding and coping strategy for what is happening or they are unable to integrate the experience and thus remain unresolved and in pain. This lack of resolution serves to increase the discord within an individual's attempts to self-organise, which will then allow a negative cycle of effort and outcome to become established [9]. Additional influences on attempts to resolve the pain event are wide ranging. Personality, support network, an individual's social group, the political climate for health and social benefits, economic issues related to productivity and life-style choices are only a few of the features composing the dynamic internal and external context of a pain experience, which can further destabilize the likelihood of successful resolution.

#### Schemata

The process of filtering sensory input though a network of past experience, influences and memory to determine degree of conscious awareness and subsequent action plan has been called *schemata* and is widely addressed in the clinical and cognitive psychology literature [1]. Schemata involve both cognitive and motor patterns (as involved, for example, in riding a bicycle) and it is believed that each time schemata are retrieved and brought to a conscious level they are altered by the new context and experiences occurring at that point in time. The bicycle riding schemata formed as a child for example, may be dramatically altered when retrieved during an emergency situation when an adult is required to ride a bicycle to get help in an emergency. For successful resolution of this unexpected situation, the individual must be able to modify his or her existing schemata and construct a new meaning and awareness of bicycle riding that allows for different behaviours to occur. The range of internal and external forces that interact when an individual activates personal schemata for pain is vast. It is possible that when there is a barrier to activating consciousness (such as lacking a cohesive explanation for the pain), the individual cannot modify his/her schemata and therefore cannot construct a contextually relevant meaning for the new pain experience. Without this new meaning construct, efforts for resolution will remain ineffective.

#### Consolidation of Theories Through the Consciousness Model

The subsequent discussion of features that present as barriers to schematic modification and construction, will be organised around the four features of consciousness [1] previously introduced, *sense of self, purposiveness, personal nature and affect,* and *coherence.* In the following discussion many studies highlight issues preventing resolution of the chronic pain experience but are not built on a constructivist model. Their theoretical underpinnings are as diverse as the range of issues presented. The following review will serve to illustrate how their diversity can coherently be organised and linked within the four features of consciousness in a constructivist analysis.

#### Sense of Self

There are two general aspects in developing and maintaining a sense of self. The first is physiological (mediated by the body systems) and cortical (particularly related to perceptual functioning in the parietal lobes). Basic functions such as respiration, heart rate, and strength all supply information about one's self. Of equal importance is the brain's ability to translate incoming sensory data into awareness of position in relation to surrounding objects, where one's individual body parts are in relation to other parts (proprioception) and overall body image. Damage to any of the cortical areas responsible for perception can result in an altered or inaccurate sense of one's own body. Examples of these problems are readily evident in the medical literature related to strokes and head injury. The Man Who Mistook His Wife for a Hat, is one widely recognised example [12] of this topic in the popular press.

Chapman and Gaurin [13] propose that persistent stimulation to the sensory pathways can result in sensitization of peripheral nerves and transmitter cells in the spinal cord so that abnormal neural firing patterns and lowered firing threshold occur. A neuroendocrine response to this stressor occurs in the hypothalamo-pituitary-adrenocortical (HPA) complex that will flood the system with maladaptive levels of neuroendocrine substances and stimulate an immune system reaction. The net result of this process can be to trigger an illness response. This response is similar to having influenza- fever, sleepiness, reduced respiration, decreased libido, muscle-ache, negative mood and cognitive functioning depression. Chapman et al outline how this promotes a cyclical process of pain, sensitisation, stressor, more pain and so on. Although a more in-depth discussion of this complex area is beyond the scope of this current work, there is a growing volume of research that supports the proposals outlined above [14, 15].

The second aspect of *self* is that of the social self. How we assume our roles within society is affected by who we are, how others interact with us, what society as a whole expects and what it allows. Cravey and colleagues' [16] recent work explores the concept that self is influenced not only by 'what' society tells us, but also by 'where' the telling occurs. Examining whether place influences peoples' receptiveness to health information (for example, someone who sees himself as a strong, dedicated worker will accept information about back pain more readily from colleagues in the workplace as opposed to the GP's surgery), they concluded that the health information people accessed to understand and explain events occurred in different locations dependant on a variety of socio-economic factors such as employment, access to close family members, and religious homogeneity of an area. People isolated from their usual socio-spatial knowledge networks through events such as loss of job, relocation, or impaired mobility will experience difficulty in adapting to experiences like chronic pain that are perceived as threats to their self-image.

The importance of social context in the construction of self and pain is illustrated through David Morris's thought-provoking story of the 1920s vaudeville performer Edward H. Gibson. Gibson claimed to feel no pain and his stage performance was a graphic demonstration of this (e.g. having himself nailed to a board and allowing the audience to stick hatpins into him). Morris states that, 'Audiences responded to Gibson in ways that tell us something about our acquired habits of thought. A person who cannot feel pain seems a kind of freak or outsider' [17:13]. People's reactions created Gibson as abnormal, a freak. Morris proposes that it is not the experience of having pain that unseats a person's sense of self. Rather, it is other people's interactions with that pain that shape an individual's understanding and behaviours. Pain is not constructed in isolation but rather within a complex social context.

There is an extensive volume of literature examining the social influences on an individual's pain experience. Issues of social conflict, normative influences, socially legitimised disempowerment and social consequences of complaining about pain are frequently recurring themes. Encandela [18] proposes that an individual's scope and creativity in creating a social construct for his or her pain is limited by what is permitted within the social context. The components of social context (for example, ethnicity, age, gender, and socioeconomics) all influence the individual's understanding of pain as it relates to him or herself. He states that the range of interpretations available to an individual is contextually bound and uses a study of two similar retirement communities in the North-eastern United States to substantiate this premise. Encandela reports how living situation was a key influence in how peoples' self-image was normed by the community. In both communities people were expected to be active and pain-free if they lived in unsupported (no nursing assistance) accommodations. These residents expressed a need to 'conceal' their pain and not 'bore other people'. They gave examples of 'being left out' and 'losing your friends' as consequences of not complying with these norms of behaviour. People living in the supported units, in the same community, were expected to be less well and were able to openly verbalise their pain and aspects of ill health. For these people, pain was a more medicalized experience, not something to fight off but rather a justification for decreased autonomy and less active participation in the social expectations of the community. Encandela also discusses how some residents used a form of 'colluded silence' about their pain experiences as a strategy to maintain personal control and privacy.

A study examining instances of when patients chose to consult a physician versus a nurse presents findings in a similar theme [19]. The researchers found that only 14.3% of people with issues related to chronic care (eg. prescription renewal) agreed with the doctor's view as to whether they needed to be seen by a physician or a nurse. They also reported that 27.7% of chronic patients felt the nurse could see them, but in the doctor's opinion only 10.2% were suitable to be seen by the nurse only. Kernick suggests that doctors and patients construe health differently; doctors are more focused on the technical aspects and patients are more concerned with psychosocial issues. If this is the case, patients who only consult with a physician may find their expression of the pain experience channelled in a bio-technical direction regardless of any issues they perceive to be more significant. Help-seeking behaviours can become less direct and the person with pain may alter his or her self-image to coincide with the physician's view as opposed to a more individualised construct. Some research indicates that service users actually have improved self-efficacy and more positive affect when their perception of pain intensity is not congruent with that offered by their physician [20]. Charles et al [21] points out that passive acquiescence to the doctor's opinion is a philosophical underpinning of the paternalistic bio-medical model. Despite calls from healthcare providers, (such as the National Health Service in the UK), to move toward a more informed process of shared decision-making, there continue to be physician, patient and caregiver socially mediated barriers to this change in direction.

Social scientists have proposed that suffering is often politicized into a medical condition. Torture and poverty can become delegitimised as social issues through a process of medicalization [22-24]. The tortured are no longer victims of political oppression and warfare but become people with post-traumatic stress. The solution no longer demands a wider social context; it is now the individual's problem. Kleinman [22] proposes that members of society find it easier to remove themselves from the overwhelming task of fixing social wrongs through medicalization of issues, and this serves to allow the status quo to remain unchallenged. Other authors have drawn similar conclusions regarding women's suffering during life changes [25], and social response to the atrocities of 20<sup>th</sup> century western history [17]. "The genocidal programs of mass murder carried out under Hitler or Pol Pot inevitably affect how we think, we act, how we feel...the mind simply blocks what it cannot comprehend...quite literally we can make nothing of it" [17:51]. These authors maintain that to resolve this inability to comprehend, western cultures have recreated suffering as an individual's own pain and medical problem, requiring personal, not political, resolution. When this social construction of suffering as personal pain comes into conflict with the individual's self perception of having suffered as a result of political activities or social negligence (unemployment, lack of education, poverty, violence), development of a consciousness and an acceptable schema is precluded and the event remains unresolved. When meaning for an event is incongruent with a person's self analysis ('I was a war hero, now I am suspected of trying to cheat the welfare system'), it leaves the individual's self identity vulnerable with no acceptable repertoire of responses and behaviours [26]. Some recent research suggests that when communication issues are interpreted as dehumanising, devaluing and disempowering within the medical encounter, people feel their personal

identity is threatened. When this occurs, reported dissatisfaction with healthcare increases [27].

#### Purposiveness

The second category of factors affecting consciousness is that of *purposiveness* [1]. Purposiveness, underpinned by the basic needs of adaptation and survival, is what captures our attention. The purpose we ascribe to the features of our internal and external environment serve to focus our attention. Focused attention (consciousness) allows for selective activity and efforts to obtain goals. Chapman and colleagues also remind the reader that we do not achieve awareness in a vacuum. Awareness of pain is a mediated state, dependent on the interaction between the individual and the bio-psycho-social context of his or her environment. Research also demonstrates that attention and cognitive functioning is affected by physiological changes during chronic pain and pain related medication. A recent study of 275 chronic pain patients in a large American urban hospital reported 20.5% of the participants stated they had difficulty finishing tasks and 18.7% felt they had problems with attention [28].

This section will review key features of the individual's psychosocial context that serve to direct or filter attention towards the experience of pain. These key features are identified as *pervasiveness, vigilance, expectation*, and *search for cure*. Fife [26] proposes that *pervasiveness* is one of the features in how a person conceptualises the meaning of chronic illness. As the impact of illness on the number of areas of a person's life grows, so does the attention it receives by that person. Fife uses the example of cancer patients who are undergoing chemotherapy. The 38 adult participants in her study reported that a form of compartmentalization is possible; the cancer did not demand all of their attention because it did not affect all aspects of life (e.g. their social life) [26]. However, when treatment caused hair loss, many people reported that their total attention was now focused on the cancer experience. The hair loss was a visible sign of cancer that was carried into all aspects of daily life. The cancer, subsequent to visible hair loss, became pervasive and highly attended to.

A number of recent theorists' work suggests that the pervasiveness of pain as a global feature of life is reinforced through personal vigilance. They propose that once an event (e.g. pain) is brought to awareness, a complex of interactions occur that serve to maintain and actually intensify vigilance to the event [29-32]. Aldrich et al [29] describes a cycle of attention, worry, maladaptive coping and problem solving, inability to resolve worry, increased vigilance, decreased activity for fear of re-injury, increased pain, increased worry and so on. People with pain may perceive that they have limited options to affect change in the situation because of financial, social or cultural issues. The only action available, according to Aldrich, remains worry. The consequence of this escalating vigilance to pain can be both psychological (diminished self-efficacy) and physiological [33]. Diagnostic uncertainty is also seen as a feature in heightened vigilance. Stenner and colleagues [32] examined patient/clinician communication and concluded that an uncertain diagnosis increases some patients' persistence in information seeking. Consequently frustration and distrust increase between both parties, and an ultimate breakdown of the therapeutic relationship occurs.

Lastly, there is the relationship between illness narrative and vigilance to pain. Hyden [30] detailed problems encountered by individuals who, for numerous reasons, were unable to construct an account of their illness. He suggests that narrative is what allows individuals, and the people who listen to the narrative, to organise understanding, formulate response and evaluate actions. When an individual is consumed by attention to pain, the narrative process suffers. Information is in constant flux, understanding is not clear and communication breaks down. Once an individual has become conscious of pain he or she needs to be able to construct an effective narrative for the experience. When this fails, positive self-evaluation and interpersonal relationships, dependent on clear and trust-worthy information exchange, will suffer. The literature regarding doctor-patient communication proposes that these exchanges are already primed for miscommunication secondary to a range of sociocultural variables. Miscommunication in turn has an effect on patient outcomes [31]. For people whose consciousness is consumed by vigilant attention to pain, the negative consequences can be far reaching.

*Expectation*, in terms of what people want and believe should be done to achieve what they want, is a third feature of purposiveness. A discussion of expectation follows logically from the preceding ideas of vigilance. For people with pain, the two are closely interrelated. Contemporary researchers propose that fear of pain and the consequent avoidance of potentially painful activities serve to heighten certain individuals' attention and consciousness of pain [29, 34-36]. Eccleston and Crombez [34] state that pain, through its novelty, interrupts attention to other aspects of daily life. For individuals who are unable to put the pain in an understandable and acceptable context, the pain experience remains dominant and does not allow functional attention to other aspects of life. People in this scenario of pain's chronic interruption will experience a range of negative consequences, one of which may be heightened fear of doing anything to increase the already pervasive pain experience [34]. This expectation of pain, paradoxically, sets up a self-perpetuating cycle of dysfunctional activity avoidance, biological deconditioning and escalating fear of activity.

Another important aspect of understanding how expectation heightens the intrusion of pain into consciousness is the patient/service provision dynamic. Research exploring the area of patient expectations has suggested that where expectations are incongruent with service provision, outcomes will be poorer [37, 38] and that dissatisfaction with service provision can have an iatrogenic affect [39]. Other research, although it does not deal specifically with a chronic pain population, shows parallels between the characteristics of social and financial marginalization. A particularly relevant example is the finding from a critical review of empirical studies of the experience of people with long standing psychiatric conditions that revealed what people expect is not necessarily the same as what they want and that people often use indirect routes to communicate their wishes to clinicians [37].

Search for cure is a further theme emerging from the literature regarding peoples' attention to the pain experience. Vrancken's provocative review, *Schools of Thought on Pain* [24] proposes that in some pain experiences the individual is unexpectedly forced to recognise (be conscious of) the body in a way not previously experienced. A part of the body (Vrancken uses the example of a painful foot) suddenly consumes 'here-and-now' awareness. This in turn, may restrict awareness of other life events and interfere with personal self-concept and the individual now stops being a person and instead becomes a pain patient. The patient assumes the role of searching for cure, fulfilling his or her need to take action, but also becoming all consumed by the search. Stenner et al [32] relates the search for cure to how

resistant a condition is to understanding. They suggest that in disorders with uncertain aetiologies, people respond through hopeless resignation or anxious, persistent, attention consuming search for cure. Stenner and colleagues compare their study of sixty people with Irritable Bowel Syndrome (IBS) with a study of chronic pain patients [40] and report a strong resemblance between the two studies' participants in their belief that there is an undetected physical cause for their current illness. Perceived efforts by health care providers to move patients away from a search for cure towards a problem management model are strongly resisted and Stenner et al. conclude with a caution for clinicians to develop more sensitive communication skills to prevent further deterioration of the healthcare relationship.

Finally, the issue of *self-awareness* should not be overlooked in this exploration of the role purposiveness plays in consciousness. Steen and Haugli [41] propose that chronic pain programmes should be based on the assumption that people must be *taught* how to be truly conscious of their body. Rather than allowing unstructured and conflicting perceptions to be forced upon the individual trying to come to terms with the painful body's intrusion on daily life, they suggest that people with pain be shown how to reflexively tap into self-knowledge based on past experience and in that way arrive at an acceptable understanding for the meaning of current event. Their analysis draws heavily on Dewey's [42] concept of awareness and Leder's [43] proposal of the split between the body and self, leading some people to experience their body as 'separate' and alien [41]. Steen's work, based on a programme of twelve sessions, every fortnight for four hours with between eight and ten participants, reports that participants learned to be aware of their bodies through a variety of techniques - movement, metaphor, guided imagery, art and written approaches to self exploration. This randomized control trail found that the intervention group reported less pain and better coping ( $P \le 0.05$ ) as compared to the control group at one year follow-up [41]. Unfortunately, the training of the programme leaders and the outcome measures employed are not identified in the paper. Despite these short falls, the work points in an interesting direction of facilitating the individual with pain's access and control over aspects of pain consciousness

#### The Personal Nature of Pain

The third component of the Consciousness model is *affect*, the emotional interpretation each individual places on the pain event. The social science literature is rich with narratives and research related to the affective dimension of the pain experience. *The Culture of Pain* [17] and *The Body in Pain* [23] can be considered seminal works in any exploration of the affective nature of pain. Morris relates a complex relationship between pain and society and gives insightful examples of how pain emotions, both negative (fear, suffering, loss, stigma) and positive (pleasure, justice, penance), are contextually bound to social events at any given time. Scarry's work explores the political and spiritual dimensions of pain and suffering. She contends that western society lacks the ability to communicate the experience of pain, making it easy to discount in others. '*To have pain is to have certainty, to hear about another's pain is to have doubt*' [23:4]. This section will discuss *fear of pain* and the *consequences of others' disbelief* as influences on the affective domain of consciousness.

Research highlights the complex relationship that exists between pain and fear. Increased pain can be a consequence of an individual acting on his or her fear of re-injury and therefore

becoming deconditioned and physiologically vulnerable [29, 34, 36]. An individual may use a personal interpretation of how severe the pain is to evaluate the degree of concern or fear that should be expressed, and the 'perceived threat or damage to one's biological integrity' [13]. Other research has focused on the wider range of consequences related to fear and anxiety about pain. Aldrich's review of this literature highlights that the outcome of pain and fear may be decreased self-confidence, inability to set future goals, deterioration of support systems, work incompetence and maladaptive problem-solving attempts [29]. Fear of pain is compounded by fear of social isolation, economic hardship and an uncertain future. Eccleston et al [40], suggest that these additional fears do not remain as separately identified issues but rather are amalgamated and become contributory to the intensity of pain experienced. They propose that this occurs because of how strongly people with pain endorse the dominant cultural belief that pain is attributable to disease or illness. In other words, people with pain believe that increased pain is caused by worsening physical condition and not by becoming unemployed. Chapman's recent review of the relationship between physiology, consciousness and pain examines how, once pain is brought to a conscious level and an affective response (like fear) is triggered, physiological mechanisms (responding to perceived biological threat) then serve to maintain consciousness and the pain experience. The mechanisms for this defence response are the physiological processes occurring at the hypothalamopituitaryadrenocortical axis (HPA) [33].

Affect is also influenced by *disbelief*. Social science research has highlighted the power imbalance between people with health conditions and service providers. This imbalance impacts on communication, satisfaction with services and concordance with treatment [31]. Ong et al., propose that when there are communication problems, patients will rely on their personal interpretation of non-verbal behaviours and that these interpretations tend to be negative as a consequence of the heightened affect experienced during health related interactions. In one study of ninty-six chronic pain patients at a university-based American hospital, 74% of the participants reported feeling angry towards themselves and 64% stated they were angry at healthcare professionals [44]. The researchers propose that there are significant secondary issues consequent to anger; depression, isolation, and perceptions of decreased functional ability. Their findings suggest that anger is targeted and in some cases can actually be helpful to motivate action (for example, more committed participation in rehabilitation and engagement in advocacy related activities).

It is also possible that this power imbalance contributes to people feeling disbelieved by service providers. Kouyanou et al., looked at iatrogenic factors in chronic pain and found that in a two-site sample of one hundred and twenty-five chronic pain outpatients, 25% reported the doctor had directly disputed the patient's reported pain intensity and that 41% stated they were made to feel that the doctor doubted their report [45]. Another study, looking at the communication issues between doctors and patients with irritable bowel syndrome (IBS), suggests that physicians dealing with difficult to diagnosis conditions like IBS may be frustrated when the condition proves resistant to treatment. In cases such as this, personal beliefs may over-ride professional 'knowledge' and behaviours, resulting in an adversarial relationship between patient and physician. The doctor blames the patient for failure to respond to treatment and the patient comes to believe the physician doubts their honesty [46]. Although this study focused on a different clinical group, it bears examination because IBS is similar to chronic pain in its resistance to diagnosis and treatment. Pain-specific research has revealed similar issues with physician frustration and patients' feelings of being distrusted

and blamed [40]. Sullivan proposes that while patients bring their first-person accounts (the experience of pain) to the encounter, the physician translates their accounts into the thirdperson (pain as a disease entity). When the translation breaks down for some reason, the firstperson account, by nature of its subjectivity, becomes suspect and open to doubt. Sullivan identifies patient behaviours that are incongruent with physician expectations, social stereotypes, and the social negotiation process around accounts of other people's pain, as being key barriers to the move from subjective reports to professionally recognised disease [47].

From the preceding discussion we see that disbelief, and the emotions it engenders, presents as a significant issue in maintaining pain at a conscious level. Chapman reiterates that emotions related to pain are individual and often difficult to communicate [1]. It is this very individualism that places such a barrier on others' comprehension, and for the person with pain the consequent weight of negative emotion serves to reinforce his or her consciousness of pain.

#### Coherence

The final category of consciousness is *coherence*. Chapman et al (2001) propose that consciousness is self-organising and focuses effort towards integrating the fragmented internal and external elements that come to awareness. This integration is required for the individual to assign meaning to events and experiences, relying on past understanding and awareness of events across a time span. Chapman calls this 'our model of reality' [11]. The literature on chronic pain is rich with discussion related to meaning construction and the socio-political and biological factors that present barriers to an individual arriving at a coherent meaning for pain that is congruent with past experience and beliefs.

Many features influence coherence and there is a wealth of literature from the social and biological sciences. For example one influence that occurs with regularity across the chronic pain literature is that of misinformation, lack of information [48], and withholding of information as a political strategy by health care providers concerned with maintaining the professional status quo [49]. Hunt et al. [50], demonstrated that people's concept of illness has a strong temporal component and that their explanatory models for illness fluctuated over time. Interestingly, these researchers conclude that changes people make to their explanation for an illness are more strongly linked to changes in personal circumstances and interactions within the social environment than to the influence of physician/patient interactions. The importance of time is also highlighted by Jenson and Kern's [51, 52] work on pain stages of change. In this model people's ability to take on new information and beliefs is related to where they are in the readiness for change stages (precontemplative, contemplative, preparation, action or maintenance). Stages of Change research suggests that for treatments to be effective they must be matched to the individual's current stage [51, 52].

A third feature of coherence relates to control of treatment decision-making. A number of authors propose that there is a significant relationship between increased perception of control and treatment outcomes [21, 53-55]. There are a number of influences on peoples' feelings of control; family and social network, access to accurate information, and the perceived seriousness of a problem, balanced against the effort required to exert control. On the service providers' part are issues of professional identity, time for information sharing and the

expense of resources to do this work [21]. Additionally, as Bailis [53] points out, there are varying degrees of ability and desire to assume control and it may be that the impact of social inequality needs to be addressed as a precursor to shared decision-making in the clinical setting.

Perhaps the area of coherence that has received the most attention is that of differing beliefs and values between service providers (SPs) and service users (SUs). The meaning a person ascribes to his or her pain is constructed through a complex interplay of socio-cultural and political factors including class, gender, economic environment, socialisation and the family. 'Human pain does not exist apart from meaning' [56 p.118]. What we are now seeing is that, in the same way as each individual with pain formulates a personally congruent meaning for the pain experience, so do service providers develop a meaning for the service user's pain. This meaning derives from an individual's own past experience, socio-cultural background, political and economic demands and other environmental contexts of the service user. It is now well demonstrated that the meanings that are constructed are not necessarily shared [10, 19, 30, 32, 46, 57-60].

In the study of SPs' and SUs' beliefs about irritable bowel syndrome (IBS) mentioned previously [32] disagreement was found between service users' and providers' beliefs regarding aetiology of the disease, the role of diet, exercise and psychological issues such as stress and childhood trauma. Other researchers have proposed that medical practitioners, while inculcated to be 'detached observers', are in fact as strongly influenced by their own subjective biases and beliefs as by the scientific evidence [10]. It has been suggested that even the assessment tools used by clinicians are flawed with this bias because they are designed by service providers and predominantly reflect what the tool creator values as opposed to the service user's desired outcomes [60].

Some research has found that while SUs rated areas like size and accessibility of the practice, the variety of treatment options offered, and social issues (such as housing and employment), as important aspects of service delivery, these areas were judged as less significant by SPs [58]. A focus group study of SUs and SPs found that, although there was agreement about the 10 most important topics in healthcare service provision, there was disagreement about the priority of each topic. For example, only 7% of service users rated the physician's knowledge as most important where as 17% of the physicians' group rated knowledge as most important [57].

In summary, Chapman [1] proposes that within a constructivist framework pain is a conscious process and that features of consciousness can either act as barriers or as facilitators of successful resolution. Some researchers propose that in addition to the previously identified barriers to meaning construction, an increasing awareness of the limitations of current medical 'knowledge' creates discord. A growing voice in the healthcare literature proposes that conventional medicine can no longer be seen as the definitive reality and that social values now reflect a shift towards holistic thinking where consumer rights, access to information via electronic media, and self-responsibility are seen by many as essential components of well-being [61-63]. Individuals attribute a range of meaning to the events they experience, interpret others' reactions as either desirable or negative and than adjust their behaviour accordingly. These theories are consistent with postmodernist thinking which proposes there are many realities depending on the lens through which an individual views the world [63]. This concept underlines that it is not only service users but also healthcare workers who are subject to the social forces in which healthcare service is provided. In

postmodernist thinking there are no universal givens and personal reality changes in relation to the socio-political, temporal and environmental context in which an individual lives [64]. The emerging social phenomenon in western cultures of challenging biomedicine's superiority has seen a growing exploration and endorsement of interventions that previously were held to be from the fringe, and 'alternative'[64-66]. Personal meaning construction eludes people as they struggle to resolve these concepts within their personal belief systems and the internal conflict that arises around a biomedical system that controls a large component of their access to care.

The preceding discussion has attempted to demonstrate how research evidence from varied schools of thought can be incorporated within a constructivist framework and the conceptualization of pain consciousness. Lastly, the emerging influence of post-modernist thinking was identified as a potentially strong influence on individual pain constructs.

The following segments of this chapter will present a research project [67-69] that employed Delphi methods to explore the usefulness of applying Chapman's Consciousness model to understand service users' and providers' decision-making in relation endorsement of treatment components for chronic pain.

#### Influences on Decision-Making

Influences on decision-making in chronic pain management were examined as part of a larger four-stage study exploring the congruence between SUs' and SPs' beliefs about treatments [67, 70, 71]. A survey of beliefs about a range of chronic pain treatment components was carried out in stage one. Specifically, participants were asked to complete a postal survey identifying whether they believed each of the sixty-two treatment options (extracted during a scoping exercise from the ain literature) were 'Important' or 'Not *important*' for chronic pain. The concept of 'importance' was carefully considered. To define 'importance' specifically (e.g. 'reducing pain', 'being able to go back to work'), potentially denied individuals the opportunity to self-define within the context of their own goals and pain experience. The rationale for using the concept of 'importance' related to the phenomenon of people making decisions based on their personal assessment of over-all desirability ('importance') as opposed to the more narrow concept of benefit to health. Several models of health-related behaviours suggest that people attach a relative value (importance) to health outcomes that then influences their degree of participation and compliance with treatment activities [72]. These models propose that people may acknowledge a certain behaviour is of benefit (for example, weight loss) but at the same time assign a low level of importance to the behaviour in the context of their own lives. Additionally, focusing on *importance* allowed a more open-ended approach so that individuals could respond from a values and beliefs perspective.

Stages two through four of the study employed a Delphi design where participants' responses were relayed back to them in written report form for validation via successive rounds of postal questionnaires. The iterative process of the Delphi method is a key strength as it allows participants an opportunity to review and reflect on their own and other participants' prior statements [73, 74]. In stage two participants were asked to identify their level of agreement with a series of statements about influences on their decision-making

171

regarding what chronic pain treatments are important. The statements were derived from the four domains (*coherence, purposiveness, self-image and affect*) presented in Chapman's model of Consciousness [1]. The ordinal scale for the questions was anchored at '*I agree*... '0- *not at all*' through to '4-*completely*'.

## Participants

For the multi-stage study a sample of convenience was recruited from the National Occupational Therapy Pain Association and the British Pain Society (multidisciplinary membership comprised of anaesthetists, general practitioners, psychologists, nurses, researchers and other allied health professionals). The stage one questionnaire was distributed to all members of these groups (N=1426) and 237 (16.6%) of the original sample volunteered for the more extensive Delphi stages two through four. By stage three 18 participants had dropped out (moved without forwarding address, deceased, or retired) leaving a study population of 209 service providers from a range of healthcare backgrounds for stage three.

## Delphi Stage Two

Delphi stage two used a postal questionnaire to determine participants' validation of the researcher's conclusions from previous rounds. The Delphi stage 2 questionnaire was distributed and a reminder letter and duplicate questionnaire was sent three weeks later to the SUs who had not yet replied. In total ninety (40.9%) SPs and sixteen (50%) SUs replied. The demographic breakdown of respondents is presented in Box 1.

Box 1: Stage two Delphi questionnaire respondents' profile				
Respondent	Age (mean	Years of experience	Female	Male
	years)			
Service User (16)	54.2	12.2	9 (56.3%)	7 (43.8%)
Service Provider (90)	44.6	19.5	47 (56.0%)	34 (40.5%)
Distribution of response between professional groups				
occupational therapist	10 (11.1%)			
physiotherapist	8 (8.8%)			
general practitioner	8 (8.8%)			
nurse	24 (26.6%)			
anaesthetist	25 (27.7%)			
psychologist	5 (5.5%)			
other & not stated	10 (11.1%)			
TOTAL	90			

## VALIDATING THE CONSCIOUSNESS MODEL

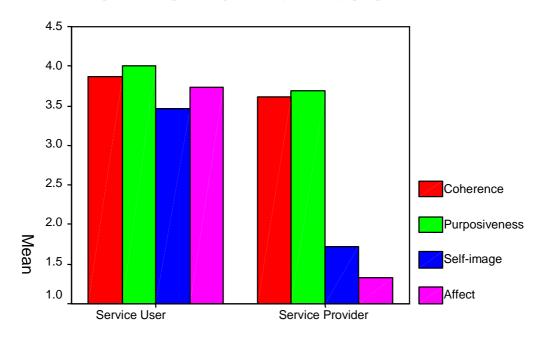
#### Decisions about 'Important'

In Part 1 of the stage two Delphi questionnaire, participants were presented with four themes representing the reasons why people decided certain treatments were important or not. The themes (*Coherence, Purposiveness, Self-image and Affect*) were derived from Chapman's model of Consciousness and pain [1] as discussed at the beginning of this chapter. Participants were presented with a list of the ten treatment components most frequently selected as *important* by their cohort. They were asked, on an ordinal scale (anchored at '0-not at all' through to '4-completely'), to identify how closely each of the four themes matched their reasons for thinking a treatment was important. Because of the small sample size of service users (16) the data were recoded into three categories ('0-not at all/a little', '2 – 'somewhat' and '4 – mostly/completely').

An independent samples t-test was conducted to compare service users' and providers' responses to the four themes. There were statistically significant differences in the scores for *Purposiveness* [(t (degrees of freedom 78) = 3.167; p=.002], *Self-image* [t (31.75) = 5.872; p= <.000 ] and Affect [t (28.764) = 7.417; p=<.000]. There was no significant difference for Coherence. Several interesting patterns emerged in the analysis (Figure 1). While the majority of service users agreed that all four themes influenced their decision-making, the service providers' responses showed strong rejection of the themes of *self-image* and *affect* as having any influence on their decision-making. Only 20.5% of service providers rated the theme of self-image (what someone like me would think) as closely matching their reasons for thinking a treatment was important or not. Service provider ratings for the influence of affect (how this treatment makes me feel) were similarly low with only 19.4% of respondents selecting a rating of 3 or 4. This is in marked contrast to the service user ratings where 73.3% selected self-image and 92.9% selected affect as a strong influence. In summary, the stage two Delphi questionnaire results showed that SUs' and SPs' beliefs about influences on their decisionmaking about treatment components for chronic pain were not congruent. SUs thought that all four of Chapman's themes within the Consciousness model (Purposiveness, Coherence, Self*image* and *Affect*) [1] influence their decision-making. However, SPs stated they are for the most part only influenced by Purposiveness and Coherence.

Based on these and other findings from the larger study reported elsewhere [67, 69, 70], a decision-making model was constructed. This decision-making model proposes that when a treatment option is presented both the service user and the service provider engage in an iterative cycle of evaluation, decision-making and testing. The treatment option is evaluated against what the individual already knows and believes. The four themes of consciousness [11]; Coherence ('does the treatment option make sense with what I already know), Purposiveness ('do I understand the purpose of this treatment'), Affect ('how do I feel about this treatment'), and Self-image ('would someone like me think this treatment was important'), characterize this process. A wide range of personal, public and contextual influence moderates individual responds to these questions and each person generates his or her meaning constructions based on a mixture of these shared and idiosyncratic influences. Service providers and service users have different roles and self-images that influence, and are influenced by, their expectations of a treatment and what factors most affect their

consciousness of the experience. As highlighted in Figure 2 SPs and SUs experience different influences on consciousness and the expected personal outcomes of their decisions about a pain treatment component. The other people and types of relationships in SUs and SPs environments are different (other service providers and professional colleagues as opposed to family/friends and social relationships). Socio-cultural roles for many SUs and SPs are also different, encouraging a dependent (patient/passive versus professional/action oriented) relationship. The bullet pointed items under the headings of *Influences on consciousness* and, particularly, *Personal outcomes* (points A and B in Figure 2) provide examples illustrating that people with pain have access to influences contingent on the experience of 'having' pain where as service providers experience pain through a 'doing' perspective [17].



Delphi 2

Figure 1. Modelling the meaning construction cycle.

Figure2 graphically portrays the relationship between two primary moderators of the influences on consciousness. The first is labelled *Control* (C) and represents the degree of control people experience (or perceive they have) [75, 76], over the outcome of actions and behaviours based on their meaning constructions. People with control are more able to engage in new behaviours, [50, 77], in turn generating additional novel information or confirming existing information that then acts as feedback to the meaning construction cycle. By nature of the role assigned to them, service providers have (and/or are perceived to have) a much higher degree of control over treatment options than service users [78-80].

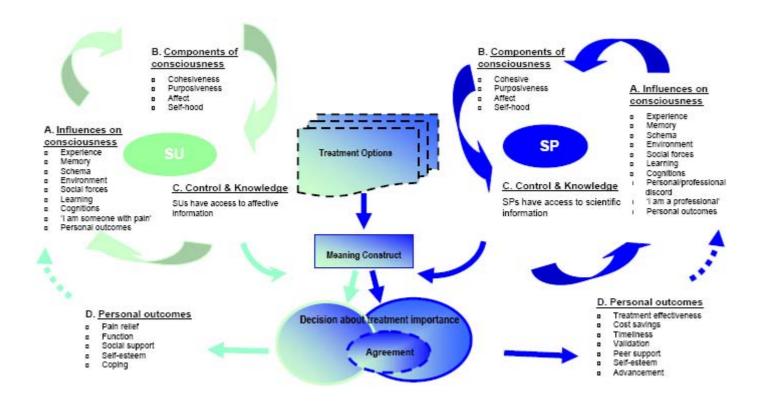


Figure 2. Control & Knowledge: decision-making process related to chronic pain treatment importance.

The second moderator proposed in the modelis labelled *Knowledge* (C). Schwandt states that the quality of constructions depend on both the degree and scope of information available and on the individual's ability to deal with that information [6]. To have one without the other can interfere with an individual's ability to formulate a construction and subsequently initiate new behaviour. Without new behaviour the individual has limited opportunity for feedback and therefore remains unable to evaluate and modify the existing construction in light of the new information afforded by responses to their actions. The model reflects that SUs predominantly have access to knowledge in the affective and personal domains where as service providers have access to scientific knowledge. The findings presented previously in this chapter and in the literature support that service users feel that, despite concerted efforts from healthcare organizations to improve patient-education resources [78, 81, 82], access to information and education about pain are perceived as unmet needs. Consequently, service users are unable to formulate meaning-constructs, and subsequent behaviours, because they lack sufficient scope of information. Similarly, service providers, while having a high degree of access to the scientific information, have limited structured opportunity to gain information about the affective domain of the chronic pain experience. Barriers to accessing affective information include, for example, little to no education in communication skills, insufficient time within appointments, competing administrative demands, and a professional ethos that teaches healthcare students to preserve personal distance and suppress emotion [83-87]. This supports the proposition that the meaning constructs of both service users and service providers are based predominately in information that is readily accessible to them, as opposed to a fuller perspective. Additionally, the domain of accessible information is largely unshared between the two groups, thereby introducing a high degree of variation into the meaning construction.

The model highlights that the cycle begins with identification of treatment options and subsequent meaning construction is moderated by the influences discussed above. Decisions about what treatments are important will then influence SU and SP behaviours and lead to a range of subsequent outcomes. The bottom left and right corners in Figure 2 (D) represent the different outcomes that SUs and SPs seek to achieve subsequent to behaviour based on decisions about which treatments are important. Service user outcomes are highly personal/private in nature and involve a range of physical, emotional and social elements (can I get back to work now?, will I be able to sleep again?, do my friends think I'm 'faking'?). Service provider outcomes also involve this complex array of elements and have a personal/private context to their meaning construction ('this job is too stressful, I can't juggle family and work full-time'). However, service providers' outcomes are also highly contextualised in the public/professional domain ('would the consultant agree with me'?, 'this patient makes me so frustrated', 'how can I defend the cost of this treatment?',' this patient thinks he knows more than I do!'). Morris reminds us that, "Pain seems the quintessential solitary experience....The isolation of pain is undeniable. Yet it is thus especially important to recognise that pain is always deeply social"[17:38]. Different constructs create different decisions and, in turn, produce different outcomes. These outcomes provide feedback to the meaning construct cycle thus sustaining an iterative and dynamic process.

Finally, the model illustrates the state of agreement between service users and providers. The *Agreement* element (centre-bottom of Figure 2) shows more agreement overlap within the treatment components that the service providers have identified as important. The model represents that, while service users will increase their agreement with service providers about which treatment components are important, service providers will not show the same trend towards increased endorsement. This conclusion is based on the existing imbalance in access to information between service users and providers. Service users, through various organizationally sponsored patient-education initiatives and the widening availability of high quality health information on the internet [88], are more likely to be successful in accessing a degree of the scientific knowledge they previously did not have. Service providers, on the other hand, may struggle to maintain what limited access to the affective domain of their patients' pain experiences they currently have. In Friends in Low Places, Willis suggests that efforts to continuously improve medical care have spawned a myriad of purportedly 'small' but essential directives and procedures. Cumulatively, these 'small' changes create an overwhelming demand on service providers and drain time and resources away from direct service provision. The demands on service providers seem to grow continually, and in each day's post, "...there is another bundle of glossy brochures about the latest additions to the ranks of mysteriously entitled bodies jostling with each other to regulate us..." [89]. As a solution he suggests, perhaps not so facetiously, that there needs to be yet one more new government branch- The MOLWA (Ministry of Leaving Well Enough Alone) [90].

The Dynamic Nature o Decisions and Opinions

"Perspectives change, and we must give the bag a good shake and see what happens"

Michael Ramsey, Archbishop of Canterbury [in 91]

Some of the findings presented earlier in this chapter suggest that for service users, an iterative process of formulating a meaning construct that influences their decision-making, is occurring. They then act on that decision and incorporate the consequent feedback into a revised meaning construction. The findings do not so clearly support this process for service providers. A dichotomy appeared in the findings regarding participants' reflection on applying Chapman's Consciousness model [1] as an organisational framework for their decision-making. Service users' responses clearly supported all four themes of consciousness. However, service providers were equally as clear in their statements that *Self-image* and *Affect* were not strong influences in their decision-making (Figures 1 and 2). This opinion is contrary to the growing literature identifying the inherent and potentially unavoidable impact of affect in relationships between service users and providers [31, 87, 92, 93].

It appears that while Chapman's constructivist theory about consciousness and decisionmaking is a useful explanatory model for the service users, it may be insufficient for fully understanding the service providers. Service providers in this study were strongly influenced by assumptions of objectivity that are more congruent with a scientific positivist paradigm. A useful explanatory model would need to more comfortably accommodate this perspective.

#### Implications of Constructivism/ Scientific Positivism in Discord

What does all this mean? Guba and Lincoln [94] remind the reader that a paradigm represents an individual's basic beliefs and way of viewing the world. They also propose that the ongoing debate about paradigms illustrates that there is no one way to define the 'truthfulness' of these sets of beliefs. Constructivist theorists propose that even the positivist paradigm is a construction based on the individual's interpretation of the available information [94]. Scientific positivists maintain that while social forces may influence events, the world is, to the greatest extent, composed of objective reality and universal truths [95]. Scientific positivism are only two paradigms amongst many that allow one to organise thoughts about the world and experiences within it. The point is not so much that different paradigms are at play, that in itself need not be problematic. Rather, the concerning issue is; 'what happens when potentially different world-views exist for different groups in a shared experience like chronic pain?'

Reframing the issue as differing paradigms (as opposed to differing opinions) helps illustrate why, despite concerted effort and expenditure by healthcare organizations to increase the amount and quality of patient education available through programmes like NHS Direct [96] and Patient.UK [97], opinions remain resistant to change and continue to collide. No amount of 'education' will change opinion if the underlying assumption is that everyone derives his or her ideas from the same basic shared belief system. It is important to know that participants in this study did not agree and that they seemed at times to be operating from different paradigms. It is important because despite the volume of research and resource devoted to the problem of chronic pain it is still, in many aspects, as much a puzzle as ever. Some peoples' pain responds to a biomedical approach, others' to a more psychosocial approach, and some peoples' pain seems determined to remain enigmatic in the face of all manners of intervention. It is evident that chronic pain does not follow a straightforward path and the search for the 'right' way forward may paradoxically be a significant part of the problem. "If things were that simple, word would have gotten round" Derrida in [66]. A growing voice in the literature suggests that we are seeking the wrong solutions to the wrong question. It is proposed that healthcare should focus on relationships between belief systems as opposed to 'solutions' generated within traditional positivist medical thinking, or what has been called the Paradigm of Order [98].

We've seen that there is a lack of congruence between what service users and service providers believe. This lack of agreement in itself is value-free; it is neither a positive nor a negative finding. Rather, the important consideration must be understanding the implication of this inevitable difference and in developing awareness of how these differences, between and within groups of service users and providers, can be influenced so as to harness the cognitive and emotional energy generated through discordant views. Discord (like musical notes) can be constructive and creative like jazz music or a cacophonous and painful event driving away supporters and participants alike. Exciting new paradigms like complex adaptive systems theory and transdisciplinary theory hold promise as vehicles to achieve new insights and actions in dealing with a currently overwhelming healthcare crisis like chronic pain.

## REFERENCES

- [1] Chapman, R.C., Y. Nakamura, and L. Flores, *Chronic pain and consciousness: a constructivist perspective*, in *Psychosocial factors in pain*, R.J. Gatchel and D.C. Turk, Editors. 1999, Guilford Press: New York.
- [2] Ziglio, E., *The Delphi Method and its Contribution to Decision-Making*, in *Gazing into the oracle*, M. Adler and E. Ziglio, Editors. 1996, Jessica Kingsley Publishers: London. p. 3-33.
- [3] Albrecht, G., S. Freeman, and N. Higginbotham, *Complexity and human health: the case for a transdisciplinary paradigm.* Cult Med Psychiatry, 1998. 22(1): p. 55-92.
- [4] Plsek, P.E. and T. Greenhalgh, *Complexity science: The challenge of complexity in health care.* BMJ, 2001. 323(7313): p. 625-8.
- [5] Caudill-Slosberg, M.A., L.M. Schwartz, and S. Woloshin, Office visits and analgesic prescriptions for musculoskeletal pain in US: 1980 vs. 2000. Pain, 2004. 109(3): p. 514-9.
- [6] Schwandt, T.A., Constructivist, interpretivist approaches to human inquiry, in Handbook of Qualitative Research, N.K. Denzin and Y.S. Lincoln, Editors. 1994, Sage Publications: London. p. 118-137.
- [7] Alderson, P., The importance of theories in health care. BMJ, 1998. 317(7164): p. 1007-10.
- [8] Cupchik, G. Constructivist Realism: An Ontology The Encompasses Positivist Approaches to the Social Sciences. 2001 February 2001 [cited 2001 May 17,2001]; 2(1):[Available from: http://qualitative-research.net/fqs/fqs-eng.htm.
- [9] Sharp, T., *Chronic Pain: a reformulation of the cognitive-behavioural model.* Behavioural Research and Therapy, 2001. 39: p. 787-800.
- [10] Wilson, H.J., The myth of objectivity: is medicine moving towards a social constructivist medical paradigm? Fam Pract, 2000. 17(2): p. 203-9.
- [11] Chapman, R.C. and Y. Nakamura, *Pain and consciousness: a constructivist approach*. Pain Forum, 1999. 8(3): p. 113-123.
- [12] Sacks, O., The man who mistook his wife for a hat. 1989, New York: Touchstone.
- [13] Chapman, R.C. and J. Gaurin, *Suffering: the contributions of persistent pain*. The Lancet, 1999. 353: p. 2233-2237.
- [14] Loeser, J.D. and R. Melzack, Pain: an overview. Lancet, 1999. 353(9164): p. 1607-9.
- [15] Melzack, R., Pain and Stress: A New Perspective, in Psychosocial Factors in Pain: Critical Perspectives, R.J. Gatchel and D.C. Turk, Editors. 1996, Guilford Press: New York.
- [16] Cravey, A., et al., Developing socio-spatial knowledge networks: a qualitative methodology for chronic disease prevention. Social Science and Medicine, 2001. 52: p. 1763-1775.
- [17] Morris, D.B., *The Culture of Pain*. 1991, Berkeley: University of California Press.
- [18] Encandela, J., Social construction of pain and aging: individual artfulness within interpretive structures. Symbolic Interaction, 1997. 20(3): p. 251-273.
- [19] Kernick, D.P., D.M. Reinhold, and A. Mitchell, *How should patients consult? A study of the differences in viewpoint between doctors and patients.* Fam Pract, 1999. 16(6): p. 562-5.

- [20] Cremeans-Smith, J.K., et al., Spouses' and physicians' perceptions of pain severity in older women with osteoarthritis: dyadic agreement and patients' well-being. Pain, 2003. 106(1-2): p. 27-34.
- [21] Charles, C., T. Whelan, and A. Gafni, *What do we mean by partnership in making decisions about treatment?* BMJ, 1999. 319(7212): p. 780-2.
- [22] Kleinman, A., Suffering and its professional transformation: towards an ethnography of interpersonal experience. Culture, Medicine & Psychiatry, 1991. 15(3): p. 275-301.
- [23] Scarry, E., *The body in pain: the making and unmaking of the world*. 1985, Oxford: Oxford University Press.
- [24] Vrancken, M.A.E., Schools of thought on pain. Social Science and Medicine, 1989. 29(3): p. 435-444.
- [25] Goldstein, D.E., 'When ovaries retire': contrasting women's experiences with feminist and medical models of menopause. Health Technol Assess, 2000. 4(3): p. 309-323.
- [26] Fife, B.L., The conceptualization of meaning in illness. Soc Sci Med, 1994. 38(2): p. 309-16.
- [27] Coyle, J., *Exploring the meaning of 'dissatisfaction' with health care: the importance of 'personal identity threat'*. Sociology of Health & Illness, 1999. 21(1): p. 95-124.
- [28] McCracken, L.M. and G.L. Iverson, Predicting complaints of impaired cognitive functioning in patients with chronic pain. J Pain Symptom Manage, 2001. 21(5): p. 392-6.
- [29] Aldrich, S., C. Eccleston, and G. Crombez, *Worrying about chronic pain: vigilance to threat and misdirected problem solving.* Behav Res Ther, 2000. 38(5): p. 457-70.
- [30] Hyden, L.-C., *Illness and Narrative*. Sociology of Health & Illness, 1997. 19(1): p. 48-69.
- [31] Ong, L.M., et al., Doctor-patient communication: a review of the literature. Soc Sci Med, 1995. 40(7): p. 903-18.
- [32] Stenner, P.H., C.P. Dancey, and S. Watts, *The understanding of their illness amongst people with irritable bowel syndrome: a Q methodological study [In Process Citation]*. Soc Sci Med, 2000. 51(3): p. 439-52.
- [33] Chapman, R.C. and Y. Nakamura, A passion of the soul: an introduction to pain for consciousness researchers. Consciousness and Cognition, 1999. 8: p. 391-422.
- [34] Eccleston, C. and G. Crombez, *Pain demands attention: a cognitive-affective model of the interruptive function of pain.* Psychol Bull, 1999. 125(3): p. 356-66.
- [35] Gibson, L. and J. Strong, Assessment of psychosocial factors in functional capacity evaluation of clients with chronic back pain. British Journal of Occupational Therapy, 1998. 61(9): p. 399-404.
- [36] Strahl, C., R.A. Kleinknecht, and D.L. Dinnel, *The role of pain anxiety, coping, and pain self-efficacy in rheumatoid arthritis patient functioning*. Behav Res Ther, 2000. 38(9): p. 863-73.
- [37] Noble, L.M., B.C. Douglas, and S.P. Newman, What do patients expect of psychiatric services? A systematic and critical review of empirical studies. Soc Sci Med, 2001. 52(7): p. 985-98.
- [38] Starfield, B., et al., *The influence of patient-practitioner agreement on outcome of care*. Am J Public Health, 1981. 71(2): p. 127-31.
- [39] Main, C.J., H.L. Richards, and D.G. Fortune, *Why put new wine in old bottles: the need for a biopsychosocial approach to the assessment, treatment, and understanding of*

*unexplained and explained symptoms in medicine.* J Psychosom Res, 2000. 48(6): p. 511-4.

- [40] Eccleston, C., A.C.d.C. Williams, and W. Stainton Rogers, *Patients' and professionals' understandings of the causes of chronic pain: blame, responsibility and identity protection.* Social Science and Medicine, 1997. 45(5): p. 699-709.
- [41] Steen, E. and L. Haugli, *From pain to self-awareness a qualitative analysis of the significance of group participation for persons with chronic musculoskeletal pain.* Patient Educ Couns, 2001. 42(1): p. 35-46.
- [42] Dewey, J., *Reconstruction in philosophy*. 1957, Boston: Beacon Press.
- [43] Leder, D., *The absent body*. 1990, Chicago: University of Chicago Press.
- [44] Okifuji, A., D.C. Turk, and S.L. Curran, Anger in chronic pain: investigations of anger targets and intensity. J Psychosom Res, 1999. 47(1): p. 1-12.
- [45] Kouyanou, K., C.E. Pither, and S. Wessely, *Iatrogenic factors and chronic pain*. Psychosom Med, 1997. 59(6): p. 597-604.
- [46] Dixon-Woods, M. and S. Critchley, *Medical and lay views of irritable bowel syndrome*. Fam Pract, 2000. 17(2): p. 108-13.
- [47] Sullivan, M., Between first person and third person accounts of pain in clinical medicine, in Pain 1999 -an updated review: refresher course syllabus, M. Max, Editor. 1999, IASP Press: Seattle.
- [48] Moffett, J.A., E. Newbronner, G. Waddell, K. Croucher & S. Spear, *Public perceptions about low back pain: a gap between expectations and reality?* Health Expectations, 2000. 3: p. 161-168.
- [49] Williamson, C., *Reflections on health care consumerism: insights from feminism.* Health Expect, 1999. 2(3): p. 150-158.
- [50] Hunt, L.M., B. Jordan, and S. Irwin, Views of what's wrong: diagnosis and patients' concepts of illness. Soc Sci Med, 1989. 28(9): p. 945-56.
- [51] Jensen, M.P., et al., Further evaluation of the pain stages of change questionnaire: is the transtheoretical model of change useful for patients with chronic pain? [In Process Citation]. Pain, 2000. 86(3): p. 255-64.
- [52] Kerns, R.D., et al., *Readiness to adopt a self-management approach to chronic pain: the Pain Stages of Change Questionnaire (PSOCQ).* Pain, 1997. 72(1-2): p. 227-34.
- [53] Bailis, D.S., et al., *Perceived control in relation to socioeconomic and behavioral resources for health.* Soc Sci Med, 2001. 52(11): p. 1661-76.
- [54] Charles, C., A. Gafni, and T. Whelan, Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med, 1997. 44(5): p. 681-92.
- [55] Jensen, M.P., J.A. Turner, and J.M. Romano, *Self-efficacy and outcome expectancies:* relationship to chronic pain coping strategies and adjustment. Pain, 1991. 44(3): p. 263-9.
- [56] Morris, D., Sociocultural and religious meanings of pain, in Psychosocial factors in pain: critical perspectives, R.J. Gatchel and D.C. Turk, Editors. 1999, Guilford Press: New York. p. 118-131.
- [57] Mori, I. *Chronic pain and the doctor-client relationship: an action research project.* 1997 [cited 2000; Available from: www.sfu.ca/~imori/writeup6.htm.
- [58] Neuberger, J., Primary care: core values. Patients' priorities. BMJ, 1998. 317(7153): p. 260-2.

- [59] Shiell, A., P. Hawe, and J. Seymour, *Values and Preferences are not necessarily the same*. Health Economics, 1997. 6: p. 515-518.
- [60] Wright, J.G., S. Rudicel, and A.R. Feinstein, *Ask patients what they want. Evaluation of individual complaints before total hip replacement.* J Bone Joint Surg Br, 1994. 76(2): p. 229-34.
- [61] Cassileth, B.R., *The social implications of questionable cancer therapies*. Cancer, 1989.
   63(7): p. 1247-50.
- [62] Mitchell, D.P., Postmodernism, health and illness. J Adv Nurs, 1996. 23(1): p. 201-5.
- [63] Siahpush, M., Postmodern values, dissatisfaction with conventional medicine and popularity of alternative therapies. Journal of Sociology, 1998. 34(1): p. 58-70.
- [64] Morris, D.B., *Illness and culture in the postmodern age*. 1998, London: University of California Press.
- [65] Raithatha, N., Medicine, postmodernism, and the end of certainty. Postmodern philosophy offers a more appropriate system for medicine. BMJ, 1997. 314(7086): p. 1044.
- [66] Cilliers, P., Complexity and Postmodernism. 1998, London: Routledge.
- [67] Brown, C., *Treatment for patients with chronic pain: therapists' beliefs*. British Journal of Therapy and Rehabilitation, 2003. 10(2): p. 46-52.
- [68] Brown, C., Consciousness, complexity and chronic pain: Exploring the occurrence and implications of incongruent beliefs about 'important' chronic pain treatment components (unpublished PhD Thesis), in Primary Care. 2004, University of Liverpool: Liverpool. p. 419.
- [69] Brown, C.A., The beliefs of people with chronic pain in relation to 'important' treatment components. Eur J Pain, 2004. 8(4): p. 325-33.
- [70] Brown, C., Service providers' perception of affective influences on decision-making about treatments for chronic pain. Disabil Rehabil, 2004. 26(1): p. 16-20.
- [71] Brown, C.A., Service users' and occupational therapists' beliefs about effective treatments for chronic pain: a meeting of the minds or the great divide? Disabil Rehabil, 2003. 25(19): p. 1115-25.
- [72] McGavock, H., *A Review of the Literature on Drug Adherence*. 1996, The Royal Pharmaceutical Society of Great Britain: London.
- [73] Walker, A. and J. Selfe, *The Delphi Method: a useful tool for the allied health researcher*. British Journal of Therapy and Rehabilitation, 1996. 3(12): p. 677-681.
- [74] Hasson, F., S. Keeney, and H. McKenna, *Research guidelines for the Delphi survey technique*. J Adv Nurs, 2000. 32(4): p. 1008-15.
- [75] Jensen, M.P., et al., Coping with chronic pain: a critical review of the literature. Pain, 1991. 47(3): p. 249-83.
- [76] Lackner, J.M. and A.M. Carosella, *The relative influence of perceived pain control, anxiety, and functional self efficacy on spinal function among patients with chronic low back pain.* Spine, 1999. 24(21): p. 2254-60; discussion 2260-1.
- [77] Robinson, C.A., *Managing life with a chronic condition: the story of normalization*. Qual Health Res, 1993. 3(1): p. 6-28.
- [78] Elwyn, G., A. Edwards, and P. Kinnersley, *Shared decision-making in primary care: the neglected second half of the consultation.* Br J Gen Pract, 1999. 49(443): p. 477-82.

- [79] Klaber Moffett, J.A., et al., *Public perceptions about low back pain and its management: a gap between expectations and reality?* Health Expect, 2000. 3(3): p. 161-168.
- [80] Rogers, W.A., Whose autonomy? Which choice? A study of GPs' attitudes towards patient autonomy in the management of low back pain. Fam Pract, 2002. 19(2): p. 140-5.
- [81] McIntosh, A. and C.F. Shaw, Barriers to patient information provision in primary care: patients' and general practitioners' experiences and expectations of information for low back pain. Health Expect, 2003. 6(1): p. 19-29.
- [82] Bodenheimer, T., E.H. Wagner, and K. Grumbach, *Improving primary care for patients with chronic illness*. JAMA, 2002. 288(14): p. 1775-9.
- [83] Farber, N.J., D.H. Novack, and M.K. O'Brien, *Love, boundaries, and the patient-physician relationship.* Arch Intern Med, 1997. 157(20): p. 2291-4.
- [84] Jones, R., Schism and heresy in the development of orthodox medicine: the threat to medical hegemony. Social Science and Medicine, 2004. 58: p. 703-712.
- [85] Levinson, W., et al., *Physician frustration in communicating with patients*. Med Care, 1993. 31(4): p. 285-95.
- [86] Caldron, J. and R. Beltran, *Pitfalls in health communication: healthcare policy, institution, structure and process.* Medscape General Medicine, 2004. 6(1).
- [87] Tait, R.C. and J.T. Chibnall, *Physician judgments of chronic pain patients*. Soc Sci Med, 1997. 45(8): p. 1199-205.
- [88] Butler, L. and N.E. Foster, Back pain online: a cross-sectional survey of the quality of web-based information on low back pain. Spine, 2003. 28(4): p. 395-401.
- [89] Willis, J., *Friends in Low Places*. 2001, London: Radcliffe Medical Press.
- [90] Willis, J., *Paradox of Progress*. 1995, Oxford: Radcliffee Medical Press.
- [91] Simpson, J. *Quoting Michael Ramsey, (then Bishop of Canterbury) #4421.* 1988 [cited 2004; Available from: www.bartleby.com/63/21/4421/html.
- [92] Zinn, W.M., Doctors have feelings too. JAMA, 1988. 259(22): p. 3296-8.
- [93] Bellini, L.M. and J.A. Shea, *Mood change and empathy decline persist during three years of internal medicine training*. Acad Med, 2005. 80(2): p. 164-7.
- [94] Guba, E. and Y.S. Lincoln, Competing paradigms in qualitative research., in Handbook of Qualitative Research, N.K. Denzin and Y.S. Lincoln, Editors. 1994, Sage: London. p. 105-117.
- [95] Hopayian, K., Why medicine still needs a scientific foundation. British Journal of General Practice, 2004. 54(502): p. 400-403.
- [96] NHS. NHS Direct. 2004 [cited 2004; Available from: http://www.nhsdirect.nhs.uk/.
- [97] NHS. *Patient.UK.* 2004 [cited 2004; Available from: http://www.patient.co.uk/nhs-info.htm.
- [98] Geyer, R., *Beyond the Third Way: the science of complexity and the politics of choice.* British Journal of Politics and International Relations, 2003. 5(2): p. 237-257.

Chapter 8

## THE PSYCHOLOGY OF USING AND CREATING VIDEO DECISION AIDS FOR ADVANCE CARE PLANNING\*

Muriel R. Gillick<sup>1,†</sup> and Angelo E. Volandes<sup>2,‡</sup>

<sup>1</sup> Department of Ambulatory Care and Prevention, Harvard Medical School/Harvard Pilgrim and Harvard Vanguard Medical Associates,

Boston, MA 02215, USA

<sup>2</sup> General Medicine Unit, Massachusetts General Hospital and Instructor in Medicine, Harvard Medical School, Boston, MA 02215, USA

## ABSTRACT

Decision-making in the medical context has traditionally focused on individual patients choosing between various medical treatments that have been proposed for an acute medical problem. For example, a pregnant woman may need to decide whether or not to undergo amniocentesis to rule out Down's syndrome in her fetus, or a man with an elevated prostate specific antigen test may need to choose between "watchful waiting" or radical prostate surgery. Much of this research has focused on how patients process information and how physicians can objectively present such information to improve decision-making.

Advance care planning is the process in which patients plan for their future medical care. Patients asked to consider the benefits and burdens of various interventions that might be considered for some hypothetical disease state, such as dementia. Advance care planning is advocated largely in order ensure that the treatment the patient receives is

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Medicine and Health Care*, edited by Thomas E. Lynch published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: Department of Ambulatory Care and Prevention, Harvard Medical School/Harvard Pilgrim and Harvard Vanguard Medical Associates, 133 Brookline Avenue, 6<sup>th</sup> floor; Boston, MA 02215 Telephone: 617-509-9977 Email: mgillick@partners.org

<sup>&</sup>lt;sup>‡</sup> Correspondence to: General Medicine Unit, Massachusetts General Hospital and Instructor in Medicine, Harvard Medical School, Boston, MA, 50 Staniford Street, 9<sup>th</sup> Floor, Room 933; Boston, MA 02114, Telephone: 617-643-4266 Email: avolandes@partners.org

consistent with his own preferences—even if he is incapable of making decisions at the time of illness.

While advocated on theoretical grounds, advance care planning has not been widely adopted, either by patients or physicians. A major barrier to engaging in advance care planning is that patients are often unfamiliar both with the hypothetical disease state and with the specific medical interventions that might be used to treat them. Video has previously been used as a decision aid for patients facing real time decisions about their medical care. The application of video to advance care planning discussions is a new approach that has the potential to help surmount the barrier to informed decision making.

Video is a powerful and widely accessible medium that engages patients in a way not easily captured with words. Video allows patients to imagine hypothetical health states and can provide an accurate portrayal of medical interventions. This chapter will review the emerging corpus of work that uses video as an adjunct to advance care planning, specifically in planning for the possible development of dementia. It will analyze the biases that can arise from the use of an aesthetic and visual medium and discuss ways those biases can be avoided. The chapter will make the case that video is useful for patients with low health literacy and patients from minority populations. Finally, other specific disease entities that lend themselves to video portraits will be proposed.

**Keywords**: Advance Care Planning, Decision Making, Video Decision Aids, Patient-Physician Communication, Dementia, Health Literacy

## INTRODUCTION

Patients must increasingly make complex decisions about their medical care.[1] This new role for patients is rooted in the primacy accorded patient autonomy: the right of the individual to determine what is done to his body.[2] It also reflects the recognition that the best approach to treatment requires a balancing of its burdens and benefits, a highly subjective process that reflects individual values and preferences. Research in decision making in medicine has typically focused on how individual patients engage in this balancing process for specific medical interventions, as well as on individual attitudes to risk.[3] Much of this work has led to the development of decision aids that attempt to facilitate medical decision making by patients and to improve the way physicians present the information that patients need in their new role.

Many of the decisions patients are called upon to make arise at a time when the patient is too confused or too overwhelmed by acute illness to be able to participate in a complex process that may have life or death consequences. To attempt to preserve patient autonomy in these situations, physicians and ethicists have developed advance care planning (ACP). ACP involves a patient making choices while cognitively intact about medical interventions that might be proposed at some future date when he is no longer decision-capable. Sometimes these hypothetical interventions represent potential treatment for a condition the patient already has: for example, a patient with emphysema may decide while cognitively intact whether he would want to be put on a ventilator should his breathing deteriorate, with concomitant impairment of his ability to think and reason. Sometimes the hypothetical intervention is one that might be used in the event the patient develops a condition he does not currently have—but in which he would, by definition, be unable to make decisions. For example, a healthy person might be asked whether he would want a feeding tube in the event he developed advanced dementia—a condition he does not currently have—and difficulty swallowing—a common complication of dementia. ACP introduces a new level of complexity compared to decisions between alternative treatments available for an existing condition, such as choosing between prostate surgery or watchful waiting for localized prostate cancer.

As if thinking about a medical intervention that is not currently indicated for a medical problem the patient does not actually have were not sufficiently abstract, ACP also asks patients to imagine the consequences of having, or not having, the treatment. For example, healthy people contemplating the pros and cons of attempted cardiopulmonary resuscitation (CPR) if they had a condition such as a stroke must consider not only what the experience of undergoing CPR would be like, but also what their lives would be like afterwards. In light of the complex nature of ACP, some thoughtful physicians, ethicists, and legislatures have advocated that patients simply select a surrogate to make decisions on their behalf if they cannot directly participate in discussions of their medical care.[4] But given extensive evidence that surrogates have difficulty determining what interventions a patient would want,[5] and given the move towards more specificity rather than less in living wills in the wake of the Terri Schiavo case, improved tools to facilitate ACP are in demand.[6]

Video decision aids are a recent innovation that have the potential to overcome the formidable barriers to informed decision-making.[7] Video is a powerful and widely accessible medium that engages patients in a way that verbal descriptions, whether written or oral, do not. Video allows patients to vividly imagine hypothetical health states while simultaneously providing an accurate portrayal of medical interventions, and it can illustrate how the intervention can affect the patient's future.

In this chapter we review the emerging corpus of work that uses video in decision making and the development of decision aids for ACP. We will specifically focus on video for ACP that requires planning for hypothetical interventions in hypothetical disease states, as this is the most problematic form of ACP. The disease state dementia and interventions often utilized in patients with dementia offer particularly rich examples. We will explore the additional benefit of video decision aids for patients with low health literacy as well as patients from minority populations. Finally, this chapter will discuss the potential biases inherent in the use of this aesthetic and visual medium and how it may influence decision making.

Video decision aids will increasingly become part of the decision aid landscape as medicine adapts its communication tools for an image-saturated society that expects its citizens to tackle complex information and decisions. The advent of internet-based vehicles for transmitting video will also aid in the dissemination of innovative video decision aids. Video is a promising means for communicating information that supplements written materials and physician-led discussions. It offers the potential to significantly improve the advance care planning process.

## Previous Uses of Video Decision Aids

The use of video in the development of decision aids is not new; a small number of decision aids have already used video technology. The Cochrane database, the most comprehensive and extensive list of decision aids available, includes 12 decision aids that use

video as part of their decision making process.[8] In addition, a number of studies published in the medical literature evaluate the use of video decision aids. Such video decision aids are used in a variety of clinical settings, including prostate surgery, back surgery, ischemic heart disease, and childhood vaccination. Many of these decision aids have been developed and made widely available by health organizations such as the Foundation for Informed Decision Making (see www.fimdm.org).

The primary goal of these video decision aids is to present information necessary for informed decision making in an unbiased and structured format. A common technique used in many of these decision aids is to present information about the intervention under consideration by using filmed interviews. These interviews provide information about the disease along with the risks and benefits of the proposed intervention. These filmed narratives are then supplemented by computer graphics that illustrate the numerical risks and benefits of the various interventions available. Such graphics can intuitively capture complex information such as the survival and life expectancy associated with various interventions.

Some decision aids have also incorporated more creative applications of video technology, including filmed interviews of patients who have undergone the medical intervention and provide testimonials. Such testimonials range from describing the actual medical intervention to offering a first-hand look at some of the possible side effects or adverse outcomes of medical interventions. Additional types of testimonials include insights from real patients into their decision making and a discussion of the satisfaction or regrets that resulted from their deliberations. Testimonials from clinicians and family members are also included in some of these video decision aids.

In general, decision aids have been well-received by patients. Numerous studies have demonstrated increased satisfaction and decreased uncertainty among individuals who used a decision aid compared to individuals receiving usual care.[9] In particular, studies have shown an increase in knowledge about the benefits and burdens of medical interventions when video decision aids supplement conventional sources of information.[10]

It is not surprising that video decision aids improve decision making to a greater extent than non-video decision aids. Consumers increasingly obtain information from television and the internet, which rely heavily on visual images. The extraordinary success of YouTube is a recent example of the widespread appeal of video. Furthermore, complex information about illness and medical interventions is more easily transmitted by a visual source than with print material, especially in people with limited health literacy or for whom English is a second language. Additionally, video images provide a context for individual decision making that is not readily communicated via print materials: they provide emotional content (facial expressions, existential distress, anxiety, etc.) and experiential insights (personal reports of negative outcomes, regrets, etc.), and they help establish trust (camaraderie with patients who have faced similar decisions, first person narratives as opposed to third person accounts, etc.). How video decision aids attempt to do this and the validity of this approach to decision making is part of an exciting and emerging corpus of research. In the remainder of this chapter, we explore this innovative line of decision making research in ACP, focusing particularly on planning for the possibility of dementia.

## Advance Care Planning in Dementia

Advance care planning is a process of communication among physicians and patients which aims to ensure that treatment plans are aligned with patients' goals and needs. ACP is a process which allows patients to remain active in their medical care even if they develop impaired decision making capacity.[11] ACP typically involves a structured discussion by the physician with patients and their families about hypothetical health states such as dementia and explores individual preferences for care.[12] Central to the process of ACP is the patient's ability to imagine and understand hypothetical health states. Additionally, patients need to have a realistic understanding of the burdens and benefits of potential medical interventions, such as cardiopulmonary resuscitation and mechanical ventilation, which may be offered as treatment of acute illness. Imagining and understanding hypothetical health states and potential medical interventions have proven to be difficult. Video is a medium that can enable clinicians to surmount both of these barriers to successful ACP.

Dementia is a common disorder, with the most common variety, Alzheimer's disease, affecting an estimated 5 million Americans.[13] The prevalence of dementia rises with age so that only 2% of people age 65-74 have dementia, compared to 42% people over the age of 85. Unfortunately, the public's understanding of the disease often reflects the portrayal of the disease in the media and lay press, which often present a tempered and palatable picture of dementia. For instance, the popular book *Elegy for Iris* and its movie-adaptation paint a romanticized and over-intellectualized portrait of a writer suffering from Alzheimer's disease. The advanced stages of the disease, which are often the most disquieting and disturbing stages to both clinicians and families, are recreated as sequences of dreamy reveries of previous life events that come to poetic closure. Although artistically satisfying, these portraits misrepresent the realities of advanced dementia, which include the inability to communicate, disorganized thoughts, and incontinence. Unfortunately, impressions of dementia originating from such portrayals are more common than from clinically accurate sources. Examples of films that provide a more accurate depiction of dementia, such as the PBS documentary *The Forgetting*, are rare.

A similar argument applies to medical interventions, such as CPR and mechanical ventilation, which are commonly part of ACP. Popular understanding of the burdens, benefits, and success rates of these procedures are more often derived from television shows than from the clinical reality; moreover, the success rates for such procedures on television are dramatically inflated.[14] It is no surprise that when surveyed, patients have an unrealistic expectation of the success rate for aggressive medical interventions. Such misconceptions and misinformation pose serious barriers to accurate and informed ACP.

## Using Videos in Advance Care Planning

A new generation of video decision aids has evolved over the last few years. These instruments are distinct from the previous generation of video decision aids in that they use video in a different conceptual role. Earlier prototypes of video decision aids focused on risks and benefits of medical interventions; the newer generation of video decision aids attempt to portray the illness and the actual medical interventions. Often using real patients and actual medical interventions, this newer generation of decision aids aims to supplement hypothetical

descriptions with filmed sequences of real patients showing actual medical interventions. These newer decision aids aim to provide a more concrete picture for patients deliberating about medical decisions, adding a sense of verisimilitude that may be lacking in filmed interviews. Using the newer generation of video decision aids, we studied the use of video in ACP for dementia. The video decision aid developed for the study depicts the salient features of advanced dementia, including the inability to communicate, the inability to ambulate, and lack of capacity to feed oneself. The video depicts an elderly female patient in the advanced stages of dementia with her two daughters in a nursing home setting. The patient fails to respond to their attempts at conversation. Next, the daughters transport the patient in a wheelchair. Finally, the daughters hand feed their mother pureed food. Along with the visual images is a voice over narration that describes the advanced stages of the disease using simple language. (The film clip is available online at: http://homepage.mac.com/avolandes/AlzheimersVideo/iMovieTheater18.html).

The video was filmed and edited with close collaboration from experts in geriatrics, neurology, and medical ethics. In attempting to recreate the experience of having dementia, the video eschewed visual effects or staged directions, favoring a cinema verite style of documentary film-making.[15] The video decision aid also used different lens angles and did not zoom-in on the face of the patient until the viewer became accustomed to the clinical scenario being portrayed. The opening shots introduce the patient's disease: the visual is a wide-angle shot that presents the patient in a nursing home setting. Such careful filming seeks to avoid visceral reactions to what is a disturbing hypothetical health state for most healthy patients.

The primary aim of the study was to determine whether enhancing the understanding of the underlying health state—dementia—with video would affect the choices made by prospective patients about the level of care they would prefer. Since the video was used to depict dementia, the decision was made to ask subjects about the overall approach to medical treatment they would favor in the event they developed dementia, rather than asking about particular medical interventions. An intervention-specific approach to ACP would require video representation of each of the interventions, for example mechanical ventilation, cardiopulmonary resuscitation, and dialysis. We opted to offer subjects a choice of three levels of care: life-prolonging care, comfort-oriented care, and an intermediate level between these two extremes.[16]

To isolate the effect of a video portrayal of dementia on decision-making, subjects were initially given a verbal description of advanced dementia, based on the Functional Assessment Staging (FAST) in Alzheimer's disease criteria [17] and asked which of the three levels of care they would select if they were in such a state. They were then shown the video, following which they were again asked to select a level of care. The study measured changes in preferences after seeing the video.

The multi-site study included 120 subjects, half of whom were from minority populations, and all of whom were over the age of 40. Before watching the video, half of the subjects preferred life-prolonging treatment and half preferred comfort care. After watching the video decision aid, close to 90% of subjects chose comfort care, a highly statistically significant finding. Almost all of the subjects who initially chose a level of care other than comfort changed their preferences to comfort care after the video. Furthermore, when subjects were asked if they felt their understanding of the disease had improved after seeing the video, more than 98% of subjects found the video decision aid "very helpful" or "somewhat

helpful." Almost all of the subjects thought that using video decision aids in other diseases such as cancer would be either "very helpful" or "somewhat helpful."

Interestingly, the video decision aid had the greatest effect in changing the preferences of Latino and African-American subjects and in those with a high school education or less. In a subsequent study that looked at the use of the same video decision aid in patients from diverse educational backgrounds, the aid was found to be most helpful in those patients with limited education and low health literacy. These observations corroborate the findings of previous research that visual images communicate complex information better than verbal or print materials, especially among patients with limited education and health literacy.[18]

The distinctive feature of this decision aid was the careful depiction of dementia by filming an elderly white woman interacting with her daughters. The voice over narrative reinforced the visual images by describing the phenomena represented in the video and corresponding to FAST stage 7: for example, inability to ambulate and inability to communicate with more than 3 words. Rather than focusing on the burdens and benefits of specific medical interventions, as was customary with the previous generation of video decision aids, our video sought to overcome what is perhaps the major barrier to ACP, the ability of the individual to imagine experiencing an illness he does not currently have. Video can supplement verbal descriptions with illustrations of real people in real circumstances, enabling individuals to better visualize a possible future condition such as dementia.

Other studies are presently underway that use a similar conceptual role for video in decision aids but which seek to depict the medical interventions that might be offered in a particular disease state. For instance, CPR and mechanical ventilation are depicted with videos of the actual interventions: clips of an intensive care unit are used to explain lifeprolonging care and clips of home hospice illustrate comfort care. The goal of these decision aids is to help individuals engaged in ACP grasp both the nature of the intervention and the experience of undergoing the procedure. Once again, a voice-over narrative reinforces the visual images by describing in simple language the interventions, including their risks and benefits. How these decision aids influence the way patients make decisions is a vital question that we address below.

## The Psychology of Decision Making with Video Decision Aids

Sound decision making is predicated on knowledge. In the medical context, this has traditionally been interpreted to mean knowledge of the risks and benefits of whatever test or treatment is under consideration, as well as an accurate understanding of the disease for which the test or treatment is proposed. For instance, a patient diagnosed with prostate cancer is asked to choose between surgery or watchful waiting. To make such a choice, he must understand the risks of surgery (impotence, incontinence, diarrhea) as well as the benefits (an increased chance of cure, peace of mind, etc.). He must also understand the natural history of prostate cancer—the way the disease might progress (for example with metastases to the bones) and the likelihood of such a progression. This type of knowledge can be readily acquired through filmed patient interviews in which patients who have undergone treatment discuss their experience. With ACP, the knowledge necessary for sound decision making is more abstract and elusive. ACP requires healthy people to imagine devastating diseases such

as dementia as well as the nature of the suffering experienced by patients afflicted with such conditions.

Individuals who have no personal experience with a condition such as dementia often assume it entails slight forgetfulness or trouble expressing oneself. They often do not appreciate the extent of the cognitive impairment or the physical dependence characteristic of advanced disease. Since afflicted individuals are not able to express their feelings, they cannot directly provide a report of their experience. Verbal descriptions often fall short of describing such a tragic condition and even the most gifted of communicators may not be able to give patients a realistic image of the advanced stages of the disease.

The following quotations from the participants in our study of video aids for ACP in dementia offer a glimpse into the power of video to help imagine the unimaginable:

You explained what dementia is and I heard you, but I didn't understand you. After the video, I now do.

You know truly what is happening, reality is not left to your imagination. You see it.

Actually seeing the disease is better than verbalizing it.

It shows you the reality [of the disease].

I had a very different conception of what dementia was. This [the video depiction of dementia] was totally different.

Yes, I had my own idea what this [dementia] was. After seeing it [dementia], it was very different.

It gives people a visual concept of what to expect if they don't have experience of having the disease. It's hard to see the video but it's reality and helpful.

Many participants in our study acknowledged that their internal image of dementia and the image depicted in the video were vastly different.

Another common theme articulated by individuals after seeing the video was that it facilitated identification with the subject in the video. They were better able to imagine what it would be like to have advanced dementia by actually seeing a patient with dementia. For instance:

I can now put myself in her position. Seeing her is very helpful.

You would know how it's like to have the disease.

You get to see real patients with the disease.

Video makes it concrete and easy to connect to.

Patients engaged in ACP put a high value on the ability to project oneself into the position of someone with dementia and to empathize with such a person. Empathy, a common theme in the ACP and palliative care literature, played an important role in patients' decision making.

Qualitative data also highlighted the theme of suffering. The video conveyed human suffering in a manner simply not captured with words. The ability to grasp the patient's suffering played an almost universal role in patients' deliberations:

You see the person, the one who is suffering.

Suffering is a common concern for patients, particularly at the end of life. The experience of suffering is difficult to encapsulate verbally and perhaps this represents the strongest aspect of videos: their ability to transmit information about suffering visually rather than verbally.

Whether it is a grimace or a groan, such information is rarely communicated with print or verbal descriptions.

Other issues arising in the care of demented patients or other individuals near the end of life may also be better appreciated through visual rather than verbal communication. Many subjects in our study expressed this sentiment:

Visual is much better than talking. People don't understand what you're saying. I'm a picture person.

The written and spoken word by the doctor, for me conjures images of disease unique to me. But when you see a real person, you physically see the disease and it has an entirely different meaning. It's visual, not some scientific conclusion.

I'm a better learner visually.

The video decision aid overcame some of the barriers to good ACP by helping individuals to imagine the disease state and by reifying suffering. Other decision aids exploring diseases such as stroke or the persistent vegetative state may benefit from the use of video. Likewise, video can be used to illustrate medical interventions that might be used in these states. Although this new generation of decision aids is very promising, they are not without their own unique challenges, which we will review in the next section.

## Procedures and Challenges to Creating Video Decision Aids

Using video in decision aids raises concerns about bias and accuracy, some of which arise with print and audio aids but others of which are unique to the medium of video. The two major challenges that we will address for video decision aids can be broadly categorized as: content-oriented concerns and film-making issues.

## Content-Oriented Concerns

One of the most crucial choices in the creation of any decision aid is what content ought to be included. Often, the amount of potentially relevant medical information about a disease or intervention is overwhelming. A useful decision aid must incorporate just enough information and data to allow a patient deliberating about a medical decision to grasp the issues and draw a reasoned conclusion. For video decision aids that attempt to capture the experience of having a disease or undergoing a medical intervention, the choice of what to include is subjective. What is included and what is omitted shape the impression of what the disease or intervention entails. Thus, judgment about what ought to be included in a video decision aid is vital to the accurate and objective portrayal of the disease and intervention.

Building on a scientifically accepted characterization of the entity being portrayed is a good starting point. For instance, in our video of advanced dementia, we enumerated the FAST criteria for this diagnosis and depicted each of these criteria. As a result, the content of the film reflected accepted standards for what constitutes advanced dementia. The Foundation for Informed Medical Decision Making's procedures for the development of decision aids provides additional suggestions for how to determine content, procedures that apply equally to video decision aids and other instruments aids (see www.fimdm.org), as well as by others.[19]

Focus groups consisting of patients, family members, caretakers and clinicians should be used early on when attempting to create a video decision aid that attempts to portray the reality of a disease like dementia or to depict medical interventions. In the case of a cognitive disorder like dementia, interviewing patients with the disease is not possible. Therefore, family members, caregivers and clinicians are the major sources for a first-hand account of the disease experience: they can describe what it is like to live with the disease and what caring for a person with this disease entails. Family members should be encouraged to elaborate on their experience of taking care of patients with diseases like dementia.

Surveying clinical experts from a variety of disciplines is also key to ensuring the inclusion of diverse perspectives. When we created our decision aid for advanced dementia, we solicited the input of geriatricians, neurologists and medical ethicists. Whenever possible, such focus groups should be filmed so as to allow including parts of the interviews in the final decision aid. The same approach can be applied to videos that portray medical interventions such as mechanical ventilation. Interviewing patients who have had the procedure, along with their caregivers and clinicians, and developing a narrative describing the experience are the first steps in creating the video decision aid.

After reviewing the content generated from the focus groups and following a review of the relevant medical literature, a list of the salient aspects of the experience of the disease and interventions should be created. At this point, a "sketching out" of the video sequences should be done, guided by the technical expertise of the film-makers and the production team. Once a rough draft of the film sequences is complete, an iterative process among the production team, patients, family members, caregivers and clinicians should take place. Close collaboration among all of these groups should occur throughout both the decision aid production and filming processes.

## Film-Making Challenges

Presenting objective and unbiased information is one of the most common expressed goals of medical decision making. Exactly what the criteria are for "objective" and "unbiased" information is part of a lively and ongoing debate in the medical decision making literature. Decision aids ought to be tools that provide information but should remain agnostic and impartial to a "right" decision when numerous reasonable choices exist. Achieving objectivity is complicated with video decision aids since video is inherently an aesthetic medium. Such media necessarily entails numerous prejudices and biases. Although not exhaustive, the VIDEOS criteria detailed below based on our own experience in creating and evaluating a video decision aid for ACP can serve as a broad outline to creating and evaluating other video decision aids.

## VIDEOS Criteria

- V Views of various caregivers (patient, family, clinician)
- I Include a variety of film angles with care given to the viewer's experience
- D Diversity of subjects (gender, race) and settings (inpatient, home)
- E Explain and reinforce images with verbal narratives
- O Other videos of the same subject matter should be compared
- S -Staged directions should be avoided in favor of real patients and scenes

## Views of Various Caregivers

Attempting to recreate the experience of a disease requires a large amount of research before beginning the film-making process. Interviewing and discussing the illness experiences of patients and their various caregivers (family members, nursing aides, and clinicians) will allow for the most complete picture of the experience of a disease.

#### Inclusion of Various Film Angles

Viewing a video of a tragic disease or intervention can be a jarring experience. Careful attention should be paid to the use of close-up (zoom) lens angles in the early segments of a video decision aid since the viewer is likely becoming accustomed to the illness experience. Wide-angle shots introducing the disease and setting are preferred in the opening scenes of a video decision aid. Attention to film angles will ensure that the informational component of the visual images is not muted by images that some viewers may find disturbing.

## Diversity of Subjects and Settings

How patients relate to other patients may be influenced by qualities extraneous to the disease, such as gender, race, ethnicity and age. Whenever possible, depicting patients from diverse backgrounds will help overcome this concern. Additionally, several non-clinical settings should be included in the video if the patient typically experiences his illness in the home, the nursing home, or other environment outside the hospital.

## Explain and Reinforce Visual Images

Voice-over narratives can often reinforce the informational content of the visual image. All narrations should use simple language at the junior high school language. Repetition of information through images and words will help viewers absorb information.

## Other Videos for Comparison

Other videos depicting the same illness experience—preferably created by different filmmakers—should be compared to the video decision aid that is being designed. Surveying patients about their viewing experiences after watching various videos of the same disease will assist in the final editing process. Furthermore, concordance of decision making using different video aids is re-assuring although not definitive.

## Staged Directions Eschewed

Documentary film-makers traditionally advocate avoiding staged directions in order to capture objective experiences. This cinema verite style of film-making is not without its difficulties, but is the preferred style of film-making for video decision aids that purport to recreate the illness experience.[16] Staged directions and actors should be eschewed in favor of real patients and real medical circumstances.

#### Future Research Areas

In the coming years, physicians and patients will increasingly use video decision aids and many will be available on the Internet. Three promising areas of future development of video decision aids for ACP include: aids for patients in the early stages of dementia; aids for minority populations; and aids for patients with limited health literacy. Research in decision making has traditionally focused on the decision making process of cognitively intact individuals, with little emphasis on the needs of cognitively impaired individuals, such as patients with mild dementia. The coming years will see an explosion in the need for tools to assist patients who are cognitively disabled – but still have the capacity to make decisions. The number of patients with Alzheimer's dementia alone is expected to reach epidemic proportions in the coming years, with 1 out of 8 patients aged 65 or over in the U.S. suffering from the disease.[20] ACP has the potential to protect families from some of the emotional burden of medical decision making for loved ones in the advanced states of the disease and can help individuals retain control over their bodies as they decline. Exploring how patients with cognitive disabilities process and understand information through a visual medium as opposed to print or verbal descriptions will be a promising avenue of research.

The decision making literature is paying increasing attention to the large numbers of patients in the health care system from minority populations, including Latinos, African-Americans, and Asian-Americans. Conventional decision aids are often problematic in these groups because they fail to address the specific concerns of patients from varying ethnic backgrounds. For example, minority populations have repeatedly been shown to reject hospice care at the end of life, in part due to lack of trust of the medical establishment.[21].An additional advantage of a video decision aid is the opportunity to portray patients and clinicians from particular backgrounds using culturally appropriate language and a culturally specific milieu.

Finally, further research is needed to better address the needs of some of the most vulnerable patients, those with limited health literacy. Over 90 million patients in the US health care system are estimated to have a literacy level at the 9<sup>th</sup> grade or below.[22] ACP with traditional tools that depend heavily on verbal skills may be especially difficult in these patients. Recent studies have suggested that video decision aids in the context of end-of-life decision making may be effective in disseminating information vital for informed decision making to patients with limited health literacy.

## CONCLUSION

Advance care planning requires patients to think about conditions they hope they will never experience: states such as disability, cognitive impairment, coma and death. It demands that patients—or rather potential patients—imagine themselves as dependent, in pain, suffering, dying, or some combination of all these distressing aspects of the human condition. Many people prefer not to think about such eventualities, asserting that they will make decisions about their medical care "when the time comes." But that point may be too late as patients are often confused, unconscious, or simply emotionally overwhelmed at the time critical medical decisions must be made. To increase the chance of determining what happens to them at important junctures in their lives, or to help the friends and family members who will take the responsibility for making decisions on their behalf, they can choose to engage in ACP.

Despite over twenty years devoted to creating living will and other advance directives, the process of ACP remains highly problematic. Although surveys repeatedly confirm that patients say they want to discuss difficult issues, both they and their physicians have difficulty bringing up the subject. Nationwide, at best 25% of older adults have completed some kind of advance directive.[23] When patients do participate in ACP, it is far from clear that their choices reflect truly informed decision making. We know that patients are often influenced by recent experiences: their 'preferences' may in fact reflect exposure to a single friend or relative who underwent the procedure in question. We also know that patients often use magical thinking: they believe that however unlikely a procedure is to be effective, it will work when applied to their particular case.[24] Video-based ACP offers the potential to overcome these psychological biases. It can help patients identify with others who have had a given condition or procedure; it can particularize the intervention, making it real in a way that verbal or written descriptions seldom do. The incorporation of video into decision aids may enable ACP to finally fill its initial promise.

## REFERENCES

- [1] Barry MJ. Health decision aids to facilitate shared decision making in office practice. *Ann Intern Med* 2002,136:127-135.
- [2] United States. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. (1983) *Deciding to forego life-sustaining treatment : a report on the ethical, medical, and legal issues in treatment decisions.* Washington, DC: President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. Washington, D.C.: U.S. G.P.O.
- [3] Weinstein MC. (1980) *Clinical decision analysis*. Philadelphia: Saunders.
- [4] Buchanan AE, Brock DW. (2004) Deciding for others : the ethics of surrogate decision making. New York: Cambridge University Press.
- [5] Shalowitz DI, Garrett-Mayer E, Wendler D. The accuracy of surrogate decision makers: a systematic review. *Arch Intern Med* 2006;166:493-497.
- [6] Hook CC, Mueller PS. The Terri Schiavo saga: the making of a tragedy and lessons learned. *Mayo Clin Proc* 2005;80:1449-1460.
- [7] Volandes A, Lehmann LS, Cook EF, Shaykevich S, Abbo ED, Gillick MR. Using video images of advanced dementia in advance care planning. *Arch Intern Med* 2007; 167:1-6.
- [8] O'Connor AM, Stacey D, Rovner D, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev.* 2001(3):CD001431.
- [9] Mitchell SL, Tetroe J, O'Connor AM. A decision aid for long-term tube feeding in cognitively impaired older persons. *J Am Geriatr Soc* 2001;49:313-316.
- [10] Morgan MW, Deber RB, Llewellyn-Thomas HA, et al. Randomized, controlled trial of an interactive video disc decision aid for patients with ischemic heart disease. J Gen Intern Med 2000; 15: 685-693.
- [11] Gillick MR. Advance care planning. N Engl J Med 2004;350:7-8.
- [12] Emanuel LL, Danis M, Pearlman RA, Singer PA. Advance care planning as a process: structuring the discussions in practice. *J Am Geriatr Soc* 1995;43:440-446.
- [13] Alzheimer's Association. *Alzheimer's Disease Facts and Figures* 2007. www.alz.org/national/documents/Report\_2007FactsandFigures.pdf

- [14] Diem SJ, Lantos JD, Tulsky JA. Cardiopulmonary resuscitation on television. Miracles and misinformation. N Engl J Med 1996;334:1578-1582.
- [15] Grant BK, Sloniowski J. (1998) Documenting the documentary : close readings of documentary film and video. Detroit: Wayne State University Press.
- [16] Molloy DW, Guyatt GH, Russo R, et al. Systematic implementation of an advance directive program in nursing homes: a randomized controlled trial. JAMA 2000;283:1437-1444.
- [17] Sclan SG, Reisberg B. Functional assessment staging (FAST) in Alzheimer's disease: reliability, validity, and ordinality. *Int Psychogeriatr* 1992;4 Suppl 1:55-69.
- [18] Houts PS, Shankar S, Klassen AC, Robinson EB. Use of pictures to facilitate nutrition education for low-income African American women. J Nutr Educ Behav 2006;38:317-318.
- [19] Elwyn G, O'Connor A, Stacey D, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ* 2006;333:417.
- [20] Hebert LE, Scherr PA, Bienias JL, Bennett DA, Evans DA. Alzheimer disease in the US population: prevalence estimates using the 2000 census. *Arch Neurol* 2003;60:1119-1122.
- [21] Greiner KA, Perera S, Ahluwalia JS. Hospice usage by minorities in the last year of life: results from the National Mortality Followback Survey. J Am Geriatr Soc 2003;51:970-978.
- [22] Nielsen-Bohlman L, Institute of Medicine Committee on Health Literacy (2004). *Health literacy : a prescription to end confusion*. Washington, DC: National Academies Press.
- [23] Lo B and Steinbrook R. Resuscitating advance directives. *Arch Intern Med* 2004; 164:1501-1506.
- [24] Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science* 1981;211:453-458.

Chapter 9

## **DECISION MAKING AND THE EMOTIONAL BRAIN**\*

José M. Martínez-Selva and Juan P. Sánchez-Navarro

School of Psychology, University of Murcia, Spain

## ABSTRACT

There is an increasing amount of empirical data which give support to the idea that affective processes play an important role when making decisions. The main purpose of this chapter is to describe the brain mechanisms that support and/or mediate the involvement of emotion in decision making processes. A brief overview is presented of the theoretical approaches explaining how emotion influences decision making, such as the prospect theory, the somatic marker hypothesis, and other recent theories. The brain mechanisms involved in decision making and the processing of emotional information are reviewed. Particular attention is given to the role of the ventromedial and dorsolateral prefrontal cortices, anterior cingulate cortex and amygdala in decision making, and to the brain defence system and the dopaminergic mesocortical-mesolimbic system, together with their projections and related structures, in emotion. Research results from clinical studies, e.g. brain injury and substance abuse, and from the field of Neuroeconomics which are relevant to the topic are also discussed.

## **1. INTRODUCTION**

We are all continually called upon to take decisions in all walks of life. It may be something quite trivial or something of the highest transcendence. Choosing from among several options may often be an easy task, but there are times when it becomes so complex that it becomes a problem in its own right. Ideally, deciding involves using all the information available, earlier experiences of similar situations, along with the knowledge we have about the behaviour and opinions of others and on the foreseeable outcomes arising from the

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Medicine and Health Care*, edited by Thomas E. Lynch, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

various options. Real life, however, is uncertain and unpredictable, and decisions are taken with incomplete data and the outcomes are frequently unknown. Many decisions are therefore taken automatically on account of the high cost of assessing the information. To a large extent, habits, social norms, moral criteria and even laws can serve to lessen the burden of continuous and at times intense analysis of information.

Decision making involves large cognitive activity, including the processing of the stimuli present in the situation, the recall of earlier experiences and the estimation of the possible outcomes arising from the various options. It may therefore be considered as a complex process which is made up of multiple sub-processes where varied information is incorporated. Different areas of the brain are implied, which, while acting together, have separable functions (Krawczyk, 2002; Tranel & Damasio, 2000).

Current research shows that decision making is very often not a mere reasoning process consisting of weighing up or comparing the losses or gains arising from a particular decision. Rather, what seems to occur is that the emotional aspects deriving from experience of similar situations, be they personal or vicarious, as well as those associated with the outcomes or with the context in which the decision is made, play an important role (Damasio, 1994). Emotions guide decision making. They simplify and speed up the process, thus reducing the complexity and easing possible conflict between similar choices. It is significant that people who do not perform well in experimental decision making tasks do not show similar emotional changes to those who perform correctly, and they present serious problems in adjusting in their social or interpersonal lives, as occurs in patients with certain injuries to the frontal cortex or in some consumers of abusive substances (Bechara, 2003; Bechara et al., 2001).

## 2. PSYCHOLOGY OF DECISION MAKING

### 2.1. Reason and Emotion

Decision making has been the subject of wide study, both in Economics and in Psychology. The *rational or expected utility theory* of decision making considers that humans are rational and therefore make weighted estimations which lead to rational decisions. The decision will be guided by estimation or rational evaluation of the *utility* or benefit that can be expected. According to this theory, people assess the information available and maximize the benefits they can expect, and consumers are, therefore, rational, economic players who select those options which offer the highest expected utility (Von Neumann & Morgenstern, 1947). Consequently, choices are based on rational assessment of the outcomes and, hence, depend on the balance between the possible losses and gains which can be anticipated in different situations, such that the cost of the decision will be equal to or lower than the expected return.

Utility is no easy concept to define and entails many subjective components. In its beginnings, *utility*, as defined by mathematicians and philosophers like Daniel Bernoulli and Jeremy Bentham, was understood as the balance between pleasure and pain, and was based on the idea that people establish their values on wealth according to the pleasure or pain this provides them. The calculation behind many economic decisions is rational, but the fact is that decisions or bets which expect a negative outcome persist, e.g. playing the lottery or taking out an insurance policy (Mellers, 2000). Many decisions are taken on the basis of

impulses, of previous experience, of what we see or believe others do. Moreover, individual differences are also very important since, for example, two people in possession of the same information may opt for different decisions. Tversky and Kahneman (1974, 1981) proposed, in this vein, the *prospect theory*, which underlines the psychological and subjective elements of decision making. The value of a *prospect* is equal to the sum of the benefits of the outcomes which may be experienced and taking into account the likelihood of their happening. The decision includes probabilities which are affected by belief, confidence or expectation of outcomes and their consequences. Unlike rational theory, not only is the expected result maximized, the *confidence* that this will occur is also taken into account. In other words, the expectation that some event, outcome or consequence will arise from the decision taken and not the real likelihood of its occurring is what is important. People do not always tend towards the rational solution to a problem but to the most convenient, to the one they like or the one which saves them the problem of having to take a decision. They will often be guided by the mood they are in.

Each person values utility subjectively and according to signs or *points of reference* that depend as much on the information available as on psychological processes. People do not have clearly ordered preferences, which often depend on the context and the type of decision. When taking a decision we evaluate the likelihood and the seriousness of the possible outcomes of the alternatives and we process this information subjectively, including biases and errors, in a calculation which is based on expectation. The prospect theory proposes (Kahneman, 2003; Shafir & LeBoeuf, 2002) that the assessment of the outcomes of a decision depends on whether these are perceived as gains or losses, and not on the total net outcome. The emotional response to the outcome of a decision depends on the perceived value and the likelihood of the outcome and also on the various alternatives (Mellers, 2000). Under conditions of certain gain, the person does not tend to take risks, while if there is a certain loss, the tendency is to take risks, even if the net outcome is the same for both cases. The function of value is more accentuated for losses than for gains. Thus, the value function is concave for gains and convex for losses, which produces an aversion to risk in the case of gains and the search for risk in the case of losses, except when the probabilities are very low. If forced to choose from among gains, the tendency is to go for the safest one, while with losses it is risk that is chosen, even though the options lead to the same outcome (Shafir & LeBoeuf, 2002).

Affect also influences decision making and causes distortions in the estimation of likelihood and in the outcomes or consequences that are anticipated. The emotional context in which the events occur is important to their mental representation and, hence, in the estimation of their likelihood. Things of an affective nature are conspicuous, easy to picture and lead to their frequency being overestimated, so the more intense an affect is, the more vivid are the mental images it generates. Events and thoughts with an emotional load are represented better, they provide mental images more readily and hence the consequences of the decisions are easier to anticipate. Decisions also depend in part on the memory of similar experiences and whether these were positive, i.e. they depend on the hedonic recollection of the event. The memory of an affective event does not depend on objective characteristics. The biases of memory, such as forgetting about more negative aspects of a situation, may lead to false expectations and a wrong decision being taken. In such a case, we would be talking about *remembered utility*, and this may guide new decisions (Berridge, 2003). Finally, it is worth mentioning that the value of the consequences changes over time, affecting the

perception of losses and gains. The mathematical models of Johnson and Tversky (1983) show a steeper probability slope for losses than for gains. Gains and losses have marginal, decreasing values but the speeds are different. The pain of a loss increases more rapidly than does the pleasure of an equivalent gain. It is because of this lack of compensation that we develop an aversion to loss.

## 2.2. Perspective Bias or "Framing"

One of the most noteworthy aspects of Tversky and Kahneman's work was the identification of a series of decision making tactics, or *heuristics*, and of *biases* that we use to save time and effort, especially in complex or uncertain situations. During our lives we often take quick decisions without sparing too much thought. We resort to mental "shortcuts" which are in the main irrational and are not based on logic or the calculation of probabilities and frequently adhere to mental economy.

Heuristics are *explanatory rules* which can be defined as simple rules for solving complex problems. It is a question of intuitive probability that we use when it is not possible to interpret rules of reasoning, when there is no opportunity for analytical deliberation, or when this is diminished. Since simplifications are involved, these distort logical inferences, they bias the subjective evaluation of the likelihood, they deviate from perfect rationality and they generate decisions which might lower wellbeing with respect to the rational option (Kahneman, 2003).

The studies on heuristics by Tversky and Kahneman (1981) have clarified some aspects of the role of emotion in decision making, especially those which depend on the "framing" of the information. The framework or perspective is the idea or image the person taking a decision about actions, consequences and contingencies associated to a specific decision has and which lead the person to make judgements and take decisions according to how the data are presented. The principle of mental economy leads to the consideration of situations just as they are offered and does not take into account the alternatives or other solutions, which clearly makes it difficult for more reasonable alternatives to be considered. The frame is controlled in part by the formulation of the problem, as well as by the norms, habits and characteristics of the decision maker. Thus, people possess "reference points" which serve to compare or contrast the information available. This reference point may be influenced by the information which is given, so if the reference point is changed, the decision may also be changed, even though the final outcome of the various options is the same. The change in perspective or framework can alter the relative attraction of a choice.

The gain or loss framework may determine the decision taken. In the initial experiments made by Tversky and Kahneman, subjects were given problems like selecting medical treatments whose efficiency was described in terms of success or failure. Thus, the group presented with the *success* version reported a greater willingness to recommend the treatment than those who had received the data in terms of failure. When the threat of a loss is experienced, more is risked, but when a certain gain is anticipated we tend to be more conservative. A situation similar to the "framing" provoked by the anticipation of a possible loss of one's gains occurs in the economic decisions in iterative games, in which cooperative behaviours tend to disappear (Rilling et al., 2002). The final goes are like a game with a single go and a risk of loss. The subject has already won an amount, and is thus in a gain

framing and, if he or she cooperates or shares the money they will not be able to recover it in other goes. The only "feedback" information if they collaborate is a loss, and so the player opts conservatively, i.e. neither to risk nor to cooperate.

According to studies by Rottenstreich and Hsee (2001), in situations of a *high likelihood* of gain the preference is for the reward which is low in affect rather than for the one which is high. One calculates more and one risks less, so that the gain is assured and the affect decreases. In situations of *low likelihood* of gains, the reward which is rich in affect is preferred. In other words, when the loss is certain the tendency is to risk.

## 2.3. Experienced Emotion as Informative Input in Decision Making

Several authors (for a review, see Loewenstein, Weber, Hsee & Welch, 2001) have proposed that emotion has the rank of informative input in decision making, especially in situations which are complex, risky or uncertain. Decisions may be influenced by affective aspects of the situation or by the consequences of the decision. When divergences exist between the cognitive, rational assessment of the risk of a decision and the emotional reaction to taking the decision, it is the emotional reaction which dominates. In such cases, emotions provide basic, imprecise or barely elaborated information about the behavioural options and lead to decisions being taken rapidly.

These authors distinguish between *anticipated* emotions, which arise from the possible consequences of the decision anticipated by the subject, and emotions *experienced* while taking the decision. The framing effect would be placed within the sphere of emotions anticipated by the foreseeable consequences of an option as a function of the status (gain or loss) of the subject. Both are important, although these authors underline the role of the affect experienced at the moment of taking a decision. One example of the *emotions experienced* during decision making occurs when the process itself generates unease. Taking difficult decisions can cause anxiety and make the situation itself aversive, which frequently leads to avoidance behaviour, which consists of delaying the decision.

According to Loewenstein et al. (2001), people in a positive mood make optimistic judgements and choices, while fear and anxiety lead to wary choices and an aversion to risk. In situations of happiness, an automatic processing of the situation predominates, together with a lower perception of risk, which leads one to risk more. When a person is optimistic or is in a good mood, it may be that they will think that everything is going to go well. It is easier to believe that they might win the lottery and they are therefore more likely to buy a ticket if the opportunity arises.

Several experimental studies support the influence of emotions experienced when taking decisions. Johnson and Tversky (1983) found that the mood has an impact on estimations of risk frequency. The negative mood generated from reading sad press stories causes an overall increase in the estimate frequency of risk and leads to overestimation of rates of illnesses, although not of crimes or disasters. In contrast, reading a story with a happy ending improves the reader's mood and the estimated frequency of risk decreases. The effect arises when the event provoking the emotion is *incidental* and not related to the decision to be taken, in this case the estimation of risk frequency, i.e. the effect is independent of similarity between the story and the risk. Assessment of risks in daily life (accidents, thefts, attacks) is made, through lack of data, on the basis of intuitive and subjective judgements. We make

judgements that are compatible with our frame of mind, even though the subject of the judgement has nothing to do with what has caused that frame of mind. Valdesolo and DeSteno (2006) reported that inducing a positive mood increased the likelihood of rational or utilitarian judgements being made for moral dilemmas with an emotional load. Since the negative affect of such dilemmas competes against rational, objective deliberation, decreasing the emotional load allows the rational analysis to prevail in the moral dilemma. Reciprocally, a hypnotic suggestion of negative moods, such as disgust, leads to more severe moral judgements (Wheatley & Haidt, 2005).

In the same line, Lerner, Small & Loewenstein (2004) induced disgust and sadness emotions through video sequences and observed that negative emotions alter the estimated values when buying and selling. In particular they influenced the "endowment effect", the overvaluation of what one has, shown by fixing the selling price above the purchase price. Disgust eliminates the endowment effect, reducing the selling and purchase values, while sadness inverts the endowment effect and increases the purchase values and reduces the selling values. The authors conclude that emotions affect economic decisions, even though the emotions may have been induced by events which are not related to the task. Induced emotions, and their effects, go beyond the situation in progress. Winkielman, Berridge & Wilbarger (2005) found that subliminal masked presentation of faces expressing happiness led thirsty subjects to drink more of a soft drink, which was also considered more desirable, and so subjects were willing to pay more for it. The opposite effect appeared in those subjects shown sad masked faces.

Different research studies reveal that many errors in decision making may be due to *emotional alterations* (Baumeister, 2003; Gray, 1999). Affective states reduce the number of favourable choices and the effect is accentuated when the emotions are more intense (Peters & Slovic, 2000). In states of high emotion affective decisions dominate, while in those of low emotion it is the cognitive decisions that dominate. Experiments by Baumeister (2003) show how angry people accept higher risk options in which the loss is higher than the expected gain. People placed in a frustrating situation, with the threat of external criticism or public shame, opted for irrational, costly choices. Baumeister found that the angry subject neither thinks nor reflects properly about the consequences of the decision. But if the consequences are clearly explained, then the person will choose more favourable options, despite the anger.

Gray (1999) reported that under emotional states involving fear, a bias is produced towards decisions with short term results but which may be bad over the long term if repeated. There exists an evolutionary justification for this tendency to act over the short term. In situations of danger the adjustment behaviour needs to respond to the immediate contingencies before future ones, whatever the cost. In a neutral context, the tendency would be to give longer reflection to long term consequences. Gray's experiment used temporally extended choice and the presentation of pictures from the International Affective Picture System (IAPS) to provoke negative affective states. He also obtained the same result with academically stressed students.

Positive feelings influence the cognitive organization and the processing of information (Isen, 1993). People in a positive mood increase the number of creative solutions to a problem, they favour the recall of a positive outcome and show a more meticulous behaviour, although only with more interesting, pleasant or neutral materials or material they consider important. This author reports that people in a positive mood show changes in cognitive organization, and in word association tasks show more cognitive flexibility and produce more

associated words. Subjects have greater intrinsic motivation; they look for variety in the options, although these must be safe and pleasant. They show a preference for risk when taking decisions when a small amount is at stake, although not when there is a large one. They also reveal an "illusion of control" or a tendency to behave as if they could somehow control random events, and they prefer to increase the risk, especially when it is hypothetical and not real. Isen defends the idea that people in a positive mood are risk averse in order to maintain their good mood (Isen, 1993; Loewenstein et al., 2001).

Sánchez-Navarro and Martínez-Selva (2006) assessed how emotions induced by IAPS pictures influenced decision making. One group of subjects were shown pleasant pictures and another, unpleasant ones. Throughout the task autonomic responses were recorded (electrodermal activity and heart rate) and also muscular responses (corrugator supercilii), with the aim of checking that the pictures provoked an emotional physiological reaction. The results showed that subjects viewing pleasant pictures did worse in the test, making more unfavourable choices than those viewing unpleasant ones. So, pleasant affective states bring with them a greater propensity to risk than do unpleasant ones. The more wary behaviours under negative emotions can be compared to the results of Lerner et al. (2004) for sad subjects, who overvalued purchase prices, indicating conservative consumption.

Sanfey, Loewenstein, McClure & Cohen (2006) propose that decision making depends on the interaction of different specialized subsystems, proceeding from two forms of processing which can compete or act synergistically:

- 1. *Automatic processes*, which are fast and efficient and often act in parallel and in specialized domains. They are relatively inflexible. They would be equivalent to an automatic system of heuristic responses and of intuitive responses, which these authors call System 1.
- 2. *Controlled processes*, which are conscious, voluntary, accessible to introspection, explicit, flexible and of limited capacity. They would be equivalent to a controlled processing system which monitors the response from System 1 and corrects it if necessary. This type of process is called System 2.

System 2 possibly fits in with the type of information processing expected by the rational theory of usefulness. Yet it would only explain a part of the behaviour, since it would frequently compete with System 1, based on specific automatic processes like those occurring in emotional processes. For these authors, emotional states lead to a decrease or lessening of the risk aversion effects, such that:

- 1. Positive emotion, caused for example by viewing images, would lead to a decrease in risk aversion and to unfavourable choices being made.
- 2. Negative emotion, caused by viewing negative images would lead to a greater risk aversion and wary, more favourable decisions.

## 2.4. Conclusions: Anticipated Emotions and Experienced Emotions

As Gray, Braver & Raichle (2002) point out, emotional states can temporally favour or prejudice different cognitive functions in a fast, flexible and reversible way, and thus bias thinking and behaviour, making them more or less adjusted.

Many decisions are taken in anticipation of their consequences. *Anticipated emotions* might help to know the consequences of a decision. Preferences towards one option or another depend on the affective reactions associated with the possible outcomes of a decision and on those occurring when the decision is being taken, and are integrated into the status or framework in which the subject is when taking the decision, which in turn may be influenced by the information received. Anticipated emotions are components of the expected consequences. They are basic, fast and in part unconscious. People assess the attractiveness and likelihood of the possible outcomes of the different choices and alternatives, with their biases and errors, integrate the information into a "calculation" based on expectations, and thus reach a decision.

*Experienced emotion* also influence decision making. In situations of sadness the tendency is to be more conservative. A pessimistic person is more sensitive to negative moods and shows more intense affective reactions. They tend to take decisions which relieve or lighten the emotional reactions, e.g. avoidance or escape. There is a greater analytical processing, with more attention to environment signals, greater distraction and slower reaction times (Mellers, 2000).

Emotions, though, may benefit or prejudice. So fear may be positive in threatening or life threatening situations, since analytical processing can help to find keys to escape from the situation and facilitate reactions to fight or to flee. In this case, the emotion is congruent with the decision. However, the same does not occur in many day to day situations. Emotion is beneficial in decision making when it is *integrated* into the task or directly related to it, but it can influence negatively if this is not the case. In these situations, emotion directs us towards clues or signals which might not be the most appropriate and they could lead us to make wrong or unfavourable decisions. The appearance of negative emotions not related to the task bias decisions towards the short term, and this is not always the best option (Gray, 1999). The integration of emotion may be necessary not only in situations requiring a fast answer but also in a wide range of social situations. If emotion is not incorporated into these, socially maladjusted behaviours may occur which prejudice the individual, as is shown by behaviours in economic games and when making moral judgements with high emotional implication (see sections 5.3 and 5.4). In a wider sense, emotions may provide enough energy to tackle and perform very difficult or very prolonged tasks. Great, vital decisions, with effects in many spheres, tend to rest on emotional arguments.

According to Beer, Knight & D'Esposito (2006) there are two processes relating to the role of emotion in decision making:

- 1. *The flexible incorporation or integration of the emotion* in cognition when this is relevant to the process.
- 2. *Reducing or impeding the influence of the emotion* in decision making when it is not relevant, which often leads to the need to put emotions aside when deciding. This process would be related to emotional control, which is dealt with in section 5.

# 2.5. How Emotion Influences Decision Making: The Somatic Marker Hypothesis

The most complete account for the role of emotion in decision making is known as the *somatic marker* hypothesis, and was put forward by Damasio (1994). A somatic marker is a bodily change which reflects an emotional state, be it a positive or a negative one, and it may influence the decision taken at a given moment. Anticipation of possible consequences of a decision produces emotion based somatic responses which guide the decision making process. The responses arising from anticipating possible consequences have their origins in the emotional reaction produced by decisions taken in the past in similar situations. The somatic marker facilitates and speeds up decision making, especially in the case of social behaviour, where situations of greater uncertainty may arise.

In this context, the emotional response is the subjective and somatic reaction of the individual to an event - e.g. the positive or negative consequences of a decision. When the reaction is associated to a situation or to a set of stimuli, it may consciously or unconsciously affect future behaviour, and is thus a somatic marker. Such markers, which are understood experimentally as autonomic, muscular, neuroendocrine or neurophysiological changes, can provide unconscious signals which precede, facilitate and contribute to taking a decision, even before the subject can explain the reason for the decision or can state conceptually the strategy being used to take the decision (Bechara & Damasio, 1997; Bechara, Damasio, Tranel & Damasio, 2005). Thus, somatic markers support cognitive processes, they permit an appropriate social conduct, they contribute to advantageous decisions being taken (by inhibiting the tendency to seek an immediate reinforcement) and they facilitate the representation of future scenarios in the working memory (Bechara et al., 2005). On the other hand, the absence or the weakening of somatic markers lead to inappropriate or disadvantageous decisions being made. This deficit occurs in patients with ventromedial prefrontal lesions or lesions to other frontal regions like the dorsolateral prefrontal cortex (PFC) and the cingulate cortex, as well as in patients with bilateral lesions to the amygdala, who are incapable of experimenting emotions appropriately or of generating autonomic responses to aversive stimuli.

The experimental study of somatic markers and their role in decision making has been made using the Iowa Gambling Task (IGT) (Bechara, Damasio, Damasio & Lee, 1999). The changes in electrodermal activity arising from the decision making situation are the most studied of somatic markers. These changes are the result of a sympathetic activation, caused by the situation, the mental effort and the emotional activity, including the anticipation of the consequences of the decision and the memory of earlier emotional reactions to the consequences of a similar option. Bechara and colleagues (Bechara & Damasio, 1997; Bechara et al., 1999; 2005) report that normal subjects show skin conductance responses caused by the consequences (gains or losses) of their decisions, which are greater when the reward or punishment is more intense. At the beginning of the task cards are selected from the disadvantageous decks -because of their higher short-term gains- but as the task progresses, subjects start to choose cards from the advantageous decks -because of their higher long-term net gains- and show higher skin conductance responses prior to choosing a disadvantageous deck. These responses have been interpreted as somatic markers associated to the choice made by the subject, which derive from experience gained from the consequences of their previous choices. The higher magnitude of the responses prior to choosing a disadvantageous deck represents the accumulated bodily signal which biases or guides the subject away from that deck. Inexistent anticipatory conductance responses or those of low intensity are associated to a higher selection of unfavourable cards and, hence, a bad performance in the task. This occurs in subjects with prefrontal lesions, especially to the ventromedial sector (Bechara & Damasio, 1997; Bechara et al., 1999; 2005).

These data have been replicated in part by other researchers (Crone, Somsen, Van Beek & Van der Molen, 2004; Suzuki, Hirota, Takasawa & Shigemasu, 2003; Tomb, Hauser, Deldin & Caramazza, 2002). The losses, in comparison with the gains, cause increases in the skin conductance along with a deceleration in heart rate. Subjects who perform well in the IGT are those who show the most pronounced deceleration, especially after losses, and a higher increase in skin conductance prior to unfavourable choices. On the other hand, subjects performing badly show no differences in their anticipatory autonomic activity prior to favourable and unfavourable choices (Crone et al., 2004). So, as the somatic marker hypothesis proposes, there exists a subgroup of normal subjects who show reduced physiological responses during the IGT and who also perform the task badly, which is a similar situation to that of patients with lesions in the ventromedial sector of the PFC. An indirect proof of the need for somatic markers in decision making is found in a series of experiments by Hinson, Jameson & Whitney (2002) in which an overload of the working memory results in diminished conductance responses and a worse performance of the task.

Other authors, e.g. Maia and McClelland (2004), do not discard the existence of emotional mechanisms or unconscious processes that guide decision making, although they clearly state that no proof of a somatic marker exists to date. The appearance of skin conductance responses associated to choices does not necessarily mean that there is an emotional activity which is guiding the decision. They also believe that the procedure used in the IGT by Damasio and colleagues to evaluate the degree of knowledge or consciousness that the subject has of the task is not an adequate one and, therefore, it cannot be proved that the subject can achieve good results without being aware of the strategy he or she is following. In this line, data from Hinson Whitney, Holben & Wirick (2006) give support to the idea that a good performance in a gambling task precedes the appearance of the somatic markers rather than the opposite.

There are other authors who believe that the somatic marker hypothesis does not concord with the principle of parsimony (Krawczyk, 2002; Kringebalch & Rolls, 2004). Emotional reaction and its incorporation in the brain may be made directly through intracerebral neuronal circuits and with no need for an imprecise somatic reaction, as in the James-Lange theory, which the brain needs to identify. Emotion occurs as the result of a stimulus which provokes reactions in different structures of the brain, and it is not clear how the brain would have to interpret these complex feedback signals.

## **3. BRAIN MECHANISMS WHICH SUPPORT THE INTERVENTION** OF EMOTION IN DECISION MAKING

Below we review the principal structures and systems involved in the recognition of stimuli, subjective experience and activation of emotional responses and their participation in decision making. Particular consideration is given to the so called brain systems of fear and reward, as well as to the structures or systems involved in integrating cognitive and emotional information, especially the PFC. Finally, results of the experiments are given where the main aspect is the inhibition or control of emotion during decision making, or the opposite circumstance: what happens when, for various reasons, emotional processes are incorporated into rational activity or when they overwhelm it?

Following other authors, e.g. Kringelbach & Rolls (2004), we understand emotions as states which are elicited by rewards and punishments and by the stimuli associated to these. Although emotions are very varied and are present in human beings to a high degree of specificity, their deep evolutionary roots leads us to think that the majority of emotions share common characteristics. The processing of stimuli and the emission of emotional responses is due to certain brain systems being activated in which different, strongly interconnected structures are taking part. An important corpus of knowledge is available which gives a fairly good idea as to the cerebral response in negative and positive affective situations.

There is also a series of cerebral structures which systematically appear in studies on the brain mechanisms of decision making: ventromedial/orbital PFC, dorsolateral PFC, amygdala, anterior cingulate cortex (ACC), striatum, insula and hippocampus (Elliott, Rees & Dolan, 1999; Ernst et al., 2002; 2003; Martínez-Selva, Sánchez-Navarro, Bechara & Román, 2006). All these structures either form part of, or are directly connected to, the brain systems involved in processing emotional stimuli.

One important difficulty in this sphere arises from the different nature of the tasks employed in decision making studies, which range from simple gambling tasks, complex or extended tasks like the IGT, to moral judgements of greater or lesser emotional implication, to which are added economic decisions which belong to the field of game theory. The variations in the complexity or uncertainty induced by the task are also critical. Data on the participation of different regions of the brain are also varied and come mainly from neuroimaging studies (functional magnetic resonance imaging fMRI, positron emission tomography, PET) in normal and brain damage subjects, and also in substance abusers.

## 3.1. Brain Defence System

Affective stimuli are rapidly processed by the brain via a distributed network of brain structures representing their different characteristics and response options (Davidson, Maxwell & Shackman, 2004). These stimuli provoke an attentional demand with greater processing effort and higher cortical activation than do non affective ones. When they are threatening, their detection follows automatic processes, which are largely unconscious, using parallel information processing systems, which are triggered by relatively simple stimulus characteristics. One example of these would be the *brain defence system*, based in the amygdala and proposed by LeDoux (1996), and in which other structures participate, like the insula, the ACC, the hippocampus and the frontal cortex. These stimuli can be processed in the amygdala with no intervention of the cortex, which allows for a rapid detection of those which are potentially or directly dangerous, even of those stimuli that are below the conscious threshold (Ishai, Pessoa, Bickle & Ungerleider, 2004; Öhman, 1993; Pessoa and Padmala, 2005). When subjects were presented with faces showing expressions of fear, Ishai et al. (2004) found that they responded faster and more precisely than when presented with neutral stimuli, provoking a greater activation in various regions of the brain, such as the inferior

occipital gyri, the lateral fusiform gyri, the superior temporal sulci, the amygdala and the inferior frontal gyri/insula. The amygdala and the insula are sensitive to facial expressions, in particular to negative emotions like fear, anger and disgust. It may be that the response of these extrastriate visual areas to negative facial expressions is modulated by the return connections of the amygdala.

## 3.1.1. Amygdala

The amygdala is composed of a set of nuclei located in the anterior part of the temporal lobe, which are involved in the recognition of aversive stimuli and the reaction to these. This structure would play a key role in the brain defence system, which is triggered by emotional stimuli, especially aversive ones, and which would set off emotional and defence responses and favour the processing of affective and significant stimuli which are important for survival. The system is triggered by stimuli from different sensory modalities, whose signals converge in the amygdala, which would participate in identifying the value of the signal from the stimuli and would process their emotional content and generate autonomic responses through the hypothalamus and nuclei of the brain stem, thus arousing approach/attack and withdrawal/flight behaviours. It also participates in the recognition of affective facial expressions, especially those of fear. All the sensory systems send projections to the amygdala, although some do not come directly from the sensory cortical areas but from the thalamus. Sensory information arrives first at the lateral nucleus of the amygdala, then at the basal and accessory basal nuclei, to which sensory information is supposedly associated, and then at the central nucleus, which is its main output pathway (Adolphs et al., 2005; Davis, 1997; Emery & Amaral, 2000; Tranel, 2000).

The amygdala is involved in learning processes with emotional content, like the acquisition of fear conditioned by signals which anticipate a danger or threat, since it is there that the sensory information from the conditioned stimulus and the aversive unconditioned stimulus converge (Le Doux, 2000). The importance of the amygdala in processing emotional stimuli with a negative affective load has also been highlighted in studies using fMRI (Critchley et al., 2005; Kuniecki, Urbanik, Sobiecka, Kozub & Binder, 2003; Sabatinelli, Bradley, Fitzsimmons & Lang, 2005). Davidson (2002) reports that the amygdala is only necessary in the initial learning stages, and is associated to an early attentional stage.

The amygdala exerts some of its effects through the stria terminalis and the ventral amygdalofugal pathway which reach the lateral and paraventricular hypothalamus, and activate autonomic and hormonal responses. Its impulses are also directed towards the central gray and set in motion the brain's activating systems, thus favouring the analysis of the potentially threatening stimuli –by increasing the level of alertness and improving the processing of the information, which allows for a more detailed analysis of the stimuli and thus prepares the suitable response. The amygdala acts on the brain stem motor regions, such as the facial motor nucleus and sets off the facial emotional expression, especially in the case of negative emotions. The hypothalamus seems to be the structure which triggers the autonomic and hormonal responses which accompany the emotions (LeDoux, 1996; Sánchez-Navarro, Martínez-Selva & Román, 2006).

The emotional activation of the amygdala occurs *to* stimuli with either positive or negative affective value, although it is accentuated in the latter. The study of the brain response to anticipation and experience of gains and losses using the fMRI reveals that the sublenticular extended amygdala is activated during the anticipation of the reinforcement and

in response to the results arising from a decision (Breiter, Aharon, Kahneman, Dale & Shizgal, 2001). Monetary rewards in a simple betting task cause the activation of the amygdala, as does the anticipation of primary rewards, such as a sweet taste (Elliott, Friston & Dolan, 2000; O'Doherty, Deichmann, Critchley & Dolan, 2002). Nevertheless, Breiter et al. (2001) report that it is activated more when the expectations of reinforcement are negative, especially in the left hemisphere. De Martino, Kumaran, Seymour & Dolan (2006) have studied the neural mechanisms of the framing effect in subjects that were in a gain frame, who tended to opt for a sure gain rather than a risk and to opt for a risk when they were in a loss frame. They find that this bias acts through emotional activation, in particular that of the amygdala, and independently of the type of frame.

Lesions to the amygdala lead to a deficit in recognition of facial expressions of fear, docility, blockade or deficits in responses conditioned by fear in animals and humans, and to the disappearance of startle reflex modulation caused by fear (Adolphs et al., 2005; Aggleton & Young, 2000; Buchanan, Tranel & Adolphs, 2004; Davis, 1996, 1997; Dolan & Morris, 2000). Patients with amygdala lesion do not show electrodermal responses as a reaction to rewards and punishments, nor autonomic responses conditioned by fear, and they do not appear to experience sufficiently the emotional aspects of affective situations, although they do possess an explicit, conscious appraisal of the situation (Aggleton & Young, 2000; Davis, 1996; 1997). They perform worse in the IGT than the control subjects and they do not show anticipatory conductance responses for unfavourable choices, which would be indicating the absence of a somatic state that might signal the consequences of an action (Bechara et al., 1999; Tranel, 2000).

According to the somatic marker hypothesis, it is in the amygdala where the associations or connections between the stimuli and their aversive consequences would be produced and which are used later in taking decisions in similar situations. This structure could process the emotional content of unfavourable choices and generate autonomic responses through the hypothalamus and brain stem nuclei, and it is therefore possible that the cognitive evaluation of the situation is made on the basis of the previous emotional one (Bechara et al., 1999; Martínez-Selva et al., 2006).

## 3.1.2. Insula

The insula is a cortical region which processes interoceptive and taste sensations, but it is also associated to affect and to negative situations and to the anticipation of aversive stimuli such as pain. Like the amygdala, it is activated (especially its anterior portion) when affective stimuli like faces and emotional expressions are being processed (Carr, Iacobini, Dubeau, Mazziotta, & Lenzi 2003; Ishai et al., 2004; Pessoa and Padmala, 2005), and it plays an important role in the cortical representation of the autonomic changes of anxiety (Davidson & Irwin, 1999). Data coming from PET studies show that a greater activation of the left insula is associated to better performance in the IGT (Ernst et al., 2003). Kuhnen and Knutson (2005) found that in tasks involving anticipation of gains and losses the insula would be associated to the prediction and anticipation of losses. The studies by Knutson, Rick, Wimmer, Prelec & Loewenstein (2007) on decisions on buying show that before taking the decision to buy, the excessive price activates the insula at the same time as it deactivates the medial PFC, with the activation of the insula being directly proportional to the excessive price. Similarly, high rewards and punishments increase emotion and also activate the insula (Elliott, Friston et al., 2000).

Shafritz, Collins & Blumberg (2006) indicate that in situations with an intense motivational or emotional component, the anterior insula incorporates the affective information into the decision making process in order to integrate this information and then proceed to the inhibition or the activation of the emotional reaction.

Another point about the participation of the insula in processing aversive information comes from its greater or lesser activation during participation in the Ultimatum game, in which a subject receives a more or less favourable offer to share an amount of money with the player who proposes the offer (Sanfey, Rilling, Aronson, Nystrom & Cohen, 2003). If the offer is accepted, both players share the stated amount, but if the offer is refused, neither player gets anything. This means that when a low offer is received there will be a conflict as to whether to accept or to refuse. The usual response to a low offer is to refuse it, which may include some anger. In such cases the anterior insula is bilaterally activated, which points to emotional activation and negative affect. The bilateral activation of the anterior insula is greater as the offer worsens. Moreover, its activation is positively correlated to the refusal and predicts it (Sanfey et al., 2003). The refusal of an offer seems to be the result of a strong emotional reaction, which would indirectly support Damasio's "somatic marker" theory. Given its role in processing visceral signals, this structure could participate in capturing and transmitting responses or somatic markers associated to the consequences of favourable and unfavourable choices by incorporating the negative somatic markers into the decision making (Paulus, Rogalsky, Simmons, Feinstein & Stein, 2003).

## 3.1.3. Anterior Cingulate Cortex

The ACC is a paralimbic region which carries out various functions related to the processing of information from different sensory modalities, including nociception. It also participates in motor, cognitive and complex emotional processes. It is involved in processing emotional information, especially in conflict situations, and in the anticipation of the consequences of a choice. It is closely related to the striatum and to other regions involved in the processing of emotional information: PFC, amygdala, central grey, nucleus accumbens (NAc) and hypothalamus (Bush, Luu & Posner, 2000). This region, especially in the right hemisphere, appears as more active when negative consequences are experienced (Gehring & Willoughby, 2002; Knutson, Westdorp, Kaiser & Hommer 2000). Functional neuroimaging studies relate the ACC with the control or monitoring of the behaviour, which includes response assessment and inhibition processes, in which the lateral orbitofrontal cortex is also involved (Volkow, Fowler & Wang, 2004). It is activated in response to sad faces and to utterings of emotional content (Shafritz et al., 2006), and in circumstances like incongruence or conflict between options (Ernst et al., 2002; Mac Donald, Cohen, Stenger & Carter 2000). Electrophysiological data place the error-related negativity potential which appears when committing errors in the cingulate cortex (Hajcak, Holroyd, Moser & Simons, 2005).

Lesions to the ACC produce disorders in behavioural control and in the capacity to evaluate the risks or effort involved in the search for rewards (Rogers et al., 2004). This region, together with the orbital PFC, seems to be more active when performing decision making tasks involving risk or uncertainty (Cohen, Heller & Ranganath, 2005; Milham, Banich, Claus & Cohen 2003). Processing the characteristics of the context may be another function of this structure. In a betting task with symbolic monetary gains, Elliott, Friston et al. (2000) find activation of the subgenual ACC (Brodmann Area, BA 25) when the reward is high, in the case a run of successes.

Nevertheless, a distinction is usually made between two sectors of the ACC, which would take part in different functions – the dorsal or "cognitive" region and the rostral-ventral region related to the ventromedial/orbital PFC, with affective functions. Executive functions are among the cognitive functions attributed to the dorsal region, e.g. selective attention, task management, which is activated when performing dual tasks along with the dorsolateral PFC (Smith & Jonides, 1999). The ACC would be activated in supervising the result of the choice and in correcting it.

The rostral-ventral portion of the ACC appears to be involved in the assessment of conflict, especially when affective aspects are intervening (Bush et al., 2000). Processing conflict situations also involves the PFC which, in turn, controls responses (Bunge, Hazeltine, Scanlon, Rosen & Gabrieli, 2002). De Martino et al. (2006) report that ACC activity increases when the subjects' tendency to respond does not, according to the loss or gain frame, correspond to their decisions. The opinion of these authors is that this reflects the opposition of two brain systems – one involved in rational and analytical decisions, and the other, based on the amygdala, related to emotional ones. In the Ultimatum game (Sanfey et al., 2003), when player one makes a low offer, his or her ACC is activated, relating it to situations of cognitive conflict (in this case the conflict lies in the fact that the player making the offer has more to win than the player receiving the offer, while there is the risk that the latter refuses the offer, which means that both players lose, or that the offer is considered offensive).

Taken together, the research leads us to think that this structure forms part of an attentional circuit which regulates both cognitive and emotional processes (Bush et al., 2000; Critchley, Mathias & Dolan, 2001). Krawczyk (2002) places the role of the cingulate cortex in decision making in ambiguous situations with a high likelihood of errors being committed. While the orbitofrontal cortex is related to associations of the stimuli with the reward, the ACC would intervene in the control and choice of the most suitable behaviours, the detecting of errors and the calculations of the likelihood of a reward. According to Shafritz et al. (2006), this control function of the ACC may include the inhibition of inappropriate responses in situations which are motivationally or emotionally important.

Both the ACC and the anterior insula represent the emotional experience of different kinds of stimulation, including pain. Both structures would contribute to forming subjective representations that allow the assessment of the effects of affective stimuli, and the choice of appropriate responses, by facilitating or inhibiting them. Singer et al. (2004) indicate that the anterior insula would act as a relay structure for the sensory information, which is later integrated and processed in other regions such as the ACC and the PFC.

## 3.1.4. Uncertainty and Risk

It is to be expected that the brain defence system would be activated in situations related to the anticipation or expectation of losses or negative results in decision making tasks. The greater the risk, complexity or difficulty of the decision, the greater the activation of structures related to the brain defence system. In situations of uncertainty, the subject cannot know which responses are right and it is difficult to infer them from the information provided or from the results of the choices. Critchley et al. (2001) studied the brain metabolic response during the anticipation of reward. Uncertainty activated both the ACC and the orbitofrontal cortex. The ACC was also activated by the arousal level. In this respect, Hsu, Bhatt, Adolphs, Tranel & Camerer (2005) distinguish between *risk* situations, where the likelihood of the result may be known from earlier data, and *ambiguous* ones, where likelihood cannot be

adequately estimated because relevant information is missing. In the first case, the authors find an increase in the activity of the dorsal striatum and the caudate, which correlates with the value expected for choices. In contrast, in ambiguous situations, of greater uncertainty, a higher activation is observed in the amygdala, indicating an emotional reaction, in the orbitofrontal cortex, related to the integration of cognitive and emotional information, and the dorsomedial PFC, related to the working memory and the modulation of amygdala activity. In ambiguous situations, the activity decreases in the striatum, which is a part of the brain reward system.

The choice of risk options is accompanied by an increase in activity in the right anterior insula (Paulus, Hozack et al., 2003). For these authors the insula processes risk during decision making and they conclude that a neural system seems to exist which assesses the uncertainty, and which is based mainly in the amygdala and the orbital PFC. These two structures receive multimodal sensory information, are interconnected and function together in assessing the value of the stimuli. They both, but especially the amygdala, detect relevant stimuli that indicate the level of uncertainty. Moreover, both are connected to the striatum and, therefore, intervene in detecting the value of the reward of the stimuli.

#### 3.2. Brain Reward System

Positive emotional states are associated to primary reinforcers like food or sexual conduct, the consumption of certain drugs, e.g. stimulants, and to secondary ones like money or signals associated to the primary reinforcers. Where positive emotions, or their anticipation, prevail, the so-called brain reward (mesolimbic-mesocortical) system comes into play. It is accepted that this system has its origins in the dopaminergic neurons of the tegmental ventral area and the substantia nigra, that project to the striatum, especially the caudate and NAc, and also to the amygdala, medial temporal lobe, hypothalamus, septum, ACC and ventromedial/orbital PFC.

The brain reward structures seem to possess different functions. Knutson, Taylor, Kaufman, Peterson & Glover (2005) find that the NAc is more activated according to the anticipated reward while the medial PFC would be more sensitive to changes in the anticipated likelihood of reward. This indicates that at the subcortical level, affective reactions are more represented, while the cortical levels take charge of more elaborate judgements. Thus, Berridge (2003) has pointed out that the cortical portions appear to be involved in the cognitive aspects of anticipated or recalled reinforcement, which implicate memory and conscious expectations. Elsewhere the dopaminergic mesolimbic system takes part in the basic aspects of reward or expected utility.

The dopaminergic neurons of the substantia nigra and of the ventral tegmental area are activated in response to natural or symbolic rewards, like monetary ones, including social stimuli and substances like amphetamine and cocaine (Breiter et al., 1997; Volkow et al., 1999). At rest, these neurons show a stable firing rate which increases considerably in the presence of significant stimuli (e.g. an unexpected reward) and of stimuli which anticipate a reward or reveal the availability of reinforcement. If the reward is associated to a signal, the response of the dopaminergic neurons may appear to the signal and not to the reward. The changes in temporal predictability, which have been studied in operant conditioning in animals, activate the NAc. When deviations in the expected reinforcement occur, the dopaminergic neurons of the ventral tegmental area and of the substantia nigra show a phasic increase in their firing rates (Berridge, 2003; Schultz, 2004; Shizgal & Arvanitogiannis, 2003).

Activation of the reward system is independent of the type of stimulus or its sensory modality. In humans, sexual activity and the observation of pictures of loved ones raise the dopamine secretion and activate the dopaminergic regions. Viewing beautiful faces activates the brain reward system, and in particular the NAc (Aharon et al., 2001). Similarly, gambling tasks are associated with a greater release of dopamine (Bechara & Damasio, 2005). Obtaining monetary rewards in tasks involving a certain degree of uncertainty brings about the activation of not only the brain reward system but also of a wide group of structures which participate in decision making, such as the orbitofrontal cortex, medial caudate and ACC (Kuhnen & Knutson, 2005).

As regards rewards of a social type, the studies by Moll, de Oliveira-Souza, Bramati, & Grafman (2006) show that when a person decides to make a donation to a freely chosen cause, the mesolimbic-mesocortical reward system is activated, especially the ventral tegmental area and the dorsal and ventral striatum, as occurs when receiving financial rewards. In situations of social interaction and games like the "trust game", the dorsal striatum is activated when there is reciprocity or when a player receives a symbolic amount as a reward for an earlier altruistic behaviour. The differences between benevolent and malevolent reciprocity (in which the generosity of the first player is rewarded or punished) are manifested by a greater activity of the caudate in anticipation of the benevolent response. The head of the caudate nucleus integrates information about the goodness of the partner's decision and the intention to correspond (King-Casas et al., 2005). For their own part, the studies by Knutson et al. (2007) on decisions to buy show that a preference for a product causes an increase in the mesolimbic activation (brain stem, NAc, prefrontal medial cortex and other dopaminergic areas). The authors find that the activation of these structures predicts subsequent purchases above and beyond the subjective reports. In contrast, the excessive price activates the insula, which is related to negative emotion, and deactivates the medial PFC before taking the decision to buy.

## 3.2.1. Anticipation of Reward

Decision making is linked to anticipation of both the positive and negative consequences of the different choices. Activity in the brain reward system increases when rewards or gains are anticipated and obtained. Variations in the different characteristics of these consequences (type of stimulation, magnitude, lag) and the fact that they are more or less predictable are especially manifested in the ventral striatum and the ventromedial PFC.

Studies using fMRI show that some of the regions of the brain reward system are activated both in anticipation of the reinforcement and when obtaining or receiving it (Breiter et al. 2001; McClure, Berns & Montague, 2003). Likewise, other studies also underline the critical role of the ventral striatum in the prediction or anticipation of the reinforcement or the punishment (Seymour et al., 2004). Anticipation of a reward activates the ventral striatum, including the NAc, and the medial prefrontal cortex (Breiter et al., 2001; Critchley et al., 2001; Knutson, Fong, Bennett, Adams & Hommer, 2003). The ventral striatum, particularly the NAc, is involved in prediction errors which occur when the reward does not appear at the expected time or is not of the expected magnitude (Pagnoni, Zink, Montague, & Berns, 2002). Anticipation of gains activates the NAc and correlates with a positive mood, possibly

associated to the release of dopamine (Knutson, Adams, Fong & Hommer, 2001; 2005; 2007). In addition, NAc activation is proportional to the size of the anticipated response in monetary rewards (Knutson et al., 2001).

Breiter et al. (2001) found a brain response to the anticipation and experience of gains and losses, and they distinguished between the expectancy and the achievement phases of the reward. The expected reinforcement would activate the amygdala and the orbitofrontal cortex, while the reinforcement obtained activates the NAc, the sublenticular extended amygdala, the hypothalamus and the ventral tegmental area. These are the same regions activated by primary reinforcers, like olfactory and taste ones. When the reward is obtained, the activation of these regions is proportional to the size of the reward. Another exception to this overlap between brain regions of anticipation of the reward and obtaining it comes from the studies by O'Doherty et al. (2002), who found that anticipation of a pleasant taste activates the midbrain dopaminergic centres, the posterior dorsal amygdala (next to the hippocampus), the striatum and the orbitofrontal cortex. The last of these was also activated on obtaining the taste. Thus, the regions intervening in anticipation can be separated from those corresponding to the reception of a reinforcement when the rewards are predictable.

Other authors have found greater activation of the ventral part of the caudate nucleus with respect to the size of the behavioural change in an instrumental learning task (Haruno et al., 2004). Activity of the medial caudate nucleus anticipates both rewards and punishments and is more related to the size than to the valence of the reinforcer (Knutson et al., 2001). Delgado, Nystrom, Fissell, Noll & Fiez (2000) found that both the ventral and the dorsal striatum, and especially the caudate nucleus, are activated both in the anticipation and in the obtaining of monetary rewards. Elliott, Friston et al. (2000) also coincide in highlighting the role of the caudate nucleus, which is activated by punishments and high rewards, together with the insula and the ventral PFC. The same authors found that a high reward also activated the pallidum, the thalamus and the subgenual ACC (in its rostral-ventral portion, BA 25), related to the orbitofrontal cortex and the processing of emotional information, as seen above.

Activation of the NAc precedes risky decisions and, consequently, errors arising from risk choices (Kuhnen & Knutson, 2005). The NAc is activated more in anticipation of monetary gains and positive affect and would be associated to the prediction or the expectation of gains. These authors relate the effects of the NAc activation to the influence of emotion in decision making. Anticipation of a gain, the result of the reward for the earlier decision, provokes positive emotions that may lead to risky choices. In contrast, negative feelings of anxiety, provoked by loss, which in the study in question activate the anterior insula, lead to risk aversion. Associated errors based on the earlier outcome, exist in both cases, due either to risk preference or to risk aversion. This study has discerned different neural systems, one for the anticipation of gain, based in the Anterior insula and which leads to risk aversion, and another for anticipation of gain, based in the NAc, which leads to risk aversion in the anterior insula when subjects received a punishment and switched to a safer option in the next trial.

Research indicates that the activation of the brain defence and reward systems reach different areas of the frontal cortex. It has been suggested that the functions of the striatum would be to integrate information related to the reward from the information arriving from the cortical and limbic regions, which could be modulated in turn by inputs from the midbrain systems. A wide number of studies have shown that the integration of cognitive and emotional information is produced in the frontal cortex, as is laid out below.

## 4. SYSTEMS AND STRUCTURES FOR THE INTEGRATION OF COGNITIVE AND EMOTIONAL INFORMATION: PREFRONTAL CORTEX

## 4.1. Ventromedial Prefrontal Cortex/Orbitofrontal Cortex

The ventromedial PFC refers to the medial ventral region of the PFC and to the medial sector of the orbitofrontal cortex, including the BA 25, the lower portion of areas 24 and 32, and the medial sector of areas 10, 11 and 12. For its own part, the orbitofrontal region occupies the ventral surface of the frontal cortex and is made up by areas 10, 11 and 47. Area 25 represents the most caudal portion of the orbitofrontal region which corresponds to the subgenual ACC, which extends to BA 10 (Bechara, 2004; Kringelbach & Rolls, 2004).

Convergent data from studies of patients with brain lesions, together with those from fMRI of healthy subjects point to the PFC, especially the ventromedial portion, as the key region where the integration of emotional and cognitive information intervening in decision making is produced (Bechara et al., 1999; Bechara & Damasio, 1997; Krawczyk, 2002). Both the brain defence and the brain reward systems run to different sectors of the ventromedial region.

Their main functions would be to process the value of the reward from both the primary stimuli (e.g. taste) and the secondary or symbolic ones (economic or social), and to adapt to the fast changes in the value of contingencies and to generate adjustable and flexible behaviours (Rolls, 2000). This function is directly related to the working memory and executive control of the behaviour. It should also be added that it would suppress responses to stimuli that were no longer reinforcing as well as inappropriate behaviours.

The ventromedial PFC, together with the dorsolateral PFC is involved in decision making in conditions of uncertainty, especially where time is a pressure factor, and in the capacity to take decisions in new situations with no indications or with insufficient data. Glimcher and Rustichini (2004) state that the ventromedial PFC is mainly involved in learning about the availability of rewards and punishments during decision making in conditions of ambiguity. It would participate more in learning and processing sensory information than in selecting the responses.

It is possible to distinguish between the medial portion of the orbitofrontal cortex (involved in the association between stimuli, responses and consequences in changing situations, which may be considered as behavioural flexibility) and the lateral portion (involved in the inhibition of previously rewarded responses). The latter region is associated with risky decisions, related to the amygdala, insula and temporal pole, and all involved in the experience and emotional expression (Elliott, Dolan et al., 2000; O'Doherty, Kringelbach, Rolls, Hornak & Andrews, 2001).

The orbitofrontal cortex is involved in social and symbolic reinforcement, especially in the gains and losses obtained, and so the magnitude of the activation reflects the size of the monetary rewards or penalties obtained (O'Doherty et al., 2001). Furthermore, the lateral orbitofrontal cortex is associated with the assessment of the contextual relevance of emotional information in decision making, as well as with the anticipation and experience of positive and negative outcomes (Beer et al., 2006; Breiter et al., 2001).

According to Moll et al. (2006), the anterior PFC, which includes the orbitofrontal region, represents complex reinforcement contingencies related to altruistic behaviours. On the other hand, the subgenual and lateral orbitofrontal cortices respectively participate both when donating and, in the opposite case, in the aversion to donate to a cause which one rejects. As pointed out by Kringelbach & Rolls (2004), the lateral orbitofrontal cortex seems to participate in aversive mechanisms, including anger and moral rejection.

## 4.1.1. Lesions to the Ventromedial Prefrontal Cortex

Patients with ventromedial lesions usually show alterations in social behaviour, decision making and emotional processing. Although they have difficulties in learning from their mistakes, they conserve their intellectual capacities, intelligence and memory, along with the other cognitive functions, including management, at a normal level (Bechara & Van der Linden, 2005). In their social, professional and economic lives, however, these patients are prone to take decisions and adopt conducts that lead to negative consequences. They lose behavioural flexibility and show impulsiveness and difficulties in adapting to changes produced in the tasks, i.e. to new reinforcement contingencies, as well as showing perseverance in their behaviour. They have difficulties in planning their daily and future activities, in choosing friends and at work. They are insensitive to both positive and negative long term consequences, and are guided instead by immediate consequences. They reason well when faced with moral problems, they know what they should do, but they make the wrong decisions and to use emotions and feelings to assist their behaviour (Damasio, 1994).

Within the context of the somatic marker theory, these patients would not have access to somatic markers signalling the various alternatives and their possible consequences (Bechara, 2004; Tranel & Damasio, 2000). Proof of this is the fact that patients with lesions in the ventromedial PFC do not perform well in the IGT and they do not develop anticipatory electrodermal responses to disadvantageous decisions (Bechara, Damasio, Tranel & Anderson, 1998). Instead, these patients experience somatic reactions to the consequences of the decisions they take, i.e. before losses or gains, as do normal subjects, although their responses are of lower intensity. Their lesions prevent or interfere in their using the somatic signals set in motion by the amygdala, the hypothalamus and the brain stem nuclei, which are integrated into future decisions and are necessary for correct decisions to be taken. The impression is that the patient with a ventromedial lesion, and with no problems of memory, does not take into account earlier experience from decisions. When the new decision making scenario arises, the patient does not show anticipatory somatic reactions, which lead them to take disadvantageous decisions.

Data from patients with ventromedial frontal lesions are confirmed by practically all researchers when the lesions are general and involve the orbital prefrontal cortex. Rogers, Everith et al. (1999) report, moreover, an increase in the deliberation time which is probably common to this type of lesions. Elsewhere, Manes et al (2002) found that patients with extensive, diffuse frontal lesions are similar to those with lesions restricted to the ventromedial PFC – poor performance of the IGT, with a clear preference towards the riskier options and a tendency to bet larger amounts in the search for higher gains. In patients with focal lesions, the problems do not appear to be so marked. Lateral lesions affect the working

memory and the executive functions. Focal orbitofrontal lesions do not produce a significant deterioration in performance of the IGT, where the patients return a level near to that of the control subjects, although they do take more time to respond (Manes et al., 2002). When patients are grouped according to criteria similar to that of Damasio and colleagues, those showing the greatest problems in taking decisions are those with a lesion to the orbitofrontal region whose damage extends to other areas. Krawczyk (2002) points out that the choices of disadvantageous decks which appear from the outset of the IGT in patients with frontal lesions may indicate a lack of flexibility.

## 4.1.2. Asymmetries in the Ventromedial Prefrontal Cortex.

A certain hemispheric specialization exists in the processing of emotional information which some authors have related to decision making. According to Davidson (2003) studies on hemispheric lateralization and emotions indicate that:

- 1. The right hemisphere shows superiority in recognizing emotional information as well as in regulating mood and affect.
- 2. Both recognition and regulation of emotion is bilateral, although the left hemisphere is more specialized in positive emotions and the right one in negative ones.

Nevertheless, Davidson himself (2003) prefers to talk about approach behaviours, governed by the anterior left hemisphere, and withdrawal behaviours, linked to the anterior right hemisphere. The supposed dominance of the right hemisphere for emotions would not be so, but rather for emotional perception and activation, and referred to posterior regions. According to Davidson, the left PFC would be more active during positive moods and would be involved in appetitive goals, related to approach behaviours. On the other hand, the right PFC would show higher activity during negative emotions and would be related to behavioural inhibition and the increased arousal of some negative affects. The imbalance in favour of the right side would be proper to depressive states. This asymmetry is more manifest in negative emotions than in positive ones. Depressive patients exhibit less left prefrontal activation than people who are not depressed (Davidson, 2002; 2003). The data obtained from fMRI show that the right orbitofrontal cortex is activated more as a reaction to punishment and during defence and withdrawal responses, while the left region is activated more in response to reward and approach to the stimulus (Clark, Manes, Antoun, Sahakian & Robbins, 2003; Rogers, Owen et al., 1999; Tranel, Bechara & Denburg, 2002). Elsewhere, Ernst et al. (2002) associate the right lateralization to affective processing, but also to the inhibitory processes. However, not all studies have found this asymmetry. Breiter et al. (2001), for example, report that in gambling tasks the right PFC is related to positive affect, while the left one is related to negative affect.

The ventromedial lesion in the right hemisphere provokes a worse execution in decision making than does the left ventromedial lesion (Tranel et al., 2002). An effect of the size of the lesion is also observed. The more the lesion extends to other areas away from the ventromedial PFC, the greater is the tendency to make disadvantageous decisions and the worse is the performance in the IGT (Clark et al., 2003).

To summarize, most research indicates that the right ventromedial PFC is more involved than the left one in social behaviour, in emotions and in decision making. It does not seem to be activated in tasks requiring working memory, unless the uncertainty present demands an additional effort. Its possible function is related to the assessment of the affective consequences or the behavioural significance of the choice (Elliott et al., 1999).

#### 4.1.3. Conclusions: Ventromedial Prefrontal Cortex in Decision Making

The review of fMRI studies by Kringelbach and Rolls (2004) underscores the role of the orbitofrontal cortex in processing and representing rewards. They agree with the previous authors that this cortical region does not only assist in evaluating the rewards but also in making quick adjustments to changes in the value of the reinforcement. Processing emotional value from the stimuli is more plastic and flexible in the orbitofrontal cortex than in the amygdala. The authors distinguish two functional axes of the orbitofrontal cortex:

- 1. *Medial-lateral axis*. The medial portion would be involved in processing rewards or gains, while the lateral portion would take part in processing punishments or losses, which might lead to changes in the current behaviour.
- 2. *Anterior-posterior axis*. Representation of abstract reinforcers, e.g. money, would take place in the anterior zones. This would be related to the complexity of the stimulus. Primary reinforcers, e.g. taste or pain, would be represented in posterior zones.

The orbitofrontal cortex therefore processes different types of rewards and makes quick changes to the behaviour in order to adapt to environmental changes (Krawczyk, 2002): this is the *adjustment* and the *flexibility* of the behaviour. This region is involved in decisions based on rules and on self-generated information. Krawczyk (2002) and Kringelbach and Rolls (2004) agree that the lateral sector of the orbital prefrontal region is more associated to punishment and to behaviour inhibition. However, data from Northoff et al. (2000) show that the medial PFC processes negative affective stimuli, while the lateral responds to positive stimuli, although more slowly and with less intensity. Elliott, Friston et al. (2000) point out that the lateral region would possess a more inhibitory character and would annul the effects of earlier rewards. This region is closely related to the amygdala and to the dopaminergic brain reward system. While the amygdala is more involved in the relation between stimuli and affective attributes, the ventromedial PFC integrates emotional information and information from other sources, and then decides from a variety of response options on the basis of the affect they generate.

## 4.2. Dorsomedial and Dorsolateral Prefrontal Cortex

The dorsal regions of the PFC play an essential role in the working memory and in executive functions (Smith & Jonides, 1999). Many of these functions assist in decision making – maintaining goals over time, the capacity to recover and hold information for a limited period of time and the mobilization and assignment of resources. The more information that exists to be processed in a task, the more this region is activated (Goldman-Racik, 1992; Gray et al., 2002; MacDonald et al., 2000; Rorie & Newsome, 2005). Although these functions are closely related to decision making, they are in fact different processes.

Paulus et al. (2001) found activation of the frontoparietal and limbic structures in situations of uncertainty and they propose that a neural frontoparietal system is activated

during these, to which regions of the prefrontal (BA 10) and parietal cortices (BA 7, 40) belong. All of them would be associated to executive functions. When changes in strategy occur the parietal and temporal lobes are activated as are the prefrontal regions and the cingulate cortex, although to a lesser extent. Similar results were reported by Critchley et al. (2001) who found that uncertainty during the wait for a reinforcement activates the orbital and medial prefrontal cortices along with the right temporoparietal region. As the uncertainty increases the activation of the orbitofrontal cortex becomes bilateral and the ACC is also activated.

The dorsolateral PFC lesion affects decision making, working memory and other executive functions. These patients make more disadvantageous decisions in the IGT, but not in simpler tests derived from it (Manes et al., 2002), some studies show that patients with wide and diffused lesions perform badly in the IGT, preferring riskier decisions (Clark et al., 2003). Bechara and colleagues report that patients with dorsolateral prefrontal lesion have problems in decision making, but these are secondary and are due to problems in executive functions, especially in the working memory. The effect is more pronounced in lesions to the right hemisphere. These lesions therefore affect decision making indirectly since they disrupt the working memory (Bechara et al., 1999; 2005).

The right ventrolateral and dorsolateral prefrontal cortices are mainly involved in the use of mnesic information (Krawczyk, 2002). Ambiguity activates the right dorsolateral PFC, while the left one participates more in simpler decisions with fewer options. Patients with right dorsomedial frontal lesions show bad performance in delayed tasks, which points to a deficit in working memory and a deficient execution of decision making tasks, in which their performance is similar to that of patients with ventromedial PFC lesions (Clark et al., 2003; Manes et al., 2002). Bechara et al. (1998), however, found that these patients do the IGT well, but with a low normal score.

To summarize, the dorsolateral PFC would form a part of the cognitive control system, together with the parietal cortex, the dorsal ACC and, in part, the temporal cortex, which analyses and manipulates the relevant information for deliberation and for conscious decision, and increases its activity in situations of uncertainty. This area has the capacity to recruit different regions of the brain and becomes more important when various sources of information have to be considered. It is activated more under pressures of time and it assists conscious thinking. It contributes to decision making where its role is, as we shall see below, more one of inhibition than of integration of emotion.

# **5. EMOTIONAL CONTROL**

#### 5.1. Performance in Cognitive Tasks by Inhibition of Emotion

The PFC can regulate other structures, e.g. the amygdala, and modify the intensity of emotions. *Emotional control* techniques habitually seek to diminish negative emotions, often by means of cognitive strategies which mentally transform the meaning of an emotional situation. At the experimental level, the subject is normally required to perform tasks that need a continuous cognitive effort, so the irruption or presentation of emotional stimuli interferes with the performance of the task and leads to a worse result, which means that tasks

cannot be carried out without a playing down of the emotional reactions. In general, the presentation of an emotional stimulus, e.g. fearful images, while carrying out cognitive tasks results in a lower activation of the amygdala.

Emotional control is manifested in studies with fMRI by an *increase* in the activation of the PFC, especially the lateral sector, associated to the working memory, and by a *decrease* in the activity of the amygdala. Ochsner et al. (2004) show that the dorsal portion of the ACC and the lateral portions of the PFC are related, respectively, to the processes of cognitive control and behavioural inhibition. The control also activates the ACC, depending on the emotional self-control strategy used. Apparently, the PFC actively intervenes in the alternative interpretation of an event, which allows us play down the negative emotion associated to it. Beer et al. (2006) find that the lateral inferior frontal cortex (BA 47) may participate both in the inhibition of emotional information and in its incorporation into decision making. These authors are of the opinion that the region would form part of a more extensive neural system of emotional control. Along these lines, De Martino et al. (2006) found that the greater the "framing" effect, whether positive or negative, the greater the activation of the amygdala; when the "framing" effect is lower, there is greater activity of the orbital and medial PFC and greater deliberation and rationality in the decisions. In this line, Hare, Tottenham, Davidson, Glover & Casey (2005) found that performing a cognitive task reduces the response of the amygdala to emotional stimuli.

Under Damasio's *somatic marker theory*, somatic states, sensory information and earlier experiences are integrated in the ventromedial PFC with information coming from the amygdala, the hypothalamus and other brain stem nuclei. The ventromedial PFC exercises its influence on the autonomic and motor activity through circuits running to the amygdala, the hypothalamus and the striatum, and from there to the brain stem nuclei. Thus, the ventromedial PFC is able to exert an inhibitory control, especially through their lateral portions. But other structures also participate in controlling emotion. Bush et al. (2000) and Shafritz, Collins & Blumberg (2006) give special importance to the ACC, which may exert the emotional control as a result of the balance between the "cognitive", dorsal and the "affective", rostral-ventral ACC.

#### 5.2. Preference for Long Term Rewards

Another type of situation where inhibition of emotional behaviour and the predominance of rational calculation are observed is in tasks where the subject chooses between short and long term rewards Subjects have to check their impulses towards an immediate gain and choose a larger, long term gain. Such tasks clearly show the "rational" control of emotional information and how emotions arising from the anticipation of short term gains are inhibited.

McClure, Laibson, Loewenstein & Cohen (2004) reported the activation of two separate brain systems, according to whether the subject chooses monetary rewards over the short or long term. In short term choices, the activation appears in sectors of the limbic system associated to the mesolimbic dopaminergic system, including the paralimbic cortex– the ventral striatum, the medial orbitofrontal cortex and the medial PFC. These are limbic structures which are strongly innervated by the brain stem dopaminergic system and are related to reinforcement (see section 3.2), to impulsive behaviour and addictions. These structures may make up the system which prevails in impulsive and disinhibited behaviour and in drug addicts, with hypersensitivity towards immediate reward and insensitivity towards future rewards. (Breiter et al.,2001). In contrast, when the preference is for monetary rewards with a six-week delay, the lateral PFC and the posterior parietal cortex are activated more. Under normal circumstances competition would arise between the two systems. The proximity of a reward would lead to context dependent impulsive reactions, while a more distant reward would activate regions committed to calculation and deliberation (McClure et al., 2004).

These long term choices are associated to abstract rewards and the balance between immediate and future benefits. Long term reward tends to activate the PFC, even more so in the parietal regions, and forms part of the frontoparietal system of executive control associated to cognitive activity and reasoning and to assessing future expectations and long term consequences of behaviour.

Bechara and Damasio (2005) point out that the brain mechanisms participating in decision making differ depending on the degree of uncertainty in the task and the delay in the reward. These authors agree with the proposal by Kringelbach and Rolls (2004) of an anterior-posterior axis in the orbitofrontal cortex related to choices with delayed and immediate reward. According to Bechara and Damasio (2005), the near future seems to be processed by the posterior ventromedial sector of the PFC, while the more anterior ventromedial sector would be involved in long term decisions. Bilateral lesions to the ventromedial PFC which affect only the anterior regions lead to bad decisions, "future myopia", and are related to the bad results in the IGT in a subgroup of people who consume abuse substances. The preference for delayed reward indicates intervention by the lateral PFC and the parietal posterior cortex, which are both related to executive functions and rational calculation. The posterior ventromedial PFC, which includes the ACC, tends to dominate the processing and use of concrete, and more predictable information, and also that related to specific and tangible reinforcers. Its activity is context-dependent, and has greater sensitivity to emotional aspects and to aversive stimuli. It would be related to the steeper evaluation slope for losses than for gains described in Tversky and Kahneman's prospect theory (see section 2.1).

The anterior ventromedial PFC, on the other hand, tends to dominate the processing and use of abstract and symbolic information which is less predictable, risky, more context independent and with less emotional modulation.

# 5.3. Cooperation, Altruism and Reciprocity in Economic Games

A growing body of research coming from Neuroeconomics, the study of the neural bases of economic decisions, shows the relevance of emotion in economic games. Economic decision making games between two or more players are characterized by usually being zerosum and iterative, which means that successive actions by the different players occur which have opposite consequences, losses or gains, for both. Rational theory proposes that a subject will always choose the most favourable option and maximize the expected utility. Yet the opposite predominates with the frequent appearance of non rational behaviours, of *cooperation*, which lead to more balanced gains for both players, and prevail over the immediate gain strategy. Matched or unmatched cooperation is a constant in social interactions. It brings with it altruistic behaviours – giving before receiving – in the expectation that the short term benefits which they forego will be rewarded in the long term. Games involving cooperation frequently require one to control the tendency to go for an immediate reward and to prefer a future one which is more advantageous to both parties. The people contrast what they expect to receive over the short and long term and subjectively calculate the cost of the cooperation. In trust and reciprocity games, the latter is indicated in the putting off of an immediate reward, in understanding the moods of the other players and the circumstances and contingencies which can lead to a bigger reward by applying cooperative behaviours.

In the Ultimatum game, which is played between two players, one player has 10 monetary units to share out. There is only one turn, so if the second player accepts the offer, the money is shared as proposed by the first player, whereas if the offer is refused, neither player receives any money. Rational decision making would lead player one to making a low offer and to player two accepting it. However, most offers are of about 50% of the amount since the first player foregoes the chance to take advantage and offers an equal share out. In fact, 50% of low offers (9:1; 8:2), which are usually about 20% of all the offers, are turned down. The second player does not want to be taken advantage of, even though this means prejudicing himself. The structures activated during the task are the anterior insula (bilaterally), the ACC and the dorsolateral PFC (Camerer, 2003; Sanfey et al., 2003). The dorsolateral PFC would assist in maintaining and representing the information necessary for the task. Its activation does not predict the behaviour in the choice, but rather reflects the calculation and maximization of the monetary reward. For the authors in question, cooperation is related to PFC activity. As seen earlier (section 3.1.1), low offers are often accompanied in the receiving player by activity of the insula and by negative emotion. The ventromedial PFC also plays a role in this task due to the fact that the ventromedial lesions cause a higher number of refusals and exaggerated reactions when low offers are received (Koenigs & Tranel, 2007). These lesions lead to disorders in the regulation of emotion, in particular in the integration of emotion into decision making.

In the iterated Prisoner's Dilemma game, subjects also tend to show cooperative behaviours. Cooperation activates structures in the brain reward system, in particular the anteroventral striatum, in NAc, in the caudate and in the ventromedial/orbital PFC and the rostral-ventral ACC. The hypothesis by Rilling et al. (2002) is that the rewarding nature of cooperation, derived from the activation of the brain reward system, allows for altruistic behaviour. Cooperation is also associated to greater subjective satisfaction and has reinforcing effects, although cooperation tends to fall off in the last rounds of the game.

According to King-Casas and colleagues (2005), who studied reciprocity in the Trust game, reciprocity also activates the dorsal striatum. In this game, the first player decides to transfer a part of the amount he receives to an investment, where it is tripled. The second player then takes control of the investment and can divide it as he *or she* chooses between the two. The first player usually puts a lot of money, proportionally, into the investment, thus renouncing an immediate profit and not prejudicing the second player. The second player corresponds by sharing out an important amount of the profit obtained from the investment. For King-Casas et al. (2005), activation of these structures is equivalent to a surprise effect, since there is a deviation from the expected behaviour – that of the rational, neutral investor – with the detection of a change in reciprocity. The intention of trusting, that is, the anticipation of a future reinforcement, leads to a greater activation of the anterior and medial cingulate

cortices (possibly related to uncertainty). Activation of the cingulate cortex is associated in social games with knowledge of another player's decision.

In the Trust game, Tomlin et al. (2006) found activation of the middle cingulate region for the decision itself and of the anterior and posterior cingulate region when the decisions of others are known. Deviations in reciprocity activate error and uncertainty signals which are processed in the dorsal ACC. McCabe, Houser, Ryan, Smith & Trouard (2001) found greater activation in the medial PFC in more cooperative subjects.

To summarize, the PFC and the ACC would play an important role in attributing moods to other players and in cooperation. The appearance of cooperative behaviours brings with it the activation of the dorsolateral PFC and the anterior pole of the frontal cortex, which are involved, respectively, in executive functions, rational control and calculations, and in long term rewards. Cooperation is accompanied by frontal cortex activity and activity of the brain reward system. Cooperation and altruism act as reinforcers. They activate the brain reward system and are associated to positive moods. Be it on account of evolutionary factors or experience, cooperation is a dominant tendency in social interactions which has led many authors, e.g. Camerer (2003) to refer to it as *social utility* which often provides long term benefits. It should be observed that if there is a conflict between the two response tendencies – altruism and selfishness – the brain defence system is activated, in particular the anterior insula and the rostral-ventral ACC.

# 5.4. Incorporation of Emotion in Decision Making: Moral Judgements

One special type of decision making with greater or lesser emotional involvement is moral decisions or judgements. These consist of valuing actions by others according to one's own norms and values and those of the social group. The moral dilemma is enhanced as the emotion aroused increases (Greene, Sommerville, Nystrom, Darley & Cohen, 2001).

Purely cognitive processes participate in impersonal moral judgements. During these deliberations, the brain areas most activated are those related to abstract thinking and problem solving, in particular the dorsolateral PFC and, in general, the cognitive control frontoparietal system (Greene, Nystrom, Engell, Darley & Cohen, 2004). If the issue to be judged has an emotional content (usually negative), the medial orbitofrontal cortex is also activated (Moll, de Oliveira-Souza, Bramati & Grafman, 2002). In comparison with moral dilemmas where there is no personal implication, Greene et al. (2001) found that this type of judgements activate regions of the brain associated to emotion - medial frontal gyrus (BA 9, 10), posterior cingulate cortex (BA 31) and bilateral angular gyrus (BA 39). Subsequent studies have included the superior temporal sulcus and different areas of the temporal cortex (Greene et al., 2004). Heekeren et al. (2005) argue for a neural substrate which is common to moral decision making and consisting of the ventromedial/ orbitofrontal PFC, the right posterior cingulate cortex, the posterior superior temporal sulcus, and the temporal pole. It is important to point out that in their experiment the authors do not find any involvement of the amygdala, nor of the posterior cingulate cortex to sentences containing expressions of physical harm, which may be due to the way the stimuli were presented (in this case, listening to sentences). The amygdala might be activated more by visual stimuli.

The moral dilemma increases when the tendency to adopt a utilitarian, or rational, decision is up against one guided by emotional aspects. The difficult personal dilemma goes

against usual moral principles and leads to a situation of conflict. If the final decision taken is utilitarian, cognitive control prevails, and there is activation of those regions associated to the working memory and abstract reasoning, in particular of the dorsolateral PFC, the bilateral inferior parietal lobe and the temporal gyri. In this situation reaction times are longer and the ACC and the lateral amygdala are activated (Greene et al., 2004). The predominance of the utilitarian option is, at the neural level, related to the preference towards delayed reward, with inhibition of emotions, as was seen in section 5.2.

It is especially important to confirm the role of the ventromedial PFC in the incorporation of emotion and its inhibition in moral dilemmas. When exposed to moral dilemmas, both with and without personal implication, patients with lesions to the ventromedial PFC show a high number of excessively rational, "utilitarian" *j*udgements, i.e. their responses to moral dilemmas tend to be impersonal. Logical reasoning is maintained and explicit moral norms persist in the patients, but when the moral dilemma is heavily charged with emotions of the type, "Would you kill your own child to save ten other lives?" these patients show excessively rational decisions. Lesions to the ventromedial PFC would seem to impede the incorporation of emotional considerations to the decisions (Koenigs et al., 2007). It often occurs in moral decisions that those which may be socially more acceptable are the ones that take into account affective aspects. The same occurs in iterative games, in which cooperation is generally the best strategy and emotions are also present in the decision. In both cases the concept of social utility mentioned would be applicable, as would the involvement of the medial PFC observed, at the neural level.

Impersonal moral dilemmas can apparently be solved by the intervention of the frontoparietal and temporal cognitive control system. On the other hand, consideration of emotional aspects requires participation by the brain defence system, the amygdala, ACC, and also activation of the ventromedial PFC.

# 5.5. Loss of Emotional Control: Drug Addiction and Decision Making

Other data as to which mechanisms participate in decision making come from studies of patients of chronic abuse of substances, in whom decision making is affected, with a preference for short term benefits and an ignorance of long term ones (Clark & Robbins, 2002). This situation is similar to that of patients with ventromedial PFC lesions, who are sensitive to immediate consequences but not to future ones. When they perform the IGT, drug addicts as a whole stand half way between the ventromedial patients and the control subjects (Bechara, 2003). They present a bimodal distribution – those scoring low are similar to the patients with ventromedial PFC lesions, with no anticipatory somatic responses to the disadvantageous decks, while those who score high are similar to the control subjects and show anticipatory somatic responses to the disadvantageous decks. In the group which performs the task badly, two thirds responded exaggeratedly to the reward, e.g. hypersensitivity to the reward or insensitivity to the punishment. This could be due to frontoparietal, insula and secondary sensory cortex anomalies. Insensitivity towards future consequences of decisions, be they positive or negative, seems to be similar to what occurs in ventromedial PFC patients, i.e. related to damage or hypoactivity in this region (Bechara, 2003; Bechara et al., 2001; Verdejo, Aguilar de Arcos & Pérez-García, 2004).

But drug addicts and patients with prefrontal lesions may perform tasks in which risk is properly rewarded better than the control subjects (Shiv, Loewenstein & Bechara, 2005). These authors, along with Bechara et al. (2001), point out that there are two processes that may arise in addiction:

- 1. Abnormal activity in the brain defence system, and hence of the amygdala, which exaggerates the processing of incentives associated to the consumption of substances and provokes hypersensitivity towards the reward. It would also affect the NAc, the bed nucleus of the stria terminalis and the mesolimbic dopaminergic system.
- 2. Abnormal activity in the ventromedial PFC system (including the ventral striatum) which would lower the value of future negative consequences of substance consumption. This causes a loss of inhibition, and a behaviour associated with immediate rewards.

Amphetamine addicts and patients with lesions to the orbitofrontal cortex do not perform well in decision making tasks. Moreover, the effect is more pronounced the longer the addiction has lasted. For Bechara (2003), there may exist a conflict within the drug addict between the primary reinforcer, the drug, and the secondary reinforcers, such as the anticipation of the negative consequences of its consumption.

# **6.** CONCLUSIONS

In all decision making tasks consequences are anticipated, which to a greater or lesser extent, have an affective value for the subject. This means that there is an overlap of the emotion processing systems with those involved in decision making. Different structures with different functions are participating. Current research allows us to propose that both anticipation and attainment of the consequences of a decision activate distributed brain systems which process punishments and rewards, and so cognitive and emotional information are represented and incorporated in different structures.

The defence brain system is activated when the emotion is negative, with the presence or the anticipation of punishments or in situations of great uncertainty. The main structures consistently involved are the amygdala, the anterior insula, the ACC and the lateral PFC. The anterior insula and the ACC can be considered as cortical regions which process and integrate emotional information, with the anterior insula being related to internal sensation and states, while the ACC takes charge of integrating cognitive and emotional information. Both are dependent on a subsequent, more elaborate, process in the ventromedial/orbital PFC, which selects behaviours and inhibits or promotes responses through these structures, through the basal ganglia and through the thalamus. An initial integration of emotion and cognition is made through in the ACC. This is activated especially in situations of conflict and uncertainty and is related both to the ventromedial/orbital PFC, as regards emotion and to the dorsolateral PFC, as regards behaviour inhibition and selection of responses. Elsewhere, the brain reward system is activated by the presence or anticipation of reinforcing stimuli and their associated signals. There is a common brain circuit for reward which runs from the dopaminergic brain stem nuclei to the ventral and dorsal striatum, especially the NAc and to the caudate nucleus, including, although to a lesser extent, the amygdala. The greater the reward, the more structures are activated, including the caudate.

The ventromedial PFC, and in particular the orbital part, is involved in both the integration and the inhibition of emotional behaviour and in a reciprocal manner with the dorsolateral PFC, amygdala, insula and ACC. The representation of rewards and punishments makes it ready to act flexibly towards risk and uncertainty and it is important when calculating the consequences of an action and in anticipating changes in reinforcement contingencies and in other contextual circumstances which may arise. The greater the emotion, risk and uncertainty, the more it will be activated. It receives inputs from the amygdala, the anterior insula and the rostral-ventral portion of the ACC, thus allowing it to integrate and assess the value of the incentive of the expected consequences of the choices, together with the internal states, and select a final behaviour. The amygdala and the ventromedial/orbital PFC assist decision making in different ways. The ventromedial PFC would integrate the somatic states with the subject's previous experience, including those emotions processed in the amygdala. The more lateral areas are more related to behaviour inhibition and to the frontoparietal system related to executive control.

The dorsolateral PFC is directly related to the superior parietal cortex, to the temporal cortex and to the dorsal ACC. It participates in executive functions like working memory and attentional control, as well as in mobilizing cognitive resources, which increase in situations of uncertainty where there is the need to recover and use both contextual information and information from long term memory. It is also activated as the emotion increases but, together with the ventromedial PFC, it can, when the situation requires, inhibit emotions. The basal ganglia and the thalamus would be a part of the efferent pathway of this inhibitory circuit.

The activity of these systems is sensitive to uncertainty, since its increase leads to more deliberation and mental effort, which is related to a higher level of activity not only in the orbitofrontal cortex, but also in the executive frontoparietal system, including the dorsal ACC. An excess of uncertainty or risk will also activate the defence system, with the participation of the amygdala, anterior insula and rostral-ventral ACC, particularly in the right hemisphere.

There exists some competition between the two cerebral reward subsystems. When the reward is expected at short term, there will be more activity in the limbic areas and the basal ganglia. When the reward is long term, the ventromedial PFC and the frontoparietal system of executive control, which may include the temporal lobe, are activated more. In this respect, Bechara (2005) proposes two different interacting systems which would intervene in decision making, and which do not work properly in people with lesions to the ventromedial PFC and in some persons addicted to drugs -an impulsive system or one from the amygdala, which indicates pleasure or pain as an immediate result of the possible options, and a reflexive system based in the ventromedial PFC and sensitive to future consequences set off by the options themselves. The preference for the delayed reward activates the dorsolateral PFC and, in general, the frontoparietal system. On the other hand, cooperation in iterative games, associated to the aim of a greater delayed reward or the satisfaction derived from the cooperation itself, is characterized by the involvement of the PFC, which seems to participate in the anticipation or interpretation of the mental states of the other players – often referred to as "theory of mind"- and in the activation of the brain reward system.

Several hypotheses can be put forward to account for how anticipated and experienced emotions influence the "framing" effect, i.e. when the subject, on anticipating the consequences of his decisions, opts for the higher risk choices in a loss frame and for more conservative ones in a gain frame. Experiments by De Martino et al. (2006) refer to participation of the amygdala in both cases. The absence of the "framing" effect is related to a greater activation of the orbital and medial PFC, indicating cognitive control of the situation.

The role of the brain mechanisms governing the incorporation of a mood, resulting from losses or gains, or from other sources outside the task, has been explored by several researchers. Positive affect states, resulting from an earlier immediate reward, anticipate riskier behaviours and are associated to activation of the NAc, which belongs to the brain reward system (Kuhnen and Knutson, 2005). In contrast, a loss result activates the anterior insula, in the defence system, and leads to risk aversion in the next trial (Paulus et al., 2003). When emotions are not related to the task, the affective carry-over effect spreads out at the moment the decision is taken, and these same systems may be activated.

As opposed to an interpretation with certain Freudian reminiscences that the cerebral cortex exerts a function of rational control on the subcortical emotional systems, the reality is the image of a PFC with several subdivisions and multiple functions. One of these is precisely that of incorporating feelings into decisions to achieve an optimum fit of the individuals in their social sphere.

In concluding, we must underline the methodological problem posed by the diversity of the tasks in studies of this type. If we are to progress with guarantee, the tasks and behaviours used must be standardized and well known, even if this requires sacrificing ecological validity for the sake of rigorous methodology which will ensure a suitable control of the variables. Finally, neuroimaging techniques have important limitations on account of their correlational nature. The use of neuropsychological data from studies of patients with brain lesions is therefore important, as is a meticulous study and assessment of the behavioural variables in play.

# References

- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P, & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*, 68-72.
- Aggleton, J. P. & Young, A. W. (2000) The enigma of the amygdala: On its contribution to human emotion. In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 106-128). New York: Oxford University Press.
- Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E. & Breiter, H. C. (2001) Beautiful faces have variable reward value: fMRI and behavioural evidence. *Neuron*, 32, 537-551.
- Baumeister, R. F. (2003) The psychology of irrationality: Why people make foolish, selfdefeating choices. In I. Brocas & J. D. Carrillo (Eds.) *The psychology of economic decisions. Vol. I, Rationality and well-being* (pp. 3-16). Oxford: Oxford University Press.
- Bechara, A. (2003) Risky business: Emotion, decision-making, and addiction. *Journal of Gambling Studies*, 19, 23-51.
- Bechara, A. (2004) The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition*, 55, 30-40.

- Bechara, A. (2005) Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. *Nature Neuroscience*, *8*, 1458-1463.
- Bechara, A. & Damasio, A. R. (2005) The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52*, 336-372.
- Bechara, A. & Damasio, H. (1997) Deciding advantageously before knowing the advantageous strategy. *Science*, 275, 1293-1295.
- Bechara, A, Damasio, H., Damasio, A. R. & Lee, S. W. (1999) Different contributions of the human amygdala and ventromedial prefrontal cortex to decision making. *Journal of Neuroscience*, 19, 5473-5481.
- Bechara, A., Damasio, H., Tranel, D. & Anderson, S. W. (1998) Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience*, 18, 428-437.
- Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. (2005) The Iowa Gambling Task and the somatic marker hypothesis: Some questions and answers. *Trends in Cognitive Science*, 9, 159-162.
- Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S. W. & Nathan, P. E. (2001) Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia*, 39, 376-389.
- Bechara, A. & Van der Linden, M. (2005) Decision-making and impulse control after frontal lobe injuries. *Current Opinion in Neurobiology*, 18, 734-739.
- Beer, J.S., Knight, R. T. & D'Esposito, M. D. (2006) Controlling the integration of emotion and cognition. *Psychological Science*, 17, 448-453.
- Berridge, K. C. (2003) Irrational pursuits: Hyper-incentives from a visceral brain. In I. Brocas & J. D. Carrillo (Eds.) *The psychology of economic decisions. Vol. I, Rationality and well-being* (pp. 17-40). Oxford: Oxford University Press.
- Breiter, H. C., Aharon, I., Kahneman, D., Dale, A. & Shizgal, P. (2001) Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, 30, 619-639.
- Breiter, H. C., Gollub, R. L., Weiskoff, R. M., Kennedy, D. N., Makris, N., Berke, J. D., Goodman, J. M., Kantor, H. L., Gastfriend, D. R., Riorden, J. P., Mathew, R. T., Rosen, B. R. & Hyman, S. E. (1997) Acute effects of cocaine on human brain activity and emotion. *Neuron*, 19, 591-611.
- Buchanan, T. W., Tranel, D. & Adolphs, R. (2004). Anteromedial temporal lobe damage blocks startle modulation by fear and disgust. *Behavioral Neuroscience*, 118, 429.437.
- Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C. & Gabrieli, J. D. E. (2002) Dissociable contributions of prefrontal and parietal cortices to response selection. *Neuroimage*, 17, 1562-1571.
- Bush, G., Luu, P. & Posner, M. I. (2000) Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, 4, 215-222.
- Camerer, C. F. (2003) Strategizing in the brain. Science, 300, 1673-1675.
- Carr, L., Iacobini, M., Dubeau, M. C., Mazziotta, J. C. & Lenzi, G. L. (2003) Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences USA*, 100, 5497-5502.
- Clark, L., Manes, F., Antoun, N., Sahakian, B. J. & Robbins, T. W. (2003) The contributions of lesion laterality and lesion volume to decision-making within the human prefrontal cortex. *Neuropsychologia*, 41, 1474-1483.

- Clark, L. & Robbins, T. W. (2002) Decision-making deficits in drug addiction. Trends in Cognitive Sciences, 6, 361-363.
- Cohen, M. X., Heller, A. S. & Ranganath, C. (2005) Functional connectivity with anterior cingulate and orbitofrontal cortices during decision-making. *Brain Research. Cognitive Brain Research*, 23, 61-70.
- Critchley, H. D., Mathias, C. J. & Dolan, R. J. (2001) Neural activity in the human brain relating to uncertainity and arousal during anticipation. *Neuron*, *29*, 537-545.
- Critchley, H. D., Rotshtein, P., Nagai, Y., O'Doherty, J., Mathias, C. J. & Dolan, R. J. (2005) Activity in the human brain predicting differential heart rate responses to emotional facial expressions. *Neuroimage*, 24, 751-762.
- Crone, E. A., Somsen, R. J. M., Van Beek, B. & Van der Molen, M. W. (2004) Heart rate and skin conductance analysis of antecedents and consequences of decision making. *Psychophysiology*, 41, 531-540.
- Damasio, A. R. (1994) *Descartes' error: Emotion, reason, and the human brain*. New York: Avon Books.
- Davidson, R. J. (2002). Anxiety and affective style: role of the prefrontal cortex and amygdala. *Biological Psychiatry*, 51, 68-80.
- Davidson, R. J. (2003). Affective neuroscience and psychophysiology: Towards a síntesis. *Psychophysiology*, 40, 655-665.
- Davidson, R. J. & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences*, 3, 11-21.
- Davidson, R. J., Maxwell, J. S. & Shackman, A. J. (2004) The privileged status of emotion in the brain. Proceedings of the National Academy of Sciences USA, 101, 11915-11916.
- Davis M. (1996) Differential role of the amygdala and bed nucleus of the stria terminalis in conditioned fear and startle enhanced by corticotrophine-releasing hormone. In T. Ono, B. L. McNaughton, S. Molotchnikoff, E. T. Rolls, H. Nishijo (Eds.) *Perception, memory and emotion: Frontiers in neuroscience* (pp. 525-548). Oxford: Elsevier.
- Davis, M. (1997) The neurophysiological basis of acoustic startle modulation: Research on fear, motivation, and sensory gating. In P. J. Lang, R. F. Simons & M. Balaban (Eds.) *Attention and orienting: Sensory and motivational processes* (pp. 69-96). Mahwah, New Jersey: Lawrence Erlbaum Associates; 1997.
- Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C. & Fiez, J. A. (2000) Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, 84, 3072-3077.
- De Martino, B., Kumaran, D., Seymour, B. & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*, 684-687.
- Dolan, R. J. & Morris, J. S. (2000) The functional anatomy of innate and acquired fear: perspectives from neuroimaging. In R. D. Lane & L. Nadel (Eds.) Cognitive neuroscience of emotion (pp. 225-241). New York: Oxford University Press.
- Elliott, R., Dolan, R. J. & Frith, C. D. (2000) Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cerebral Cortex*, 10, 308-317.
- Elliott, R., Friston, K. J. & Dolan, R. J. (2000) Dissociable neural responses in human reward systems. *The Journal of Neuroscience*, 20, 6159-6165.
- Elliott, R., Rees, G. & Dolan, R. J. (1999) Ventromedial prefrontal cortex mediates guessing. *Neuropsychologia*, 37, 403-411.

- Emery, N. J. & Amaral, D. G. (2000) The role of the amygdala in primate social cognition. In R. D. Lane & L. Nadel (Eds.) *Cognitive neuroscience of emotion* (pp. 156-191). New York: Oxford University Press.
- Ernst, M., Bolla, K., Mouratidis, M., Contoreggi, C., Matochik, J. A., Kurian, V. S., Cadet, J.-L., Kimes, A. S. & London, E. D. (2002) Decision-making in a risk-taking task: A PET study. *Neuropsychopharmacology*, 26, 682-691.
- Ernst, M., Kimes, A. S., London, E. D., Matochik, J. A., Eldreth, D., Tata, S., Contoreggi, C., Leff, M. & Bolla, K. (2003) Neural substrates of decision making in adults with attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 160, 1061-1070.
- Gehring, W. J. & Willoughby, A. R. (2002) The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, 295, 2279-2282.
- Glimcher, P. W. y Rustichini, A. (2004) Neuroeconomics: The consilience of brain and decisión. *Science*, 306, 447-452.
- Goldman-Rakic, P. S. (1992) Working memory and the mind. *Scientific American*, 267, 111-117.
- Gray, J. R. (1999) A bias towards short-term thinking in threat-related negative emotional states. *Personality and Social Psychology Bulletin*, 25, 65-75.
- Gray, J. R., Braver, T. S. & Raichle, M. E. (2002) Integration of emotion and cognition in the lateral prefrontal cortex. *Proceedings of the National Academy of Sciences USA*, 99, 4115-4120.
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M. & Cohen, J. D. (2004). The neural bses of cognitive conflict and control in moral judgement. *Neuron*, 44, 389-400.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M. & Cohen, J. D. (2001) An fMRI investigation of emotional engagement in moral judgement. *Science*, 293, 2105-2108.
- Hajcak, G., Holroyd, C. B., Moser, J. S. & Simons R. F. (2005) Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*, 161-170.
- Hare, T. A., Tottenham, N., Davidson, M. C., Glover, G. H. & Casey, B. J. (2005) Contributions of amygdala and striatal activity in emotion regulation. *Biological Psychiatry*, 57, 624-632.
- Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., Imamizu, H. & Kawato, M. (2004) A neural correlate of reward-based behavioral learning in caudate nucleus: A functional magnetic resonance imaging study of a stochastic decision task. *The Journal of Neuroscience*, 24, 1660-1665.
- Heekeren, H. R., Wartenburger, J., Schmidt, H., Prehn, K., Schwintoski, H. P. & Villringer, A. (2005) Influence of bodily harm on neural correlates of semantic and moral decision making. *Neuroimage*, 24, 887-897.
- Hinson, J. M., Jameson, T. L. & Whitney, P. (2002) Somatic markers, working memory and decision making. *Cognitive, Affective, & Behavioral Neuroscience*, 2, 341-353.
- Hinson, J. M, Whitney, P., Holben, H. & Wirick, A. K. (2006) Affective biasing of choices in gambling task decision making. *Cognitive, Affective, & Behavioral Neuroscience*, 6, 190-200.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D. & Camerer, C. F. (2005) Neural systems responding to degrees of uncertainty in human decision-making. *Science*, 310, 1680-1683.

- Isen, A. M. (1993) Positive affect and decision making. In M. Lewis & J. M. Haviland (Eds.) Handbook of emotions (pp. 261-277). New York: Guilford Press.
- Ishai, A., Pessoa, L., Bickle, P. C. & Ungerleider, L-G. (2004) Repetition suppression of faces is modulated by emotion. *Proceedings of the Nacional Academy of Sciences USA*, 101, 9827-9832.
- Johnson, E. y Tversky, A. (1983) Affect, generalization, and the perception of risk. Journal of Personality and Social Psychology, 45, 20-31.
- Kahneman, D. (2003) Experienced utility and objective happiness: A moment-based approach. In I. Brocas & J.D. Carrillo (Eds.) *The psychology of economic decisions. Vol. I, Rationality and well-being* (pp. 187-208). Oxford: Oxford University Press.
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R. & Montague, P. R. (2005) Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, 308, 78-83.
- Knutson, B., Adams, C. M., Fong, G. W. & Hommer, D. (2001) Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience*, 21, RC159 (1-5).
- Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., Hommer, D. (2003) A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: Characterization with rapid event-related fMRI. *Neuroimage*,18, 263-272.
- Knutson B, Rick, S., Wimmer, G. E., Prelec, D. & Loewenstein, G. (2007) Neural predictors of purchases. *Neuron*, 53, 147-156.
- Knutson, B., Taylor, J., Kaufman, M., Peterson, R. & Glover, G. (2005) Distributed neural representation of expected value. *The Journal of Neuroscience*, 25, 4806-4812.
- Knutson, B., Westdorp, A., Kaiser, E. &Hommer, D. (2000) FMRI Visualization of brain activity during a monetary incentive delay task. *Neuroimage*, 12, 20-27.
- Koenigs, M. & Tranel, D. (2007) Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the Ultimatum game. *The Journal of Neuroscience*, 27, 951-956.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M. & Damasio, A. R. (2007) Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, 446, 908-911.
- Krawczyk, D. C. (2002). Contributions of the prefrontal cortex to the neural basis of human decision making. *Neuroscience and Biobehavioral Reviews*, *26*, 631-664.
- Kringelbach, M. & Rolls, E. T. (2004) The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, 72, 341-372.
- Kuhnen, C. M. & Knutson, B. (2005) The neural basis of financial risk taking. *Neuron*, 47, 763-770.
- Kuniecki, M., Urbanik, A., Sobiecka, B., Kozub, J., Binder, M. (2003) Central control of heart rate changes during visual affective processing as revealed by fMRI. Acta Neurobiologiae Experimentalis, 63, 39-48.
- Le Doux, J. E. (1996) The emotional brain. New York: Simon & Schuster.
- LeDoux, J. E. (2000) Cognitive-emotional interactions. In R.D. Lane & L. Nadel (Eds.) Cognitive neuroscience of emotion (pp. 129-155) New York: Oxford University Press; 2000.

- Lerner, J., Small, D. A. & Loewenstein, G. (2004) Heart strings and purse strings: Carryover effects of emotions on economic decisions. *Psychological Science*, 15, 337-341.
- Loewenstein, G. F., Weber, E. U., Hsee, C. K. & Welch, N. (2001) Risk as feelings. Psychological Bulletin, 127, 267-286.
- Mac Donald, A. W., Cohen, J. D., Stenger, V. A. & Carter, C.S. (2000) Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288, 1835-1838.
- Maia, T. & McClelland, J. L. (2004) A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences USA*,101, 16705-16080.
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M. & Robbins, T. (2002) Decision-making processes following damage to the prefrontal cortex. *Brain*,125, 624-639.
- Martínez Selva, J. M., Sánchez-Navarro, J. P., Bechara, A. & Román, F. (2006) Mecanismos cerebrales de la toma de decisiones [Brain mechanisms involved in decision making]. *Revista de Neurología*, 42, 411-418.
- McCabe, K., Houser, D., Ryan, L, Smith, V. & Trouard, T. (2001) A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Acadeny of Sciences USA*, 98, 11832-11835.
- McClure, S. M., Berns, G. S. & Montague, P. (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, 38, 339-346.
- McClure, S. M., Laibson, D. I., Loewenstein, G. & Cohen, J. D. (2004) Separate neural systems value immediate and delayed monetary rewards. *Science*, 306, 503-507.
- Mellers, B. A. (2000) Choice and the relative pleasure of consequences. *Psychological Bulletin*, *126*, 910-924.
- Milham, M. P., Banich, M. T., Claus, E. D. & Cohen, N. J. (2003) Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. *Neuroimage*, 18, 483-493.
- Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R. & Gafman, J: (2006) Human fronto-mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences USA*, 103, 15623-15628.
- Moll, J., de Oliveira-Souza, R., Bramati, I. E. & Grafman, J. (2002) Functional networks in emotional moral and nonmoral social judgements. *Neuroimage*, *16*, 696-703.
- Northoff, G., Richter, A., Gessner, M., Schlagenhauf, F., Fell, J., Baumgart, F., Kaulisch, T., Kötter, R., Stephan, K. E., Leschinger, A., Hagner, T., Bargel, B., Witzel, T., Hinrisch, H., Bogerts, B., Scheich, H. & Heinze, H. J. (2000) Functional dissociation between medial and lateral prefrontal cortical spatiotemporal activation in negative and positive emotions: A combined fMRI/MEG study. *Cerebral Cortex*, 10, 93-107.
- Ochsner, K. ., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D. E. & Gross, J. J, (2004). For better or for worse: Neural systems supporting the cognitive down- and up-regulation of negative emotion. *Neuroimage*, 23, 483-499.
- O'Doherty, J., Deichmann, R., Critchley, H. D. & Dolan, R. J. (2002) Neural responses during anticipation of a primary taste reward. *Neuron*, *33*, 815-826.
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J. & Andrews, C. (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, *4*, 95-102.

- Öhman, A. (1993) Fear and anxiety as emotional phenomena: clinical phenomenology, evolutionary perspectives, and information-processing mechanisms. In M. Lewis & J.M. Haviland (Eds.) *Handbook of emotions* (pp. 511-536). New York: Guilford Press.
- Pagnoni, G., Zink, C.F., Montague, P. R. & Berns, G. S. (2002) Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*, 5, 97-98.
- Paulus, M. P., Hozack, N., Zauscher, B., McDowell, J. E., Frank, L., Brown, G. G. & Braff, D. L. (2001) Prefrontal, parietal, and temporal cortex networks underlie decision-making in the presence of uncertainity. *Neuroimage*, 13, 91-100.
- Paulus, M. P., Rogalsky, C., Simmons, A., Feinstein, J. S. & Stein, M. B. (2003) Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *Neuroimage*, 19, 1439-1448.
- Pessoa, L. & Padmala, S. (2005). Quantitative prediction of perceptual decisions during nearthresholds fear detection. *Proceedings of the National Academy of Sciences USA*, 102, 5612-5617.
- Peters, E. & Slovic, P. (2000) The springs of action: affective and analitical information processing in choice. *Personality and Social Psychology Bulletin*, *26*, 1465-1475.
- Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S. & Kilts, C. D. (2002) A neural basis for social cooperation. *Neuron*, 35, 395-405.
- Rogers, R. D., Everith, B. J., Baldacchino, A., Blackshaw, A. J., Swainson, R., Wynne, K., Baker, N. B., Hunter, J., Carthy, T., Booker, E., London, M., Deakin, J. F. W., Sahakian, B. J. & Robbins, T. W. (1999) Dissociable deficits in the decisión-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: Evidence for monoaminergic mechanisms. *Neuropsychopharmacology*, 20, 322-339.
- Rogers, R. D., Owen, A. M., Middleton, H. C., Williams, E. J., Pickard, J. D., Sahakian, B. J. & Robbins, T. W. (1999) Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *The Journal of Neuroscience*, 20, 9029-9038.
- Rogers, R. D., Ramnani, N., Mackay, C., Wilson, J. L., Jezzard, P., Carter, C. S. & Smith, S.M. (2004) Distinct portions of anterior cingulated cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological Psychiatry*, 55, 594-602.
- Rolls, E. T. (2000) The orbitofrontal cortex and reward. Cerebral Cortex, 10, 284-294.
- Rorie, A. E. & Newsome, W. T. (2005) A general mechanism for decision-making in the human brain? *Trends in Cognitive Sciences*, *9*, 41-43.
- Rottenstreich, Y. & Hsee, C. K. (2001) Money, kisses, and electric shocks: On the affective psychology of risks. *Psychological Science*, *12*, 185-190.
- Sabatinelli, D., Bradley, M. M., Fitzsimmons, J. R. & Lang, P. J. (2005). Parallel amygdala and inferotemporal activation reflect emotional intensity and fear relevance. *Neuroimage*, 24, 1265-1270.
- Sánchez-Navarro, J. P. & Martínez-Selva, J. M. (2006) Emotion-induced changes in decision making. *Journal of Psychophysiology*, 20, 248.
- Sánchez Navarro, J. P., Martínez Selva, J. M. & Román, F. (2006). Uncovering the relationship between defence and orienting in emotion: Cardiac reactivity to unpleasant pictures. *International Journal of Psychophysiology*, 61, 34-46.

- Sanfey, A. G., Loewenstein, G., McClure, S. M. & Cohen, J. (2006) Neuroeconomics: crosscurrents in research on decision-making. *Trends in Cognitive Sciences*, 10, 108-116.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E. & Cohen, J. D. (2003) The neural basis of economic decision-making in the Ultimatum game. *Science*, 300, 1755-1758.
- Schultz, W. (2004) Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Current Opinion in Neurobiology*, 14, 139-147.
- Seymour, B., O'Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K. J. & Frackowiak, R. S. (2004) Temporal difference models describe higherorder learning in humans. *Nature*, 429, 664-667.
- Shafir, E. & Le Boeuf, R. A. (2002) Rationality. Annual Review of Psychology, 53, 491-517.
- Shafritz, K. M., Collins, S. H. & Blumberg, H. P. (2006) The interaction of emotional and cognitive neural systems in emotionally guided response inhibition. *Neuroimage*, 31, 468-475.
- Shiv, B., Loewenstein, G. & Bechara, A. (2005) The dark side of emotion in decisionmaking: When individuals with decreased emotional reactions make more advantageous decisions. *Cognitive Brain Research*, 23, 85-92.
- Shizgal, P. & Arvanitogiannis, A. (2003) Gambling on dopamine. Science, 299, 1856-1858.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J. & Frith, C. D. (2004) Empathy for pain involves the affective but not sensory components of pain. *Science*, 303, 1157-1162.
- Smith, E. E. & Jonides, J. (1999) Storage and executive processes in the frontal lobe. *Science*, 283, 1657-1661.
- Suzuki, A., Hirota, A., Takasawa, N. & Shigemasu, K. (2003) Application of the somatic marker hypothesis to individual differences in decision making. *Biological Psychology*, 65, 81-88.
- Tomb, J., Hauser, M., Deldin, P. & Caramazza, A. (2002) Do somatic markers mediate decisions on the gambling task? *Nature Neuroscience*, 5, 1103-1104.
- Tomlin, D., Kayali, M. A., King-Casas, B., Anen, C., Camerer, C. F., Quartz, S. R. & Montague, P. R. (2006) Agent-specific responses in the cingulated cortex during economic exchanges. *Science*, *312*, 1047-1050
- Tranel D. (2000). Electrodermal activity in cognitive neuroscience: Neuroanatomical and physiological correlates. In Lane, R.D. & Nadel, L. (Eds.) Cognitive neuroscience of emotion. (pp. 192-224). Nueva York: Oxford University Press.
- Tranel, D. & Damasio, A. R. (2000) Neuropsychology and Behavioral Neurology. In J.T. Cacioppo, L.G. Tassinary & G.G. Berntson (Eds.) *Handbook of Psychophysiology 2nd Edition* (pp 119-141). New York: Cambridge University Press.
- Tranel, D., Bechara, A. & Denburg, N. L. (2002) Asymmetric functional roles of right and left ventromedial prefrontal cortices in social conduct, decision-making, and emotional processing. *Cortex*, 38, 589-612.
- Tversky, A. y Kahneman, D. (1974) Judgement under uncertainity. Science, 185, 1124-1131.
- Tversky, A. y Kahneman, D. (1981) The framing of decisions and the psychology of choice. Science, 211, 453-458.
- Valdesolo, P. & DeSteno, D. (2006) Manipulations of emotional context shape moral judgement. *Psychological Science*, 17, 476-477.

- Verdejo, A., Aguilar de Arcos, F. & Pérez-García, M. (2004) Alteraciones de los procesos de toma decisiones vinculados al córtex prefrontal ventromedial en pacientes drogodependientes [Alterations to the decision-making processes linked to the ventromedial prefrontal cortex in drug-abusing patients]. *Revista de Neurología*, 38, 601-606.
- Volkow, N. D., Fowler, J. S. & Wang, G-J. (2004) The addicted human brain viewed in the light of imaging studies: Brain circuits and treatment strategies. *Neuropharmacology*, 47, 3-13.
- Volkow, N. D., Wang, G-J., Fowler, J. S., Logan, J., Gatley, J., Wong, C., Hitzemann, R. & Pappas, N. R. (1999) Reinforcing effects of psychostimulants in humans are associated with increases in brain dopamine and occupancy of D<sub>2</sub> receptors. *The Journal of Pharmacology and Experimental Therapeutics*, 291, 409-415.
- Von Neumann, J. & Morgenstern, O (1947) *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton Uiversity Press.
- Wheatley, T. & Haidt, J. (2005) Hypnotic disgust makes moral judgements more severe. *Psychological Science*, *16*, 780-784.
- Winkielman, P., Berridge, K. C. & Wilbarger, J. L. (2005) Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgements of value. *Personality and Social Psychology Bulletin*, 31, 121-135.

# ACKNOWLEDGMENT

We thank the financial support of the Spanish Ministry of Education and Science, Grant SEJ2004-06062.

Chapter 10

# A NEUROSCIENTIFIC APPROACH TO THE PSYCHOLOGY OF DECISION MAKING<sup>\*</sup>

# Michael X. Cohen

Department of Epileptology and Center for Life and Brain, University of Bonn, Germany Department of Psychology, University of California, Davis, California, USA

# ABSTRACT

Considerable research in cognitive psychology, economics, and marketing has focused on the psychological mechanisms of decision-making. Coming from a different perspective, the past few decades have seen considerable research on the neurobiological mechanisms of reward processing, learning, and addiction. Can these neuroscience findings shed light on the psychology of decision-making? Here I review the neuroscience literature, ranging from physiological studies of single neurons in rats to noninvasive neuroimaging in humans, as it relates to and informs human reward-guided decision-making. I focus specifically on the nucleus accumbens, orbitofrontal cortex, and hippocampus, structures with critical roles in the flexible adaptation of decision-making behavior based on rewards and novelty. The major neurochemical system involved in reward-guided decision-making is dopamine, which originates in the midbrain, is transmitted throughout the striatum and frontal cortex, and provides a global signal about rewards and errors in reward prediction. Dysfunctions of the dopamine system might underlie impairments in decision-making such as impulsivity and irrational decisionmaking, as seen in, for example, substance abuse. These structures do not operate in isolation, but rather form a tightly linked network, in which information is passed from structure to structure, and is dynamically gated in the nucleus accumbens, such that incoming information is permitted throughput only during specific neural events. Consideration of these systems as functional circuits has the potential to shed light on the neurobiological processes that underlie psychological decision making, as well as impairments in decision-making, such as those seen in addiction disorders.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Medicine and Health Care*, edited by Thomas E. Lynch, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

# INTRODUCTION

Psychology encompasses many processes spanning the gamut of behavior, from low level processes such as visual perception to high level processes such as adapting to the rules of a new culture. Indeed, merely glancing at the chapter titles of an introductory psychology textbook reveals the wide range of topics covered under the general term psychology. Presumably, all psychological processes are rooted in the complex chemical, electrical, and mechanical interactions between neural tissues, which simultaneously occur and interact at molecular, synaptic, neural, and systems levels. It seems a daunting task to explain such complex psychological processes as adolescent group interactions at the neurobiological mechanistic level. Nonetheless, the study of the neural bases of psychological processes is at its infancy, especially when compared to other sciences such as physics and biology, which date back thousands of years and yet continue to make new developments and discoveries today. It is not feasible to provide an overview of the neuroscience of all aspects of the psychology of decision-making; I will therefore focus on specific aspects related to the psychological processes of learning about rewards, orienting behavior towards potentially rewarding situations, and biasing action selection towards obtaining future rewards. The neurobiological and neurocomputational mechanisms that might drive these psychological processes are beginning to be mapped out, and this chapter will provide an overview of the structures involved, their functions, and their interactions at the systems level.

There are several other aspects of decision-making that are being actively investigated that will not be covered here. For example, the neural processes underlying perceptual decision-making (e.g., is this turquoise color more blue or green?) are beginning to be mapped out. To the extent that similar neural mechanisms might underlie many kinds of decisions, this research has the potential to shed light on more complex psychological aspects of decision-making. Interested readers are referred to the works of Shadlen, Newsome, Romo, among others, for reviews and empirical studies related to the neuroscience of perceptual decision making (Deco and Rolls 2006; Gold and Shadlen 2002; Horwitz et al 2004; Leon and Shadlen 1998; Romo et al 2006; Sugrue et al 2005). Readers are also referred to the work conducted in the lab of Paul Glimcher, in which the neuroscience of simple economic decision-making is investigated (Glimcher 2002; Glimcher 2004; Glimcher et al 2005). These experiments also shed light into the psychology of decision-making, but are not reviewed extensively here.

# **OVERVIEW OF THE CHAPTER**

One could argue that the primary function of all neural tissue is to transform incoming information into an output signal – that is, to weigh evidence and make a decision accordingly. This implies that the entire brain is involved in some kind of decision-making. However, here I focus on brain circuits that are most relevant to psychological aspects of goal-directed, reward-seeking behavior. In particular, I will selectively focus on a small subset of limbic structures—the nucleus accumbens, orbitofrontal cortex, hippocampus, and midbrain dopamine system—without implying that other structures or neurochemical systems are not important for goal-directed decision-making.

First, I review the anatomical properties of this subset of brain structures. No understanding of the brain can be complete without knowledge of at least the basic anatomical connectional properties. Indeed, new theories and experiments are often inspired based on observed anatomical connectivity among brain regions. Next I review the basic functions of these brain regions, drawing from evidence in animals and humans that have used lesion, single cell recording, and neuroimaging techniques. Finally, I will attempt to integrate these regions into larger functional circuits that might underlie certain aspects of human decision-making. Clearly there will be some overlap between all sections.

# I. FUNCTIONAL ANATOMY OF STRUCTURES INVOLVED IN REWARD-GUIDED DECISION-MAKING

The Nucleus Accumbens: A Gateway Between Emotion and Action

The nucleus accumbens is a small structure located deep inside the brain, in the ventral part of the striatum. The rat nucleus accumbens comprises two major subregions, the shell and the core. The shell is a crescent-shaped piece that lies more ventromedial, and the core appears like a ball, resting on top of the shell. In the rat, these subregions can be functionally dissociated: Whereas the core is widely implicated in choice behavior, impulsivity, and delayed reinforcements, the shell is implicated in assessments of the hedonic value of rewards and in unlearned reward-seeking behavior such as feeding and locomotion (Berridge 2003; Cardinal 2006; Cardinal et al 2004; Di Chiara 2002; Robbins and Everitt 1996). However, such functional dissociations have not been demonstrated in humans, and, indeed, even anatomical dissociations based on histochemical studies have revealed that the human nucleus accumbens is more functionally and anatomically homogeneous than is the rat accumbens. Thus, I will refer to the human nucleus accumbens as a single functional unit.

Mogenson proposed that the nucleus accumbens acts as a "gateway" between emotion and action (Mogenson et al 1980; Mogenson and Yang 1991). That is, the accumbens might receive information about possible rewards and possible actions, and help facilitate specific actions that might lead to specific rewards. Mogenson developed this hypothesis based on the anatomical properties of the nucleus accumbens: Its major inputs come from the amygdala, orbitofrontal and medial frontal cortices, hippocampus, and midbrain dopamine structures including the ventral tegmental area (VTA) and substantia nigra, pars compacta (SNc). In turn, the accumbens projects to output structures of the basal ganglia associated with motor and homeostatic regulation, including the globus pallidus, external segment, hypothalamus, and substantia nigra, pars reticulata (SNr) (Haber et al 2000; Haber et al 2006; Haber and McFarland 1999). The SNr exerts an inhibitory control over the SNc, and thus the accumbens can activate the SNc and induce dopamine release via inhibiting the SNr's inhibitory control. Since Mogenson's proposal, considerable work has been conducted to refine our understanding of the anatomical connectivity of the primate nucleus accumbens, although his original proposal remains widely accepted and influential.

The nucleus accumbens comprises mainly medium spiny neurons (MSNs), so named because they contain many synaptic boutons. About 95% of the neurons in the nucleus accumbens are MSNs. Connections from multiple inputs synapse onto the same MSNs (Finch 1996; French and Totterdell 2002). That is, incoming information from, for example, the hippocampus, prefrontal cortex, and midbrain dopamine system can arrive at the same accumbens neurons. Thus, the accumbens is ideally situated to integrate information from many different sources. The remainder of the neurons in the accumbens are termed tonically active interneurons.

New findings about the electrophysiological properties of the nucleus accumbens confirms its role as a gateway. In particular, neurons in the nucleus accumbens oscillate between two bistable electrical states: A depolarized "up" state, in which incoming information from, for example, the prefrontal cortex, can cause action potentials, and a hyperpolarized "down" state, in which incoming information generally does not cause action potentials (Brady et al 2005; Kasanetz et al 2006; Lewis and O'Donnell 2000; Murer et al 2002; O'Donnell and Grace 1995; O'Donnell et al 1999; Wilson and Kawaguchi 1996). The terms hyperpolarized and depolarized refer to the membrane potential of the neuron as being relatively more or less negative, respectively. These up-down states oscillate with a frequency of around .5-1 Hz (i.e., once every 1-2 seconds), and create a functional gateway: When accumbens neurons are in their down state, incoming information does not "pass through" into motor output structures of the basal ganglia. However, when accumbens neurons are in their up state, incoming information can easily induce action potentials, which are then passed on, perhaps in the service of biasing or reinforcing actions associated with rewards. Does this mean that the accumbens converts emotion into action only once every 2 seconds? No. Certain input signals can force accumbens neurons from a down state into an up state. For example, electrical stimulation of the hippocampus or medial frontal cortex induces up states in accumbens neurons (Kasanetz et al 2006; O'Donnell and Grace 1995). In this way, certain brain structures can "open the gates" to allow activity from other brain structures to pass through. Although these up-down fluctuations typically are recorded intracellularly, they can also be observed when recording extracellular potentials, demonstrating that populations of nucleus accumbens neurons fluctuate in synchrony (Goto and O'Donnell 2001; Leung and Yim 1993). Interestingly, up-down cycles are not observed in single slice preparations. That is, when a slice of living rat brain containing the nucleus accumbens is placed on a dish, those neurons will remain in a perpetual down state (Surmeier and Kitai 1993; Wilson 1993). Further, up states are not observed in decorticated rats (Wilson 1993; Wilson et al 1983). This supports the notion that down states are maintained by local neural processes whereas up states are achieved by a bombardment of input from afferent structures such as the hippocampus or medial frontal cortex (Wolf et al 2005).

# Nucleus Accumbens And Reward-Guided Decision-Making

The involvement of nucleus accumbens in reward processes is demonstrated by copious reports showing increases in accumbens activity and dopamine efflux following various kinds

of rewards, ranging from primary (e.g., food or sex) to symbolic (e.g., money that will not actually be received), in species from rats to humans (Berridge 2003; Breiter et al 2001; Cardinal 2006; Cohen et al 2005; Ernst et al 2005; Knutson et al 2001; Knutson and Cooper 2005; Komisaruk and Whipple 2005; Osaka and Osaka 2005; Robbins and Everitt 1996). Studies have also implicated the accumbens in delayed rewards and risky decision-making (Huettel 2006; Kuhnen and Knutson 2005; Matthews et al 2004). Indeed, lesions to the nucleus accumbens lead rats to prefer small but immediate rewards over large but delayed rewards (Cardinal and Howes 2005; Cardinal et al 2001; Cardinal et al 2004). Consistent with a role in representing future rewards, studies have demonstrated significant increases in nucleus accumbens activity during delays prior to receiving rewards (Knutson et al 2001; Knutson and Cooper 2005; Risinger et al 2005). Nucleus accumbens activity also increases following administration of cocaine, which is perceived as rewarding (Breiter et al 1997; Breiter and Rosen 1999), as well as to attractive faces, which are perceived as rewarding (Aharon et al 2001).

But does the nucleus accumbens simply respond to rewards? The nucleus accumbens also increases its activity following non-rewarding salient events, punishments, and painful stimuli (Aharon et al 2006; Becerra et al 2001; Zink et al 2003). These findings are inconsistent with a simple role of the accumbens in reward processing. Indeed, a consideration of its connectional and functional anatomy suggests that the accumbens may be involved in more than just helping to identify rewards. The nucleus accumbens receives convergent inputs from a range of cortical and subcortical structures that are involved in emotion, motivation, memory, novelty detection, and rule representation, and in turn has output fibers to basal ganglia structures involved in the regulation of actions and dopamine flow. These findings, in addition to the observed up-down gating signals, suggest that the accumbens may be especially involved in using reinforcement information to guide flexible goal-directed behavior (Cools et al 2007). That is, the nucleus accumbens might process incoming reward information in a way that promotes or biases those actions that are more likely to lead to future rewards. This would explain why nucleus accumbens lesions produce impulsive behavior: Appropriate "stop" signals are not sent to nucleus accumbens targets, thus preventing the biasing of actions that lead to larger rewards.

One concern of the observations of the gating mechanism of up-down cycles is that they typically are recorded while rats are under anesthesia, raising the question of whether this gating mechanism exists during waking behavior as well. This is further challenged by the observation that stimulation of medial frontal cortex at physiologically relevant frequencies disrupts nucleus accumbens up-down cycles, impelling them into a perpetual up state (Kasanetz et al 2002; Kasanetz et al 2006). These findings suggest that during goal-directed behavior, when medial frontal cortex as well as other inputs to the accumbens (e.g., hippocampus, which can also impel accumbens neurons into an up state) are likely to be highly active, slow up-down cycles might not be observed. One possibility is that although the slow up-down cycle may exist primarily during periods of relative inactivity, faster oscillations (e.g., around 10 Hz, or alpha) might have a similar functional mechanism. That is, the electrophysiological gating mechanism might be a general functional property of such neurons, and the precise frequency with which this mechanism operates might change in different situations. Consistent with this, cortical neurons are also most likely to emit action potentials during their up states, which have been observed in a broader frequency range (Stern et al 1997).

# Anatomy of the Orbitofrontal Cortex

The ventral aspect of the prefrontal cortex is termed orbitofrontal cortex, so named because of its proximity to the orbit of the eyes. Cytoarchitectonically, orbitofrontal cortex is defined as the region that receives input from the medial nucleus of the mediodorsal nucleus of the thalamus (Fuster 1997; Pandya and Yeterian 1996). The development of the orbitofrontal cortex is relatively recent in phylogenetic terms (Kringelbach and Rolls 2004). Even among primates, humans have a relatively larger orbitofrontal cortex than do other nonhuman primates. And within humans, the orbitofrontal cortex is the last brain structure to fully develop. Indeed, the orbitofrontal cortex is likely not fully developed until around age 18 (Gogtay et al 2004). Even in fully developed adults, there is considerable inter-individual variability in the anatomical organization (e.g., location of sulci). Given these findings, it is no surprise that the orbitofrontal cortex has been implicated in complex and high-level social decision-making functions (Kringelbach 2005).

The orbitofrontal cortex is a relatively large and multi-functional structure. Most conceptualizations of the orbitofrontal cortex distinguish between the medial and lateral aspects. This differentiation is supported by differential patterns of anatomical connectivity and functional properties. The medial orbitofrontal cortex has dense projections to the posterior cingulate and retrosplenial cortex, hippocampus, and ventral striatum including the nucleus accumbens (Ferry et al 2000; Ongur and Price 2000). The lateral orbitofrontal cortex, on the other hand, has projections to premotor and motor cortices, and sensory cortices including gustatory, somatic, and visual (Carmichael and Price 1994; Carmichael and Price 1995). However, the medial-lateral distinction is more a soft gradient than a strong divide: Many projections into the orbitofrontal cortex reach all sectors, notably inferior temporal visual areas, but also visceral and auditory inputs as well.

In addition to functional and anatomical distinctions along the medial-lateral dimension, there might also be an anterior-posterior gradient. Simple or concrete reinforcers such as taste or smell are represented in the posterior orbitofrontal cortex, and more complex or abstract reinforcers such as money are represented in the anterior orbitofrontal cortex (Kringelbach 2005).

# Functions: Orbitofrontal Cortex and Adaptive Decision-Making

The orbitofrontal cortex has a major role in forming and modifying associations between stimuli or actions and their outcomes, and using these associations to flexibly guide behavior. Regions of the orbitofrontal cortex, especially the posterior medial sections, are important for the representation of rewarding tastes, especially of sweet and fatty foods (Kringelbach et al 2003; Rolls 2005; Rolls 2006). The orbitofrontal cortex also becomes active to rewards such as money, even if the money will not actually be received (O'Doherty et al 2001). However, the orbitofrontal cortex also serves more decision-related functions than merely representing reward value; it houses response-reward and stimulus-reward associations, and seems to be able to modify those associations when the environmental contingencies change. These are more than just simple response-reward associations (i.e., habits), which can be formed by other structures as well, including the amygdala and striatum. The orbitofrontal cortex seems

to form more abstract and flexible associations that can be used dynamically to adjust behavior according to rewards and punishments (Frank and Claus 2006; Kringelbach 2005). For example, animals with lesions to the orbitofrontal cortex are able to distinguish between rewarding over nonrewarding stimuli, but are unable to adapt their behavior to changes in reward value when contingencies in the environment change (Baxter et al 2000; Izquierdo and Murray 2004). Humans with lesions to the orbitofrontal cortex are often described as being perseverative, impulsive, and impaired at inhibiting their behavior (Damasio et al 1985; Eslinger and Damasio 1985). These qualities can be conceptualized as an impairment in adjusting behavior in a flexible and dynamic manner. In particular, if orbitofrontal cortex cannot maintain the rewarding value and motivational significance of actions and goals, behavior would become rigid, inflexible, and driven largely by immediately available rewards.

To achieve this role, the orbitofrontal cortex must be able to maintain active representations of reinforcements and response-reinforcement or stimulus-reinforcement associations over time (Frank and Claus 2006; Kringelbach 2005; Kringelbach and Rolls 2004; Schoenbaum et al 2006; Schoenbaum et al 2003). This information can be used as a top-down biasing signal to influence nucleus accumbens activity during decision-making. This is consistent with findings that the orbitofrontal cortex becomes active during working memory tasks, when subjects must maintain information over the a brief delay, as well as during long-term memory encoding (Frey and Petrides 2000; Frey and Petrides 2002; Frey and Petrides 2003).

These studies demonstrate that activity in orbitofrontal cortex reflects a combination of actions and reward associations. If the orbitofrontal cortex is important for flexibly guiding behavior, its activity should reflect not only the current reward and behavior, but also how future choices are adjusted according to current outcomes. Recent fMRI work confirms that reward- and punishment-related activity in the orbitofrontal cortex reflects decisions to be made in the subsequent trials during decision-making tasks. For example, Cohen and Ranganath (2005a) had subjects choose between high-risk and low-risk gambles. They found that activity in the orbitofrontal cortex, among other regions, during one trial predicted the gamble chosen in the subsequent trial. Interestingly, they found that this relationship was moderated by individual differences in mathematically estimated learning parameters related to how local expectations of rewards changed following different kinds of gambles.

## Anatomy of the Hippocampus

The hippocampus is a "C"-shaped structure located in the medial temporal lobe. It contains several subregions, including the four CA fields, and the dentate gyrus. The latter wraps around the former fields in a "U"-like fashion. Surrounding the hippocampus ventrally is the entorhinal cortex. Information that is sent from the rest of the brain to the hippocampus enters through the entorhinal cortex (Swanson 1979; Swanson and Cowan 1979), although it has been argued that some cortical regions directly project to the hippocampus (Schwerdtfeger 1979) of the marmoset. In contrast, projections exiting the hippocampus eminate from the hippocampus itself without traversing through the entorhinal cortex. For example, the CA1 field projects to regions in the prefrontal cortex and ventral striatum

including nucleus accumbens (Ferino et al 1987; Floresco et al 2001; Jay and Witter 1991; O'Donnell and Grace 1995; Schacter et al 1989; Swanson 1981; Yang and Mogenson 1985).

The hippocampus also receives dopaminergic projections from midbrain dopamine structures (Gasbarri et al 1994a; Gasbarri et al 1997; Gasbarri et al 1994b). Hippocampal dopamine may support memory formation and learning by facilitating long-term potentiation, which is thought to be a cellular mechanism of learning (Frey et al 1989; Frey et al 1991; Otmakhova and Lisman 1996).

## Hippocampus as Novelty Detector

The hippocampus has long been implicated in the formation of lasting episodic memories. Countless studies of humans, nonhuman primates, rats with hippocampus lesions, fMRI studies of healthy humans, single-unit recordings in animals, as well as computational models, have focused on the role of the hippocampus in the binding of sensory information into relatively permanently stored and accessible representations. Perhaps the most well-known study of the role of the hippocampus in memory is patient H.M., who had a bilateral anterior temporal lobectomy to treat epilepsy. Following the surgery, which removed the hippocampus as well as other surrounding structures, H.M. was unable to form new memories, despite a relatively intact ability to recall episodes from his life prior to his surgery, as well as intact intellectual and language abilities.

Memory formation is obviously critical to reward-guided decision-making. For example, one must remember the locations or situations in which particular actions or decisions led to particular positive or negative outcomes. The hippocampus also subserves other functions that are particularly relevant for reward-guided decision-making. In particular, the hippocampus might contain a novelty detection mechanism, by which it would respond to novel or salient features of the environment. Many lines of research, ranging from single unit recording studies in the rat to fMRI studies and lesion studies in humans (Kerr et al 1996; Kishiyama et al 2004; Knight and Nakada 1998; Strange et al 2005; Vinogradova 2001), support the claim that activity in the hippocampus increases in response to environmental novelty. How could the hippocampus detect novelty in the environment? Although the precise mechanism remains unknown, the following mechanism has been proposed (Hasselmo and Wyble 1997; Lisman and Grace 2005; Vinogradova 2001). The CA1 field receives input from the CA3, and also from high-level sensory areas such as inferior temporal cortex (via the entorhinal cortex). The former regions contain on-line mnemonic representations of the environment; the latter regions contain real-time representations of the current sensory world. CA1 can thus compare the actual information about the sensory environment with that expected from the CA3. Any difference between the actual and expected sensory world might be termed a "sensory prediction error," and is a basis for novelty. That is, a hippocampal novelty signal reflects the divergence of the sensory world from the mnemonic representation of the sensory world. Novelty is an important signal for orienting attention (Knight 1996; Laurens et al 2005; Sokolov 1990) and enhancing memory formation (Kishiyama et al 2004; Parker et al 1998). As discussed in a later section, this hippocampal-dependent novelty signal may be especially important in gating reward information in the nucleus accumbens and dopamine system (Lisman and Grace 2005).

#### 245

## Dopamine: The Tie That Binds

Dopamine is a molecule that is released into widespread brain regions from neurons originating in midbrain structures including the VTA and SNc. There are approximately 100,000 SNc neurons in each hemisphere of the macaque monkey (German et al 1988), and thus likely more in the human. Given there are approximately 31 million striatum neurons, and that the entire striatum is innervated by dopamine, each dopamine neuron projects to an average of over 300 striatum neurons, utilizing the approximately 500,000 varicosities per dopamine neuron from which dopamine is released into extra-synaptic space (Anden et al 1964; Anden et al 1965). These numbers are probably significantly higher in the human. Midbrain dopamine projections also reach widespread regions of the frontal cortex, hippocampus, and amygdala (Williams and Goldman-Rakic 1998). There are also dopamine projections to the parietal, occipital, and temporal cortices, although these projections are relatively sparse. When a dopamine neuron fires, dopamine is released into the synapse extremely rapidly, with a latency of around 40 s (0.00004 seconds). Peak dopamine concentration is achieved by around 6 ms, and within 75 ms, the dopamine concentration has reached both its temporal and spatial peak concentration (Schultz 1998). Dopamine is also rapidly reuptaken, and within around 200 ms, the local dopamine concentration returns to baseline levels. Thus, the dopamine signal is temporally precise but spatially diffuse, somewhat like a radio station's broadcast: The message is sent very quickly, in real-time, but is nearly equally available anywhere within range.

In order to receive the dopamine message, post-synaptic neurons are equipped with dopamine receptors. There are at least five classes of dopamine receptors (D1-D5), but in terms of reward and decision-making, most research has focused on D1 and D2 receptor subtypes. Receptors can be classified as high-affinity or low-affinity, meaning that they require a relatively smaller or larger concentration of dopamine, respectively, to induce post-synaptic changes. Most (~80%) dopamine receptors in the striatum are D1; the remainder are largely D2 (Richfield et al 1989). Of the D1 receptors, most (~80%) are low affinity. In contrast, most (~90%) D2 receptors are high affinity. This means that D2 receptors are highly sensitive to small concentrations of dopamine whereas D1 receptors generally require larger concentrations of dopamine to be activated. D1 and D2 receptors are usually (but not always) located on different neurons in the striatum (Richfield et al 1989). D1 receptors are found primarily on striatal neurons that project to the GPi and SNr, whereas D2 receptors are found primarily on neurons that project to the GPe.

## **Tonic Versus Phasic Dopamine Modes**

Dopamine has two modes of action: phasic and tonic. Phasic refers to actions that are rapid, occurring on a timescale of milliseconds; tonic refers to actions that are slower, occurring on a timescale of seconds to minutes or hours.

The phasic dopamine response was largely described above. In this mode, dopamine is rapidly released into the synapse immediately following action potentials in midbrain dopamine neurons. It is a very temporally precise signal. Although at the global level it is widespread in the sense that the dopamine signal is sent simultaneously to many neurons through the striatum, at the local level it appears to be quite focal. Indeed, dopamine released via the phasic mode normally does not leave the synaptic region into extra-synaptic space (Floresco et al 2003). However, some conditions, such as certain dopamine-activating drugs, appear to block this spatial restriction. This occurs because some drugs block dopamine reuptake receptors, thus allowing dopamine to flow outside the synapse.

# Phasic Dopamine and Reward Signals

As described previously, the phasic dopamine response is a rapidly onsetting, short-lived burst of dopamine that is simultaneously released in widespread regions of the limbic system. It occurs under a range of behaviorally significant events, primarily those that involve rewards or neutral stimuli that predict rewards (unconditioned and conditioned reinforcers, respectively) (Schultz 1998; Schultz 2006). However, the phasic dopamine response also occurs following events not associated with rewards, including environmental novelty and aversive stimuli such as electrical shocks (Bardo et al 1990; Guarraci and Kapp 1999; Hooks and Kalivas 1995; Strecker and Jacobs 1985). Therefore, although the phasic dopamine response is clearly linked to rewards and reward-related behaviors, it seems to convey other messages as well.

The most prominent and influential theory of the phasic dopamine response is the reward prediction error theory. This theory continues to drive novel research and has provided an exciting means of connecting neuroscience and computational learning mechanisms. The reward prediction error hypothesis states that encoded in the activity of midbrain dopamine neurons is a difference between a reward that is expected, based on previously learning associations that a particular action or stimulus precedes a particular reward, with a reward that is actually received. Thus, when an unexpected reward is received, dopamine neurons increase their firing rate; when expected rewards are received, dopamine neurons do not change their firing rate; and when expected rewards are not received, dopamine neurons decrease their firing rate. Increasing evidence obtained from recording action potentials of individual midbrain dopamine neurons supports this assertion (Bayer and Glimcher 2005; Daw and Doya 2006; Hollerman and Schultz 1998; Nakahara et al 2004; Schultz 1998; Schultz et al 1997).

Prediction errors can thus guide learning and decision-making by increasing the attention or salience to, and memory for, those actions or environments that lead to rewards. Dopamine also enhances both behavioral and cellular correlates of learning (Floresco et al 2001; Frey et al 1991), and decreases in dopamine can promote "unlearning," or decoupling of, for example, actions that were once associated with rewards but are now associated with punishments (Calabresi et al 1992; Ronesi and Lovinger 2005; Sajikumar and Frey 2004; Wolf et al 2004).

This theory is attractive not only for its simplicity and ability to account for many aspects of reward-guided behavior (Egelman et al 1998; Montague et al 2004; Schultz 1998), but also for its ability to tie together research in multiple disciplines. For example, the computational machine learning literature has long focused on how artificial systems might learn to optimize their behavior. One commonly used mechanism is teacher-guided learning, in which actions are selected based on a maximal expected "rewards" associated with each action, correct or

incorrect decisions are reinforced or not, and learning algorithms utilize prediction errors to update action selection tendencies (see Sutton and Barto 1998, for a review). The discovery that activity of dopamine neurons bears a strikingly similarity to prediction error terms in computational theories of learning has led to a marriage of computational learning mechanisms and neuroscience research, and often each field inspires novel hypotheses in the other (Cohen and Ranganath 2005b; Cohen and Ranganath 2007; McClure et al 2003; Montague et al 2004; O'Doherty J et al 2006; O'Doherty et al 2003). Excellent details of this theory are provided by Wolfram Schultz (Schultz 1998; Schultz 2001; Schultz 2006; Schultz and Dickinson 2000).

It is worth noting that this theory has been challenged (Berridge 2007; Redgrave and Gurney 2006; Redgrave et al 1999). Among the challenges are that the phasic dopamine response might be too homogenous and rapid to signal prediction errors in a range of situations outside those studied in the laboratory (for example, if it takes time to extract rewarding properties of a stimulus), and that genetically engineered rats who lack dopamine can still learn stimulus- and action-reward associations. Readers are referred to the Berridge and Gurney articles for an overview of challenges to this theory.

# Tonic Dopamine and Regulation

The tonic mode refers to a slower, more modulatory process. Whereas dopamine in the phasic mode is transmitted in rapid bursts that are quickly cleared out of the synapse and might not escape into extra-synaptic space, dopamine in the tonic mode is released slowly from extra-synaptic sites, has a more diffuse spatial extent, and is maintained homeostatically to have a relatively stable concentration. Whereas the phasic dopamine mode is critical for rapidly delivering information about rewarding or salient properties of stimuli or events in the environment, the tonic dopamine mode appears to be critical for regulating the information flow within the striatum. That is, tonic dopamine in the striatum can modulate the efficacy of incoming information from the hippocampus, amygdala, and prefrontal cortex. The functional implications of this distinction are discussed in later sections. Grace (1991) provides an excellent overview of the anatomical and functional properties of tonic vs. phasic dopamine as it relates to understanding schizophrenia, and some years later as it relates to substance abuse (Grace 2000).

In contrast to the phasic dopamine response, which continues to generate considerable empirical, computational, and theoretical research into its role in learning and decisionmaking, the tonic dopamine mode has received relatively little attention, despite its demonstrated importance in regulating the flow of information in the limbic system. Tonic dopamine refers to the global level of dopamine concentration in extrasynaptic space. It is fairly homogenously distributed throughout space, changes only slowly over time, and therefore represents a background level of dopamine. Tonic dopamine seems to be largely independent of phasic dopamine firing (except under certain circumstances such as particular drugs; Grace, 2000), and instead is regulated by a glutamatergic-mediated pre-synaptic mechanism under the control of the prefrontal cortex, and possibly other glutamatergic afferents as well. Thus, incoming signals from the prefrontal cortex can increase tonic levels of striatal dopamine. Tonic dopamine is thought to regulate the intensity of the phasic dopamine response in the striatum, in an inverse fashion. This occurs because high levels of tonic dopamine stimulate presynaptic autoreceptors, which in turn inhibit the release of dopamine that is driven by bursts of midbrain dopamine neuron activity (Grace 1991; Helmreich et al 1982; Lehmann et al 1983; Stoof et al 1982). Other compensatory mechanisms might also support the decreased phasic dopamine that results from increased tonic dopamine, including decreased receptor concentration or sensitivity (Frey et al 1987; Grace 1991; Huffman and Ticku 1983; Pan et al 1983). Thus, increased tonic dopamine, which is regulated in part by the prefrontal cortex, can dampen the phasic dopamine reward prediction error response. In this sense, tonic dopamine has been suggested to "shift the balance" between prefrontal and subcortical inputs into the striatum (Goto and Grace 2005b).

# **II. INTEGRATION**

No brain region is an island, and no individual can dynamically and actively engage in goal-directed, reward-seeking behavior with a brain containing, say, only a nucleus accumbens. Systems in the brain must operate in concert to produce meaningful behavior, and a full understanding of how neural circuits take in information, make decisions, and act upon them, will come only from considering how structures in the brain interact at the anatomical and functional levels. Indeed, clever lesion studies in animals demonstrate the necessity of neural intercommunication: By lesioning, for example, the left nucleus accumbens and the right orbitofrontal cortex, an animal can have a functioning accumbens and a functioning orbitofrontal cortex, but almost no means of communicating information between the two (Block et al 2006). Such studies demonstrate that to produce normal reward-directed behavior, it is insufficient to have functioning brain structures; these structures must be anatomically and, therefore, functionally, connected as well (for other examples, see: Gaffan and Murray 1990; Gaffan et al 1993; Izquierdo and Murray 2004).

These regions—the orbitofrontal cortex, nucleus accumbens, and hippocampus, strung together by dopamine—form an anatomically and functionally interconnected network. Only through interactions between these structures can flexible, adaptive, reward-guided decision-making arise. This circuit appears to serve several related functions.

One function, recently outlined by Lisman and Grace (2005), is to gate novel stimuli into long-term memory. It is quite established that information presented in novel contexts is better remembered than is information presented in more familiar environments. As an extreme example, if there were an elephant in your bedroom, you would probably remember that incident quite well; you might be less likely to remember a particular visit to the zoo in which you saw an elephant. This enhanced memory for contextually deviant information is sometimes referred to as the Von Restorff effect, after the psychologist who first described it (Von Restorff 1933). In a laboratory setting, the Von Restorff effect is typically induced by having subjects view a series of images. A majority of the images will be colored, for example, green, whereas a small minority of images will be colored red. When questioned later, the Von Restorff effect predicts that subjects' memory for the novel, red images will be better than their memory for the standard, green images. These novelty effects are increasingly being studied in cognitive psychology and cognitive neuroscience (Kishiyama et

al 2004; Knight 1996; Parker et al 1998), as well as in the study of memory for emotional events (Wiswede et al 2006).

What is the neural mechanism that allows novel information to be better encoded into memory? Lisman and Grace (2005) proposed that this occurs via the hippocampus—nucleus accumbens-midbrain dopamine system loop. First, the hippocampus detects environmental novelty by comparing its stored memory representations in the CA3 with incoming sensory information. Any discrepancies between the incoming information and the "expected" information based on memory patterns is detected and transmitted by the CA1 field to the nucleus accumbens. Hippocampal input to the accumbens can impel accumbens neurons into an up state, thus "opening the gates" to activate the substantia nigra, which in turn causes dopamine to be released into the hippocampus. This is important because dopamine in the hippocampus facilitates long-term potentiation, which might underpin memory formation (Frey et al 1989; Frey et al 1991; Otmakhova and Lisman 1996). Thus, the Von Restorff effect might occur by the hippocampus detecting novelty (the unusual red image among all the other green images), activating the nucleus accumbens, which in turn activates the dopamine system, allowing dopamine signals to enter the hippocampus and facilitate memory formation. This dopamine signal might also serve to orient attention in other regions of the brain.

Taking this hypothesis one step further, the orbitofrontal and other prefrontal cortical regions might fit into this loop as well. Specifically, when accumbens neurons are in their down state, they are not receptive to prefrontal cortical input. However, when hippocampal input drives accumbens neurons into an up state, they can become receptive to prefrontal input. In this way, the hippocampus might regulate the flow of information from the orbitofrontal cortex through the accumbens. Given the hippocampus' role in memory, this hippocampal regulation of prefrontal-accumbens throughput might manifest itself in guiding reward-seeking behavior based on learned reward associations. For example, if an animal learns that a particular decision leads to a particular reward under specific circumstances, the hippocampus might provide the critical learned contextual information into the accumbens that allows the accumbens to pass prefrontal cortical messages through to the basal ganglia. Consistent with this, action potentials in the hippocampus precede action potentials in the nucleus accumbens by around 10 ms (consistent with a direct connection from the hippocampus to the accumbens), and, importantly, these functional interactions increase when rats approach locations associated with rewards (Tabuchi et al 2003; Tabuchi et al 2000). Although these researchers did not additionally record field potentials of accumbens neurons to determine whether they were in functional up or down states, one possibility is that the increased hippocampal input to the nucleus accumbens enabled a functional up state, thus allowing prefrontal goal-directed representations to be converted into actions by the basal ganglia.

A breakdown of the nucleus accumbens' gating mechanism might contribute to psychiatric conditions linked with decision-making disorders, such as addiction. It is long established that drugs of abuse, including cocaine, nicotine, and alcohol, activate the nucleus accumbens (Grace 2000; Hurd et al 1999; Kalivas and Volkow 2005; Kelley 2004; Koob 1992; Volkow et al 2002; Wise 1984). In rats, long-term use of drugs leads to long-term changes in plasticity and dopamine receptor availability. In humans, abnormal activity in the nucleus accumbens in substance abusers relative to controls can be observed with a variety of methods. For example, PET imaging has shown that substance abusers, compared to healthy

controls, have lower levels of dopamine in the striatum (Martinez et al 2005; Volkow et al 2004). These dopamine levels are correlated with the severity of alcohol craving in abstinate alcoholics when viewing pictures associated with their addiction (e.g., beer bottles for alcoholics) (Heinz et al 2004). Additionally, fMRI evidence shows that substance abusers have abnormal activity in the ventral striatum, even when processing non-drug rewards such as money, suggesting the substance abuse may be associated with long-term changes in the nucleus accumbens that generalize outside the domain of substance abuse. Finally, other addictive disorders such as pathological gambling have also been associated with abnormal nucleus accumbens activity (Comings and Blum 2000; Reuter et al 2005; Tamminga and Nestler 2006). In terms of the gating hypothesis of the nucleus accumbens, drugs of abuse such as methamphetamine can induce unusually long periods of up states in nucleus accumbens neurons of anaesthetized rats (Brady et al 2005). Thus, an interesting possibility is that drugs of abuse cause impairments in decision-making at least partly by forcing nucleus accumbens neurons to remain in a constant up state, thus constantly allowing signals from reward structures in the brain to bias behavior.

The balance between tonic and phasic dopamine might also underlie some aspects of addiction (Grace 2000). Specifically, drugs of abuse likely cause their rewarding and reinforcing effects at least partly by increasing the phasic dopamine response in the nucleus accumbens (Grace 2000; Schultz 1998). During drug intake, signals from the hippocampus and prefrontal cortex-information about context and goals-converge in the accumbens (Finch 1996; French and Totterdell 2002), and due to enhanced plasticity from increased dopamine levels (Goto and Grace 2005a), these context/goal/reward/action associations become quickly solidified. In this way, associations between the reinforcing properties of drugs and the environments/actions that led to them are formed in a rapid and long-lasting manner. Long-term drug use also increases tonic dopamine levels in the nucleus accumbens (Grace, 2000). These increases in tonic dopamine are homeostatically regulated (and thus remain at high concentration for long periods of time), and inhibit the phasic response by inhibiting dopamine autoreceptors. This has two implications: (1) dysphoria might be induced when the chemical effects of the drug wear off because the reward-related phasic dopamine response is attenuated (Koob and Le Moal 1997), and (2) larger amounts of the drug will be required to obtain the level of phasic dopamine response that was previously rewarding (Grace 2000). Thus, some aspects of addiction can be accounted for by homeostatically regulated high levels of tonic dopamine. As stated above, high tonic dopamine reduces the phasic, reward-related dopamine response. Over time and repeated use, increased doses of the drug are required to increase phasic dopamine levels beyond the levels suppressed by tonic dopamine. This model can also account for withdrawal symptoms. Specifically, during withdrawal, exposure to situations or behaviors associated with the drug might activate glutamate-mediated increases in tonic dopmaine, which further suppress phasic dopamine, thus requiring the individual to seek out the drug again to compensate for decreased phasic dopamine levels (Grace, 2000).

One caveat of theories involving the up-down cycles of nucleus accumbens neurons is that these activities are often studied in anaesthetized rats. Therefore, we must await confirmation of the existence of such functions in awake, behaving primates.

# **III.** CONCLUSION

Here I presented an overview of a subset of the brain circuits and neurochemicals involved in human decision-making. A brain circuit involving the orbitofrontal other prefrontal regions, ventral striatum including the nucleus accumbens, and hippocampus all interact to subserve flexible, reward-guided decision-making. Dopamine plays a unique and critical role in decision-making through its phasic and tonic modes of action. Clearly, these systems are not the only ones involved in decision-making. Much work has been done linking, for example, the parietal cortex to decision-making, as well as other neurochemical systems, notably serotonin and oxytocin. The neuroscience of decision-making is a nascent and budding field. Our understanding of how the brain considers options and optimizes choices, though currently limited to relatively simple decision-making situations, increases with nearly every new issue of neuroscience journals.

# REFERENCES

- Aharon I, Becerra L, Chabris CF, Borsook D (2006): Noxious heat induces fMRI activation in two anatomically distinct clusters within the nucleus accumbens. *Neurosci Lett* 392:159-64.
- Aharon I, Etcoff N, Ariely D, Chabris CF, O'Connor E, Breiter HC (2001): Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* 32:537-51.
- Anden NE, Carlsson A, Dahlstroem A, Fuxe K, Hillarp NA, Larsson K (1964): Demonstration and Mapping out of Nigro-Neostriatal Dopamine Neurons. *Life Sci* 3:523-30.
- Anden NE, Dahlstroem A, Fuxe K, Larsson K (1965): Further Evidence for the Presence of Nigro-Neostriatal Dopamine Neurons in the Rat. *Am J Anat* 116:329-33.
- Bardo MT, Bowling SL, Pierce RC (1990): Changes in locomotion and dopamine neurotransmission following amphetamine, haloperidol, and exposure to novel environmental stimuli. *Psychopharmacology (Berl)* 101:338-43.
- Baxter MG, Parker A, Lindner CC, Izquierdo AD, Murray EA (2000): Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J Neurosci* 20:4311-9.
- Bayer HM, Glimcher PW (2005): Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129-41.
- Becerra L, Breiter HC, Wise R, Gonzalez RG, Borsook D (2001): Reward circuitry activation by noxious thermal stimuli. *Neuron* 32:927-46.
- Berridge KC (2003): Pleasures of the brain. Brain Cogn 52:106-28.
- Berridge KC (2007): The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)* 191:391-431.
- Block AE, Dhanji H, Thompson-Tardif SF, Floresco SB (2006): Thalamic-Prefrontal Cortical-Ventral Striatal Circuitry Mediates Dissociable Components of Strategy Set Shifting. Cereb Cortex.
- Brady AM, Glick SD, O'Donnell P (2005): Selective disruption of nucleus accumbens gating mechanisms in rats behaviorally sensitized to methamphetamine. *J Neurosci* 25:6687-95.

- Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P (2001): Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30:619-39.
- Breiter HC, Gollub RL, Weisskoff RM, et al (1997): Acute effects of cocaine on human brain activity and emotion. *Neuron* 19:591-611.
- Breiter HC, Rosen BR (1999): Functional magnetic resonance imaging of brain reward circuitry in the human. *Ann N Y Acad Sci* 877:523-47.
- Calabresi P, Maj R, Mercuri NB, Bernardi G (1992): Coactivation of D1 and D2 dopamine receptors is required for long-term synaptic depression in the striatum. *Neurosci Lett* 142:95-9.
- Cardinal RN (2006): Neural systems implicated in delayed and probabilistic reinforcement. *Neural Netw* 19:1277-301.
- Cardinal RN, Howes NJ (2005): Effects of lesions of the nucleus accumbens core on choice between small certain rewards and large uncertain rewards in rats. *BMC Neurosci* 6:37.
- Cardinal RN, Pennicott DR, Sugathapala CL, Robbins TW, Everitt BJ (2001): Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science* 292:2499-501.
- Cardinal RN, Winstanley CA, Robbins TW, Everitt BJ (2004): Limbic corticostriatal systems and delayed reinforcement. *Ann N Y Acad Sci* 1021:33-50.
- Carmichael ST, Price JL (1994): Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J Comp Neurol* 346:366-402.
- Carmichael ST, Price JL (1995): Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol* 363:615-641.
- Cohen MX, Ranganath C (2005a): Behavioral and neural predictors of upcoming decisions. *Cogn Affect Behav Neurosci* 5:117-26.
- Cohen MX, Ranganath C (2005b): Behavioral and Neural Predictors of Upcoming Decisions. *Cognitive, Affective, and Behavioral Neuroscience* 5:117-126.
- Cohen MX, Ranganath C (2007): Reinforcement learning signals predict future decisions. J Neurosci 27:371-8.
- Cohen MX, Young J, Baek JM, Kessler C, Ranganath C (2005): Individual differences in extraversion and dopamine genetics predict neural reward responses. *Brain Res Cogn Brain Res* 25:851-61.
- Comings DE, Blum K (2000): Reward deficiency syndrome: genetic aspects of behavioral disorders. *Prog Brain Res* 126:325-41.
- Cools R, Lewis SJ, Clark L, Barker RA, Robbins TW (2007): L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacology* 32:180-9.
- Damasio AR, Eslinger PJ, Damasio H, Van Hoesen GW, Cornell S (1985): Multimodal amnesic syndrome following bilateral temporal and basal forebrain damage. *Arch Neurol* 42:252-9.
- Daw ND, Doya K (2006): The computational neurobiology of learning and reward. Curr Opin Neurobiol 16:199-204.
- Deco G, Rolls ET (2006): Decision-making and Weber's law: a neurophysiological model. *Eur J Neurosci* 24:901-16.
- Di Chiara G (2002): Nucleus accumbens shell and core dopamine: differential role in behavior and addiction. *Behav Brain Res* 137:75-114.
- Egelman DM, Person C, Montague PR (1998): A computational role for dopamine delivery in human decision-making. *J Cogn Neurosci* 10:623-30.

- Ernst M, Nelson EE, Jazbec S, et al (2005): Amygdala and nucleus accumbens in responses to receipt and omission of gains in adults and adolescents. *Neuroimage* 25:1279-91.
- Eslinger PJ, Damasio AR (1985): Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology* 35:1731-41.
- Ferino F, Thierry AM, Glowinski J (1987): Anatomical and electrophysiological evidence for a direct projection from Ammon's horn to the medial prefrontal cortex in the rat. *Exp Brain Res* 65:421-6.
- Ferry AT, Ongur D, An X, Price JL (2000): Prefrontal cortical projections to the striatum in macaque monkeys: evidence for an organization related to prefrontal networks. J Comp Neurol 425:447-70.
- Finch DM (1996): Neurophysiology of converging synaptic inputs from the rat prefrontal cortex, amygdala, midline thalamus, and hippocampal formation onto single neurons of the caudate/putamen and nucleus accumbens. *Hippocampus* 6:495-512.
- Floresco SB, Blaha CD, Yang CR, Phillips AG (2001): Modulation of hippocampal and amygdalar-evoked activity of nucleus accumbens neurons by dopamine: cellular mechanisms of input selection. *J Neurosci* 21:2851-60.
- Floresco SB, West AR, Ash B, Moore H, Grace AA (2003): Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat Neurosci* 6:968-73.
- Frank MJ, Claus ED (2006): Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev* 113:300-26.
- French SJ, Totterdell S (2002): Hippocampal and prefrontal cortical inputs monosynaptically converge with individual projection neurons of the nucleus accumbens. *J Comp Neurol* 446:151-65.
- Frey JM, Ticku MK, Huffman RD (1987): GABAergic supersensitivity within the pars reticulata of the rat substantia nigra following chronic haloperidol administration. *Brain Res* 425:73-84.
- Frey S, Petrides M (2000): Orbitofrontal cortex: A key prefrontal region for encoding information. *Proc Natl Acad Sci U S A* 97:8723-7.
- Frey S, Petrides M (2002): Orbitofrontal cortex and memory formation. Neuron 36:171-6.
- Frey S, Petrides M (2003): Greater orbitofrontal activity predicts better memory for faces. *Eur J Neurosci* 17:2755-8.
- Frey U, Hartmann S, Matthies H (1989): Domperidone, an inhibitor of the D2-receptor, blocks a late phase of an electrically induced long-term potentiation in the CA1-region in rats. *Biomed Biochim Acta* 48:473-6.
- Frey U, Matthies H, Reymann KG (1991): The effect of dopaminergic D1 receptor blockade during tetanization on the expression of long-term potentiation in the rat CA1 region in vitro. *Neurosci Lett* 129:111-4.
- Fuster JM (1997): The Prefrontal Cortex. New York: Raven.
- Gaffan D, Murray EA (1990): Amygdalar interaction with the mediodorsal nucleus of the thalamus and the ventromedial prefrontal cortex in stimulus-reward associative learning in the monkey. *J Neurosci* 10:3479-93.
- Gaffan D, Murray EA, Fabre-Thorpe M (1993): Interaction of the amygdala with the frontal lobe in reward memory. *Eur J Neurosci* 5:968-75.

- Gasbarri A, Packard MG, Campana E, Pacitti C (1994a): Anterograde and retrograde tracing of projections from the ventral tegmental area to the hippocampal formation in the rat. *Brain Res Bull* 33:445-52.
- Gasbarri A, Sulli A, Packard MG (1997): The dopaminergic mesencephalic projections to the hippocampal formation in the rat. *Prog Neuropsychopharmacol Biol Psychiatry* 21:1-22.
- Gasbarri A, Verney C, Innocenzi R, Campana E, Pacitti C (1994b): Mesolimbic dopaminergic neurons innervating the hippocampal formation in the rat: a combined retrograde tracing and immunohistochemical study. *Brain Res* 668:71-9.
- German DC, Dubach M, Askari S, Speciale SG, Bowden DM (1988): 1-Methyl-4-phenyl-1,2,3,6-tetrahydropyridine-induced parkinsonian syndrome in Macaca fascicularis: which midbrain dopaminergic neurons are lost? *Neuroscience* 24:161-74.
- Glimcher P (2002): Decisions, decisions: choosing a biological science of choice. *Neuron* 36:323-32.
- Glimcher P (2004): *Decisions, Uncertainty and the Brain: The Science of Neuroeconomics.* Boston: MIT Press.
- Glimcher PW, Dorris MC, Bayer HM (2005): Physiological utility theory and the neuroeconomics of choice. *Games Econ Behav* 52:213-256.
- Gogtay N, Giedd JN, Lusk L, et al (2004): Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci U S A* 101:8174-9.
- Gold JI, Shadlen MN (2002): Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36:299-308.
- Goto Y, Grace AA (2005a): Dopamine-dependent interactions between limbic and prefrontal cortical plasticity in the nucleus accumbens: disruption by cocaine sensitization. *Neuron* 47:255-66.
- Goto Y, Grace AA (2005b): Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nat Neurosci* 8:805-12.
- Goto Y, O'Donnell P (2001): Network synchrony in the nucleus accumbens in vivo. J Neurosci 21:4498-504.
- Grace AA (1991): Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* 41:1-24.
- Grace AA (2000): The tonic/phasic model of dopamine system regulation and its implications for understanding alcohol and psychostimulant craving. *Addiction* 95 Suppl 2:S119-28.
- Guarraci FA, Kapp BS (1999): An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. *Behav Brain Res* 99:169-79.
- Haber SN, Fudge JL, McFarland NR (2000): Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci* 20:2369-82.
- Haber SN, Kim KS, Mailly P, Calzavara R (2006): Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *J Neurosci* 26:8368-76.
- Haber SN, McFarland NR (1999): The concept of the ventral striatum in nonhuman primates. *Ann N Y Acad Sci* 877:33-48.
- Hasselmo ME, Wyble BP (1997): Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. *Behav Brain Res* 89:1-34.

- Heinz A, Siessmeier T, Wrase J, et al (2004): Correlation between dopamine D(2) receptors in the ventral striatum and central processing of alcohol cues and craving. Am J Psychiatry 161:1783-9.
- Helmreich I, Reimann W, Hertting G, Starke K (1982): Are presynaptic dopamine autoreceptors and postsynaptic dopamine receptors in the rabbit caudate nucleus pharmacologically different? *Neuroscience* 7:1559-66.
- Hollerman JR, Schultz W (1998): Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1:304-9.
- Hooks MS, Kalivas PW (1995): The role of mesoaccumbens--pallidal circuitry in noveltyinduced behavioral activation. *Neuroscience* 64:587-97.
- Horwitz GD, Batista AP, Newsome WT (2004): Representation of an abstract perceptual decision in macaque superior colliculus. *J Neurophysiol* 91:2281-96.
- Huettel SA (2006): Behavioral, but not reward, risk modulates activation of prefrontal, parietal, and insular cortices. *Cogn Affect Behav Neurosci* 6:141-51.
- Huffman RD, Ticku MK (1983): The effects of chronic haloperidol administration on GABA receptor binding. *Pharmacol Biochem Behav* 19:199-204.
- Hurd YL, Svensson P, Ponten M (1999): The role of dopamine, dynorphin, and CART systems in the ventral striatum and amygdala in cocaine abuse. *Ann N Y Acad Sci* 877:499-506.
- Izquierdo A, Murray EA (2004): Combined unilateral lesions of the amygdala and orbital prefrontal cortex impair affective processing in rhesus monkeys. *J Neurophysiol* 91:2023-39.
- Jay TM, Witter MP (1991): Distribution of hippocampal CA1 and subicular efferents in the prefrontal cortex of the rat studied by means of anterograde transport of Phaseolus vulgaris-leucoagglutinin. *J Comp Neurol* 313:574-86.
- Kalivas PW, Volkow ND (2005): The neural basis of addiction: a pathology of motivation and choice. *Am J Psychiatry* 162:1403-13.
- Kasanetz F, Riquelme LA, Murer MG (2002): Disruption of the two-state membrane potential of striatal neurones during cortical desynchronisation in anaesthetised rats. *J Physiol* 543:577-89.
- Kasanetz F, Riquelme LA, O'Donnell P, Murer MG (2006): Turning off cortical ensembles stops striatal Up states and elicits phase perturbations in cortical and striatal slow oscillations in rat in vivo. J Physiol 577:97-113.
- Kelley AE (2004): Memory and addiction: shared neural circuitry and molecular mechanisms. *Neuron* 44:161-79.
- Kerr JE, Beck SG, Handa RJ (1996): Androgens selectively modulate C-fos messenger RNA induction in the rat hippocampus following novelty. *Neuroscience* 74:757-66.
- Kishiyama MM, Yonelinas AP, Lazzara MM (2004): The von Restorff effect in amnesia: the contribution of the hippocampal system to novelty-related memory enhancements. *J Cogn Neurosci* 16:15-23.
- Knight R (1996): Contribution of human hippocampal region to novelty detection. *Nature* 383:256-9.
- Knight RT, Nakada T (1998): Cortico-limbic circuits and novelty: a review of EEG and blood flow data. *Rev Neurosci* 9:57-70.
- Knutson B, Adams CM, Fong GW, Hommer D (2001): Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci* 21:RC159.

- Knutson B, Cooper JC (2005): Functional magnetic resonance imaging of reward prediction. *Curr Opin Neurol* 18:411-7.
- Komisaruk BR, Whipple B (2005): Functional MRI of the brain during orgasm in women. Annu Rev Sex Res 16:62-86.
- Koob GF (1992): Neural mechanisms of drug reinforcement. Ann N Y Acad Sci 654:171-91.
- Koob GF, Le Moal M (1997): Drug abuse: hedonic homeostatic dysregulation. *Science* 278:52-8.
- Kringelbach ML (2005): The human orbitofrontal cortex: linking reward to hedonic experience. *Nat Rev Neurosci* 6:691-702.
- Kringelbach ML, O'Doherty J, Rolls ET, Andrews C (2003): Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cereb Cortex* 13:1064-71.
- Kringelbach ML, Rolls ET (2004): The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog Neurobiol* 72:341-72.
- Kuhnen CM, Knutson B (2005): The neural basis of financial risk taking. Neuron 47:763-70.
- Laurens KR, Kiehl KA, Ngan ET, Liddle PF (2005): Attention orienting dysfunction during salient novel stimulus processing in schizophrenia. *Schizophr Res* 75:159-71.
- Lehmann J, Smith RV, Langer SZ (1983): Stereoisomers of apomorphine differ in affinity and intrinsic activity at presynaptic dopamine receptors modulating [3H]dopamine and [3H]acetylcholine release in slices of cat caudate. *Eur J Pharmacol* 88:81-8.
- Leon MI, Shadlen MN (1998): Exploring the neurophysiology of decisions. *Neuron* 21:669-72.
- Leung LS, Yim CY (1993): Rhythmic delta-frequency activities in the nucleus accumbens of anesthetized and freely moving rats. *Can J Physiol Pharmacol* 71:311-20.
- Lewis BL, O'Donnell P (2000): Ventral tegmental area afferents to the prefrontal cortex maintain membrane potential 'up' states in pyramidal neurons via D(1) dopamine receptors. *Cereb Cortex* 10:1168-75.
- Lisman JE, Grace AA (2005): The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46:703-13.
- Martinez D, Gil R, Slifstein M, et al (2005): Alcohol dependence is associated with blunted dopamine transmission in the ventral striatum. *Biol Psychiatry* 58:779-86.
- Matthews SC, Simmons AN, Lane SD, Paulus MP (2004): Selective activation of the nucleus accumbens during risk-taking decision making. *Neuroreport* 15:2123-7.
- McClure SM, Daw ND, Montague PR (2003): A computational substrate for incentive salience. *Trends Neurosci* 26:423-8.
- Mogenson GJ, Jones DL, Yim CY (1980): From motivation to action: functional interface between the limbic system and the motor system. *Prog Neurobiol* 14:69-97.
- Mogenson GJ, Yang CR (1991): The contribution of basal forebrain to limbic-motor integration and the mediation of motivation to action. *Adv Exp Med Biol* 295:267-90.
- Montague PR, Hyman SE, Cohen JD (2004): Computational roles for dopamine in behavioural control. *Nature* 431:760-7.
- Murer MG, Tseng KY, Kasanetz F, Belluscio M, Riquelme LA (2002): Brain oscillations, medium spiny neurons, and dopamine. *Cell Mol Neurobiol* 22:611-32.
- Nakahara H, Itoh H, Kawagoe R, Takikawa Y, Hikosaka O (2004): Dopamine neurons can represent context-dependent prediction error. *Neuron* 41:269-80.

- O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001): Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 4:95-102.
- O'Doherty J P, Buchanan TW, Seymour B, Dolan RJ (2006): Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron* 49:157-66.
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003): Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329-37.
- O'Donnell P, Grace AA (1995): Synaptic interactions among excitatory afferents to nucleus accumbens neurons: hippocampal gating of prefrontal cortical input. *J Neurosci* 15:3622-39.
- O'Donnell P, Greene J, Pabello N, Lewis BL, Grace AA (1999): Modulation of cell firing in the nucleus accumbens. *Ann N Y Acad Sci* 877:157-75.
- Ongur D, Price JL (2000): The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* 10:206-19.
- Osaka N, Osaka M (2005): Striatal reward areas activated by implicit laughter induced by mimic words in humans: a functional magnetic resonance imaging study. *Neuroreport* 16:1621-4.
- Otmakhova NA, Lisman JE (1996): D1/D5 dopamine receptor activation increases the magnitude of early long-term potentiation at CA1 hippocampal synapses. *J Neurosci* 16:7478-86.
- Pan HS, Frey KA, Young AB, Penney JB, Jr. (1983): Changes in [3H]muscimol binding in substantia nigra, entopeduncular nucleus, globus pallidus, and thalamus after striatal lesions as demonstrated by quantitative receptor autoradiography. *J Neurosci* 3:1189-98.
- Pandya DN, Yeterian EH (1996): Comparison of prefrontal architecture and connections. *Philos Trans R Soc Lond B Biol Sci* 351:1423-32.
- Parker A, Wilding E, Akerman C (1998): The Von Restorff effect in visual object recognition memory in humans and monkeys. The role of frontal/perirhinal interaction. J Cogn Neurosci 10:691-703.
- Prensa L, Richard S, Parent A (2003): Chemical anatomy of the human ventral striatum and adjacent basal forebrain structures. *J Comp Neurol* 460:345-67.
- Redgrave P, Gurney K (2006): The short-latency dopamine signal: a role in discovering novel actions? *Nat Rev Neurosci* 7:967-75.
- Redgrave P, Prescott TJ, Gurney K (1999): Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci* 22:146-51.
- Reuter J, Raedler T, Rose M, Hand I, Glascher J, Buchel C (2005): Pathological gambling is linked to reduced activation of the mesolimbic reward system. *Nat Neurosci* 8:147-8.
- Richfield EK, Penney JB, Young AB (1989): Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. *Neuroscience* 30:767-77.
- Risinger RC, Salmeron BJ, Ross TJ, et al (2005): Neural correlates of high and craving during cocaine self-administration using BOLD fMRI. *Neuroimage* 26:1097-108.
- Robbins TW, Everitt BJ (1996): Neurobehavioural mechanisms of reward and motivation. *Curr Opin Neurobiol* 6:228-36.
- Rolls ET (2005): Taste, olfactory, and food texture processing in the brain, and the control of food intake. *Physiol Behav* 85:45-56.

- Rolls ET (2006): Brain mechanisms underlying flavour and appetite. *Philos Trans R Soc Lond B Biol Sci* 361:1123-36.
- Romo R, Hernandez A, Zainos A, et al (2006): Decoding the temporal evolution of a simple perceptual act. *Novartis Found Symp* 270:170-86; discussion 186-90, 232-7.
- Ronesi J, Lovinger DM (2005): Induction of striatal long-term synaptic depression by moderate frequency activation of cortical afferents in rat. *J Physiol* 562:245-56.
- Sajikumar S, Frey JU (2004): Late-associativity, synaptic tagging, and the role of dopamine during LTP and LTD. *Neurobiol Learn Mem* 82:12-25.
- Schacter GB, Yang CR, Innis NK, Mogenson GJ (1989): The role of the hippocampalnucleus accumbens pathway in radial-arm maze performance. *Brain Res* 494:339-49.
- Schoenbaum G, Roesch MR, Stalnaker TA (2006): Orbitofrontal cortex, decision-making and drug addiction. *Trends Neurosci* 29:116-24.
- Schoenbaum G, Setlow B, Ramus SJ (2003): A systems approach to orbitofrontal cortex function: recordings in rat orbitofrontal cortex reveal interactions with different learning systems. *Behav Brain Res* 146:19-29.
- Schultz W (1998): Predictive reward signal of dopamine neurons. J Neurophysiol 80:1-27.
- Schultz W (2001): Reward signaling by dopamine neurons. Neuroscientist 7:293-302.
- Schultz W (2006): Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87-115.
- Schultz W, Dayan P, Montague PR (1997): A neural substrate of prediction and reward. *Science* 275:1593-9.
- Schultz W, Dickinson A (2000): Neuronal coding of prediction errors. Annu Rev Neurosci 23:473-500.
- Schwerdtfeger WK (1979): Direct efferent and afferent connections of the hippocampus with the neocortex in the marmoset monkey. *Am J Anat* 156:77-82.
- Sokolov EN (1990): The orienting response, and future directions of its development. *Pavlov J Biol Sci* 25:142-50.
- Stern EA, Kincaid AE, Wilson CJ (1997): Spontaneous subthreshold membrane potential fluctuations and action potential variability of rat corticostriatal and striatal neurons in vivo. J Neurophysiol 77:1697-715.
- Stoof JC, De Boer T, Sminia P, Mulder AH (1982): Stimulation of D2-dopamine receptors in rat neostriatum inhibits the release of acetylcholine and dopamine but does not affect the release of gamma-aminobutyric acid, glutamate or serotonin. *Eur J Pharmacol* 84:211-4.
- Strange BA, Duggins A, Penny W, Dolan RJ, Friston KJ (2005): Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Netw* 18:225-30.
- Strecker RE, Jacobs BL (1985): Substantia nigra dopaminergic unit activity in behaving cats: effect of arousal on spontaneous discharge and sensory evoked activity. *Brain Res* 361:339-50.
- Sugrue LP, Corrado GS, Newsome WT (2005): Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat Rev Neurosci* 6:363-75.
- Surmeier DJ, Kitai ST (1993): D1 and D2 dopamine receptor modulation of sodium and potassium currents in rat neostriatal neurons. *Prog Brain Res* 99:309-24.
- Sutton RS, Barto AG (1998): Reinforcement Learning. Cambride: MIT Press.
- Swanson LW (1979): The hippocampus. New anatomical insights. TINS:9-12.
- Swanson LW (1981): A direct projection from Ammon's horn to prefrontal cortex in the rat. *Brain Res* 217:150-4.

- Swanson LW, Cowan WM (1979): The connections of the septal region in the rat. J Comp Neurol 186:621-55.
- Tabuchi E, Mulder AB, Wiener SI (2003): Reward value invariant place responses and reward site associated activity in hippocampal neurons of behaving rats. *Hippocampus* 13:117-32.
- Tabuchi ET, Mulder AB, Wiener SI (2000): Position and behavioral modulation of synchronization of hippocampal and accumbens neuronal discharges in freely moving rats. *Hippocampus* 10:717-28.
- Tamminga CA, Nestler EJ (2006): Pathological gambling: focusing on the addiction, not the activity. Am J Psychiatry 163:180-1.
- Vinogradova OS (2001): Hippocampus as comparator: role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus* 11:578-98.
- Volkow ND, Fowler JS, Wang GJ (2002): Role of dopamine in drug reinforcement and addiction in humans: results from imaging studies. *Behav Pharmacol* 13:355-66.
- Volkow ND, Fowler JS, Wang GJ, Swanson JM (2004): Dopamine in drug abuse and addiction: results from imaging studies and treatment implications. *Mol Psychiatry*.
- Von Restorff H (1933): Über die Wirkung von Bereichsbildungen im Spurenfeld (The effects of field formation in the trace field). *Psychologie Forschung* 18:299-334.
- Williams SM, Goldman-Rakic PS (1998): Widespread origin of the primate mesofrontal dopamine system. *Cereb Cortex* 8:321-45.
- Wilson CJ (1993): The generation of natural firing patterns in neostriatal neurons. *Prog Brain Res* 99:277-97.
- Wilson CJ, Chang HT, Kitai ST (1983): Disfacilitation and long-lasting inhibition of neostriatal neurons in the rat. *Exp Brain Res* 51:227-35.
- Wilson CJ, Kawaguchi Y (1996): The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. *J Neurosci* 16:2397-410.
- Wise RA (1984): Neural mechanisms of the reinforcing action of cocaine. *NIDA Res Monogr* 50:15-33.
- Wiswede D, Russeler J, Hasselbach S, Munte TF (2006): Memory recall in arousing situations an emotional von Restorff effect? *BMC Neurosci* 7:57.
- Wolf JA, Moyer JT, Lazarewicz MT, et al (2005): NMDA/AMPA ratio impacts state transitions and entrainment to oscillations in a computational model of the nucleus accumbens medium spiny projection neuron. *J Neurosci* 25:9080-95.
- Wolf ME, Sun X, Mangiavacchi S, Chao SZ (2004): Psychomotor stimulants and neuronal plasticity. *Neuropharmacology* 47 Suppl 1:61-79.
- Yang CR, Mogenson GJ (1985): An electrophysiological study of the neural projections from the hippocampus to the ventral pallidum and the subpallidal areas by way of the nucleus accumbens. *Neuroscience* 15:1015-24.
- Zink CF, Pagnoni G, Martin ME, Dhamala M, Berns GS (2003): Human striatal response to salient nonrewarding stimuli. *J Neurosci* 23:8092-7.

Chapter 11

# **RISK AVOIDANCE AND RISK TAKING IN MENTAL HEALTH SOCIAL WORK<sup>\*</sup>**

## Shulamit Ramon

Interprofessional Health and Social Studies, School of Community Health and Social Studies, Anglia Polytechnic University, Cambridge, UK

### ABSTRACT

This chapter will look at issues related to risk in mental health social work, especially the over-focus on risk avoidance and the need to create space for calculated risk taking. Examples from Britain and continental Europe will be utilised.

### INTRODUCTION

The discourse on risk in British mental health social work (MHSW), and beyond it into the multidisciplinary context, is presently prolific. Nearly all of it focuses on *risk avoidance*. This focus can be observed in training for MHSW, research, policy and legislation.

This is less so in continental Europe. We therefore need to understand better the factors underlying the British preference, which largely omits serious consideration of *risk taking* in its policy and practice.

This chapter will explore the above-mentioned issues, and look at some relevant practice examples from the perspectives of social workers, mental health service users and their carers.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

## THE GENERIC RISK DISCOURSE

The discourse about risk does not emanate from the mental health system. Conceptually it is located in the debate about freedom and control, citizenship, crime and deviance, health and illness, with modernity and postmodernity (Giddens, 1991, Beck, 1992, Furedi, 1997, Rose, 2000). Its protagonists argued that each Western society copes with the increased uncertainty, which has been brought about by the growing gap between technological development and social fallibility, by increasing risk control, or by avoiding it. Thus control has become a synonym of avoidance. These eminent writers see risk control as an extension of social control, something to be lamented but to be taken for granted.

One does not need to disagree with the assumption of increasing, and ever more refined, social control to notice that risk is never considered in their discourse as worth taking individually and socially, as a buzz which stimulates people to give their best to a shared objective. Instead, the focus is on the increase in visible and less visible means of social control as an inevitable consequence of living with uncertainty.

This position goes against the grain of everyday life in the same societies in which ordinary people admire risk taking by a minority, and at times have a go at it themselves. Although a cliché, it is virtually impossible to live by adhering only to risk avoidance as a major strategy for living, without the utilisation of risk taking as a complimentary framework. This applies to the most banal risk such as crossing the road, extending to the more serious risks such as choosing a partner, or whether to be operated on when seriously ill. Some would go further to say that life is not worth living if it does not have some elements of risk taking, which put the spice of excitement into living. While some risk taking is what makes non-conforming or even illegal behaviour seen as worthwhile taking, most risk taking behaviour is within the narrow legal boundaries and the wider ones of conformity.

As a society we admire some types of risk taking, such as:

- Mountain climbing
- Explorers to unknown territories
- Marathon runners
- Diving
- Motor racing
- Writing a first successful novel
- Business gambling which pay off
- Inventors who risk all but succeed in becoming famous and successful
- Surgeons
- "Falling in love"

Moreover, these sociological approaches to risk negate the fact that risk taking also offers *the* promise of new opportunities to become what one feels s/he has the potential to be. This is well analysed by Harry Ferguson (2001) in his paper "Social work, individualisation and life politics".

Interested in emancipatory politics, Ferguson looks at how the lack of permanency in life in our time offers us also more diverse options, such as different types of intimate relationships, work, leisure and communication. All of these choices necessitate a greater degree of personal responsibility.

In his framework, a normal biography can become:

- the elective biography
- the reflexive biography
- the do-it-yourself biography
- the "tight-rope" biography

While aware that the possibilities we can exercise are limited by our past and present history, we have also become more aware that we usually have more than one option at most junctions in our life, and that to make a choice a reality risk would need to be taken. These risks can be financial (e.g. re-mortgaging our house to take a trip around the world); social (loss of status if we opt for a lower ranked occupation, a partner of the same sex); physical (being an explorer entails at times physical endurance); and psychological (the success of a first novel cannot be assured; moving to another country can lead to isolation). Yet more people are ready to take these risks for the sake of turning their dreams into reality. Thus life at present has become more risky for most people, in part due to the opening up of more possibilities for them to choose from.

It could be argued that the increase in possibilities also leads to increased stress levels, and inevitably to an increase in ill mental health, as it undermines the sense of coherence we need in order to live at peace with ourselves and others. Yet we know from the world of work that taking risk while *being in control* enables people to do more, do it better, and be more able to innovate (Bunce and West, 1996).

### MENTAL HEALTH RISK DISCOURSE

The risk discourse is particularly developed within Anglo-Saxon societies, where it is assumed that the ideology of individual freedom and autonomy is based on *regulation from within*, in which those who do not conform are regulated through a series of inclusionary and exclusionary mechanisms and processes. Prominent writers (e.g. Rose, 2000) further assume that the inclusionary discourse of people with disabilities (e.g. inclusion within education and employment), including those with mental illness, is no more than an attempt to ensure their conformity and hence the reduction of risk to their society and to themselves. Within this perspective, empowerment is no more than a fiction and a con, because it is aimed to offer the illusion of being in control over one's life, when one is merely toeing the line in terms of being a good – i.e. conforming – citizen.

Given this line of argument, it is not surprising that people who embody mental illness are perceived only as a threat to social cohesion and hegemony. Professionals working with them are seen as mandated to ensure the reduction of the threat, the return to the fold of responsible, self-regulated citizens, and if this is not possible then the application of exclusionary sanctions follows. The fascination within post-modernity with courting the irrational in particular, as expressed in literature, films, plays, poetry and the visual arts, has escaped the notice of these sociologists, as did emphatic approaches to understanding mental illness and the (temporary) excitement/liberation it offers – side by side with suffering – to those experiencing it. Even the possibility of the co-existence of threat with excitement is not recognised, let alone that of crediting users with having valuable expertise in experiencing this ambiguity to offer to the rest of us.

There are *important* reasons for focusing on risk avoidance in mental health, which include:

- Fear of others being hurt by people suffering from ill mental health.
- Fear of harm to self due to mental ill health.
- Media coverage of mental health users and of professional activities has focused mainly on homicide, and a little on suicide, by users.
- Politicians are very sensitive to issues related to public protection.
- Professionals have responded to the media and the politicians' stand by becoming defensive and have been instructed to assess and manage risk avoidance.
- As a result of the above, fear of the unpredictable element in some mental illness has been extended to all mental ill health categories.

Politicians in the Western world are perceiving risk policies and measures as one of the two most important core issues of mental health policy, the other being cost containment (Shera, Healy, Aviram and Ramon, 2002). The current debate in the UK on the proposals for the introduction of community treatment orders (CTOs) and of the preventive removal from the community of people diagnosed as having antisocial personality disorder highlights the direction in which the British government wishes to go, following that of many North American states and Australia (Brophy, Campbell and Healy, 2003, Hiday and Scheid-Cook, 1989, Home Office and the Department of Health, 1999). Continental Europe has not gone down the same route as the English-speaking world in terms of risk avoidance. None of the EU countries apart from Britain has legislated for CTOs, and the issue of risk avoidance does not seem to preoccupy their politicians, professionals or informal carers. The statistics of self harm and harm to others are often unreliable, but there is little to suggest that the rate of such incidents is lower in other Western European countries. Some of them -e.g. the Netherlands and the Scandinavian countries – have more developed alternatives to hospitalisation in a crisis, while others do not (e.g. Spain, France, Germany, Greece). Some countries - such as Italy – have developed more extensive methods of both solidarity and attention to underlying social issues, even though the level of financial benefits is lower than in Northern Europe. These approaches explain some of the reasons for the relative lack of interest in risk avoidance, but not all. There is some evidence that the media in Italy demonises people with severe mental illness less than in the UK (Ramon and Savio, 2000).

British MHSWs are the only group among social workers in Europe to have extended legal power and duties pertaining to risk avoidance in mental health. Since the introduction of the role of the Approved Social Worker (ASW) in 1983, ASWs have had to focus on their role in compulsory admission, which entails assessment through interviewing and documentary evidence, and reaching the decision as to whether to recommend such an admission or not in the light of availability of the least restrictive alternatives to hospitalisation, in conjunction with the decisions taken by psychiatrists and GPs (Barnes, Bowel, Newton and Fisher, 1990).

Insufficient time is left to engage even in the other tasks required by the legislation, such as following up clients, let alone in risk taking professional activities. While in the majority of cases the social worker is likely to recommend compulsory admission when a request for assessment has been made, in those minority cases in which the social worker does not recommend compulsory admission, s/he has to calculate – and take – the risk of what will happen to the client, to relatives and friends near to him/her, and to unknown others without such an admission. S/he needs also to calculate at any such application for compulsory admission the risk to the client from the admission itself – his/her self and social identity, relationships with others, the deprivation of civil rights during the admission, and the likelihood of drifting into a cycle of admissions and chronicity.

Care managers in countries such as Australia, Britain and Canada – likely to be social workers or nurses – collate evidence on risk and make decisions as to the boundaries of everyday life activities of service users.

Risk taking and risk avoidance issues are particularly acute for those, such as assertive outreach teams, or voluntary sector organisations (e.g. Turning Point, a British organisation focused on personalised work in the community with people who have mental illness and/or addiction). They are working with clients described as "hard to engage" in the community – people who have opted not to be in touch with statutory mental health services, usually after years of unsatisfactory contact from their perspective, who are nevertheless assessed as having serious mental health difficulties, and often also have substance misuse problems too.

In a recent study of four Turning Point projects, the key workers expressed their concerns about being on their own with the clients, while praising the ethos of the service and being mainly satisfied with the work they willingly engage with (Ramon, 2003b).

A recent research project (Williams, 2002) has highlighted the low morale among ASWs and the shortage in numbers in the north of England. If anything, one can predict that the situation in the rest of the country is at most no better, and is probably much worse in London, where the number of applications for compulsory admission is higher than elsewhere. The study does not go on to analyse the reasons for the low morale beyond the cumulative impact of an insufficient number of ASWs. I would argue that in addition to this burden, it is the nature of focusing on one narrow aspect of mental health work which contributes significantly to the low morale.

While ASW work has given MHSW more power (albeit largely of coercive nature), higher salary and greater prestige among mental health professionals, in its current stage it has also deprived them of the opportunity of engaging in the more caring, therapeutic, empowering and innovative work of which social workers are capable. Risk taking ranks high among the types of work ASWs have been discouraged from undertaking, as will be discussed below.

MHSW in continental Europe covers a wide range of emphases. In some countries these social workers focus on the financial eligibility of clients; in others they are part of a therapeutic team focusing on in-depth work with users and their families (e.g. Sweden), and/or they engage in building the links with relevant community groups (France, Germany), or in initiating and managing voluntary organisations (The Netherlands).

## **RISK AVOIDANCE PRACTICE AND POLICY ISSUES FOR USERS**

It is important to identify risk avoidance practice and policy issues, as risk taking would be about reversing them in some instances. They include admission to a special environment where the user is under constant observation and supervision and deprived of his/her freedom of movement and association with others. The more confined the setting, the greater the threat to other patients and staff, as well as to the index user, as there are fewer alternatives in which to channel the tension.

Although it is recognised that the more occupied the patient is with pleasant activities the lower the level of risk, little by way of fun or stimulating activities is available in acute admission settings for high risk patients, as distinct from some more long-term secure settings.

Thus the *organisational facet* feeds on unhelpful attitudes and modes of working (Bowers, 2002). An erosion of belief in the value of the system and of professional contribution to it is likely to follow, leading to burn-out and cynicism. Indeed, the rate of staff retention on acute wards and high security services is poor. Consequently, what they offer to users is a distancing, controlling practice.

The importance of the *physical environment* to the sense of safety and consequently to feeling empowered has emerged as one of the most significant dimensions in a study of empowerment components on different wards (acute, rehabilitation, long term), in which the dimensions originated from interviews with users (Schafer, 2003).

Social workers are not usually based in acute admission wards, and hence are spared their impact. However, this turns them into outsiders from the ward staff's and users' perspectives.

The realisation that the high tariff settings are not the most effective ones for the purpose of negative risk reduction has led to the introduction of community treatment orders (CTOs) in Australia (Brophy, et al, 2003), Canada, New Zealand, the United States (Hiday, 1989) and the United Kingdom (Foster, 2002) (where they are currently titled "supervision orders" and "guardianship"). These are compulsory orders focused on where people should live, their presence in intervention settings, their daily activities, the control of their finances, and their entitlement to support additional to that provided to other people with mental health problems who are not subject to these orders. The major sanction for not obeying is a return to hospitalisation.

Mental health service users in all of the countries which have established CTOs of one type or another oppose this new invention, primarily on grounds of the deprivation of civil liberties. They see the CTOs as yet another facet of the *medicalisation* of mental illness as much as a mechanism of greater coercive control, and prefer interventions focused on personalised relationships between staff and users, mutual support, and safe crisis facilities in the community as alternatives to hospitalisation (May, Hartley and Knight, 2003).

*Women* users have been vocal in expressing their views about acute admission wards and high secure hospitals as *unsafe* places for women (Mental Health Media, 2003). When Labour came to power in Britain in 1997, the government promised the end of wards in which accommodation is shared by men and women, and the forced socialisation of women with men on such premises. Thus far, however, little has been achieved in phasing out mixed wards. This raises the issue, which is embarrassing and difficult to confront, of users harming other users, and of a philosophy which women see as harmful.

*Black* users' and carers' associations, and a number of both black and white professionals, have protested since the 1980s against the fact that black people, especially young black men, are more likely to be compulsorily admitted, over-medicated, and less likely to be given talk therapies than white users (Thomas, Stone, Osborn, Thomas and Fisher, 1993).

People to whom a diagnosis of *personality disorder* has been attributed have become the newest group to be stigmatised by the British media, but not in continental Europe. A study in which user researchers with personality disorder interviewed other users with the label has demonstrated this (Ramon, Castillo and Morant, 2001, Castillo, 2002). Most of the respondents reported being treated considerably less well after being given the diagnosis than before. More subtle differences emerge when their own perceptions of their difficulties are compared with the diagnosis. Eighty per cent of the respondents reported depression and impact of child abuse as their major problems, and the same percentage had depression or anxiety attributed as their secondary diagnosis, while having personality disorder as their first (Castillo 2002). It is likely that what has tilted the professional decision in the direction of PD rather than that of depression relates to assumed risk to others and self, coupled with moral apprehension about their socially unacceptable behaviour.

The issue of fear from potential violence lies behind the political agenda in the context of personality disorder. Fear of black young men is more rooted within professional practice, where it is not articulated as it is embarrassing for mental health professionals, many of whom are from ethnic minorities, to admit to accepting racist beliefs. Although women constitute a clear majority of mental health professionals, the overwhelming focus on risk within acute and secure facilities leads to a particular type of blindness to the impact of admission to these same facilities.

### **POLICY ISSUES FOR CARERS**

A recent study of ASWs, users and the Nearest Relative (NR) (the nominated relative who has some legal powers within the British legislation to ask and agree to compulsory admission and discharge) (Rapaport, 2003) has illustrated the element of poor communication and poor information-giving in potential risk situations between the NRs and the ASWs who have the responsibility to provide this information. This process is taking place against a background of increased tension between users and their relatives, which in part has led to the request for an assessment for compulsory admission, at times to a sense of being betrayed by the users, and being neglected and unheard by the NR, who feels unable to achieve any positive conflict resolution. In addition, the legal information is written in a language incomprehensible by lay people, and receiving it in at a time of crisis is doomed to failure.

Carers perceive professionals' attempts to not share information with them, and to doubt their perspective on what is happening to their family member, as indications of disrespect of their contribution and of attempts to exclude them as key people in this context (Hugman and Pearson, 1995).

Carers in Anglo-Saxon countries seem to welcome CTOs as providing a much needed structure and support. In fact, organisations such as NAMI (National Alliance of the Mentally III) in the US have led the demand for CTOs. Yet the largest British carers' organisation, –

Rethink (formally the National Schizophrenia Fellowship) – has opposed the latest proposal to allow preventive hospital detention for people with antisocial personality disorder as a step too far in deprivation of civil liberties.

## **RISK TAKING**

Value and Theory Framework

Attention was paid in the introduction to the integral part of risk taking to our lives, and especially to its emancipatory component. Given the justified focus on risk avoidance in the work with people experiencing mental ill health, the question as to the place of risk taking within this context becomes central. Risk taking within mental health is based on values which are an integral part of universally accepted social work values. The theoretical framework and the specific values are embedded in the following:

- The Strengths approach
- Social role valorisation
- The right to self determination
- The right to fail

### The Strengths Approach

This approach was developed in the US by Saleebey (1992) and by Rapp for working with users of mental health services (Rapp, 1998), in the wake of psychiatric hospital closure in the late 1960s and early 1970s. The closure was in part premised on the belief that the resettled patients do not only have weaknesses, but also have abilities and potential that have not been utilised within the institution.

Within this approach the following is assumed:

- The focus is on individual strengths rather than pathology.
- Client-worker relationships are primary and essential.
- Interventions are based on client self-determination.
- Community is viewed as an oasis of resources, not as an obstacle.
- Aggressive outreach is the preferred mode of intervention.
- People suffering from severe mental illness can continue to learn, grow and change.

The focus on strengths rather than on pathology is not a tautology. It is a shift in emphasis which does not deny that people may have difficulties and problems, but accepts that unless we focus simultaneously on the abilities that they possess, we will not manage to reveal the latter and will not motivate users, carers and professional providers to focus on the strengths. This was revolutionary at the time, and remains so for a large number of lay and professionals who are yet to shift from the pathology-only position. There are consequences of these assumptions in everyday practice, which entail:

- Paying more attention than before to enhancing the individualisation of clients
- Facilitating of partnership rather than adversary
- Fostering empowerment
- Blending societal, programmatic, and client goals

We will look at how these are in fact implemented in the practice examples section.

The protagonists of the strengths approach do not discuss social control or risk taking issues *per se*. The unabashed focus on assertive outreach perhaps indicates that the critique of this intervention as too intrusive in people's lives has either not been heard by them, or has been rejected without an open debate.

### Social Role Valorisation

Social Role Valorisation (SRV) was developed within the field of learning difficulties by Wolfensberger in 1983 (Wolfensberger, 1983). This principle pinpoints disabled people's right to a socially valued place amidst the rest of us. It analyses the tragedy of institutionalisation as one which segregated residents from ordinary living and ordinary people, thus depriving them of opportunities to develop further their social and personal knowledge and skills, and in turn rendering them incapable and thus justifying the stigma attached to their disability by society (Ramon, 1991). The framework takes it for granted that people, including those disabled, have both weaknesses and strengths, abilities and disabilities. Those recognised as disabled may need more support to bring to the fore their abilities, and, for them to have a genuine place, our emphasis on economic success and work prestige needs to be revised to include basic human qualities such a loving and giving.

SRV has developed an audit of services which looks at whether these are enabling or disabling; do away with stigma or perpetuate it in the way they use language in connection with their users; the way the physical environment is giving the message of respect or lack of it; whether users have real choices in what is on offer to them or not. Wolfensberger used harsh words in describing the damage professionals do to disabled people, especially mental health professionals, by treating them as bundles of pathology.

This has impacted on the process and outcomes of psychiatric hospital closure and resettlement in some British settings (Wainwright, 1992), as well as on care management (Brandon D., Atherton, and Brandon A., 1996; Brandon, 1997).

SRV has been criticised as being too white and middle-class in its aspirations for disabled people. Also, it does not have a critical dimension concerning the type of social control which the adherence to a socially valued role brings with it. Issues of risk avoidance, however, have been part of its application to practice with a variety of disabled people. There, risk avoidance has been treated pragmatically, with minimum fuss and maximum respect to the disabled person.

SRV developed initially in Denmark, Sweden and Norway in the late 1950s and early 1960s, under the heading of normalisation and in the field of learning difficulties. Scandinavian social workers have therefore taken for granted this philosophy.

Although Italian social workers were not using the same terminology, the emphasis in the Italian psychiatric reform, from its inception in the late 1960s onwards, on full deinstitutionalisation, de-stigmatisation, and full and respectful integration of people with mental illness into the community implied a very similar conceptual and value framework. If anything, the Italian perspective has been less psychologically focused in relation to mental ill health and more preoccupied with removing social barriers to the rehabilitation of people with mental illness, and the disruption of the connections between poverty and mental illness.

#### Self-determination and the Right to Fail

The right to self-determination is a core value specific to social work, guiding workers to enable service users to exercise choice as to every aspect of their lives. Rooted in an individualistic and liberal society, the potential conflict with socially desirable choices vs. individual choices has been handled early on in the literature of social work by arguing that there is always a range of choices which are compatible with both demands/requirements. This amounts to accepting social preferences as morally better than individual choices in principle. This principle has been criticised as "naïve" or hypocritical insofar as social workers are only too aware of the many constraints put on clients wanting to exercise their self-determination as against the views of professionals, not to mention constraints which arise out of their family and social background. Yet there are social workers who take this principle seriously and who do their best to enable clients to exercise it (McDermott, 1975).

This right is also a part of the legacy of social work, having been recognised as a principle subsumed under the right to self-determination in social work in Sawyer's 1975 article (Sawyer, 1975). Sawyer argued that it is both difficult and unrealistic to expect clients to grow and develop without taking calculated risks, and without failing from time to time. His argument is echoed in Brandon's writings (1991) for people with mental health problems.

I do not know of any other helping profession which accepts failure as a *right*, rather than as an inevitable price we may pay when our calculated risk-taking strategy has failed. This recognition within social work does not necessarily mean that most social workers enable the exercise of this right, but it does mean that the issue is acknowledged and at times debated. In the original discussion, Sawyer looked at the underlying reasons which prevent social workers from enabling their clients to exercise this right. These include:

- the wish to protect the client from new failure, given that many of them would come with a history of failures;
- the wish to protect one's professional reputation as someone who fosters success and prevents failure;
- the fear that the client will collapse if s/he failed yet again.

All of these reasons are highly relevant within the field of mental health, and underline the focus on risk avoidance. However, those who accept the principles of strengths and SRV would consider that people experiencing mental ill health can take calculated risks with more support than other people would need, rather than that they should not take any risks. It is only when *calculated* risk taking is allowed to be considered that the right to fail can be exercised. Socially, our attitude to people who fail when they take a risk is less favourable than it should be logically. More often than not they are blamed for miscalculating the risk and therefore for the failure. Often, those who knew them well at the height of their success are ashamed to be seen with them at the time of failure. As a society, it would seem that we are thus afraid of failure.

## **TYPES OF RISK TAKING**

Risk taking is necessary in each aspect of mental health where the primary purpose is that of improving the quality of life of service users; yet each aspect brings with it somewhat different issues and considerations. A secondary purpose is that of improving the quality of the working life of front-line workers.

Below, I shall look at a number of relevant examples of projects established by British MHSWs, attempting to tease out the issues related to risk taking.

a. *Training and engaging local women* as leaders of groups of isolated women by the Family Welfare Association (FWA) in the 1970s.

This project was the first to apply the lessons of George Brown's work on the social origins of depression (Brown and Harris, 1978) by training local women in two London boroughs to lead groups of women who were single parents and isolated. The meetings took place at the leaders' homes, and were run as a social event which enabled discussion. Refreshments and child-minding were on offer. The projects were successful in enabling the participants to create their own social networks and thus to reduce the level of isolation and vulnerability by using very modest means.

Running these groups entailed risk taking for the leaders, the participants, and the social workers at the FWAs who were responsible for them. The main risk taken was believing that a six-week training module was sufficient to enable the lay leaders to run a group of vulnerable people, yet it worked well and empowered the local community (Knight, 1978).

This has been one of the few preventive measures taken at a level in which the lay community could actively participate.

b. *Creating the first user forum* in mental health (Iris Nutting, then team leader of Camden mental health team at Friern Hospital) in the early 1980s.

Responding to the beginning of the process of hospital closure, Ms. Nutting – literally on her own with one volunteer who did not work for social services – decided to establish this first forum. She invited some of the more articulate expatients of Friern, who were also members of Camden Mind, for a consultation process around the closure policy, aiming to achieve a policy recommendation document.

The results surpassed these expectations. Not only did the users prepare the document, but many of them became involved in the management of resettlement projects in the area, and a number moved on to become leaders of the British user

movement. One became a qualified social worker; another a researcher who is running a user-led research unit.

The risk taken here was that the users might not live up to the challenge, might not come on a regular basis, might not find the policy exercise useful, or might have a mental health crisis and not come back (indeed some participants had a crisis, but most of them did come back).

c. *The Chesterfield support network,* in which users run their own groups within a large community centre and are members of the centre's management committee since 1982 (Hennelly 1990).

This service enables the users to run their own activities on their own, inclusive of budget and constitution. Paid workers are available for individual consultation if needed. It is based in a generic community centre, and its users send representatives to the management committee of that centre.

The risk taken is to put to the test the Strengths approach by letting users lead their own groups, within an environment which is ordinary and not geared specifically to the needs of users with mental health problems.

Users are proud of the achievements of the network, which are theirs not less than the workers'.

d. Residential crisis facilities (Greenwich, 1980s, Kettering, 1990s).

These facilities were established by social services, offering both a drop-in and a respite option.

Although psychiatrists come on a sessional basis, they are not staff members. People in crisis require considerable attention, and things can go wrong in terms of harm to self and property, and less so in harm to others. Indeed, the Greenwich facility closed down after a couple of years due to the burnout of its staff group (Brangwen, 1990), while the Kettering facility is still going strong. Perhaps the differences lie in the type of support provided to the workers.

Such facilities top the users' agenda as an alternative to hospitalisation when in crisis.

e. *The Building Bridges project* (Diggins, 2000) began in the late 1990s. It offers children and parents (who have a mental illness) tailor-made support of different kinds.

This project was initiated in a south London borough by senior social workers in consultation with users and members of other professions, in response to the high number of black children going into care when their parents had a mental health crisis. It is run by a voluntary organisation with joint funding and initial Department of Health funding. The project offers casework, counselling, group work, weekends away and liaison with schools. The main feature is ensuring that each member of the family is attended to and gets the intervention method appropriate to her/his needs.

The risk taken here include child protection going wrong, parents who may harm their children, and children whose growth may be harmed by staying with the parent who has a mental illness. The project takes care of young carers, and enables parents to remain in the parenting role.

This project provides both measures of risk avoidance and risk taking, preventive and crisis response.

f. The Family Group Conference in mental health project as a risk taking enterprise.

The Family Group Conference follows the New Zealand Maori custom in which the community elders made joint decisions, having listened to each individual involved. Family and friends are invited to a meeting or series of meetings if need be, chaired by an independent chairperson. They are informed of the concerns raised by the professionals and the range of available options to meet these concerns. The family group decisions are to be followed by the professionals too.

In the Essex Social Services project (2002) the focus has been on an adult with serious mental health problems in Essex, and his/her family and friends. These meetings are perceived as high, but worthwhile, risk due to a three-fold baseline reasoning. When they are taking place, the family group has reached the point of "no-return", as the adult, or a child, may have to be taken out of the family, symbolising the family group's failure to look after itself and its own. It is judged that without a real shake-up no change is going to take place; the meetings can shake up and unstuck the "status quo" in which the group has been paralysed or acting in a destructive way, even though this is a high risk-taking strategy.

Negative risks may entail the following:

- Some family members have come reluctantly.
- Painful secrets may be revealed.
- The level of pain will be so high that the family group will never be able to work together again.
- The group is unable to reach a decision.
- The decision reached is beyond the menu of possibilities outlined by the professionals.
- However, if the risk taking is successful, then:
- Group members came because they care about each other.
- Good experiences will be shared too.
- The group discovers it has more strengths than it thought before the meeting.
- The group reached a decision shared by most members.
- The decision is within the menu of possibilities outlined by professionals.

It is assumed that with the shake-up that the family group conference brings about, the abilities within the family group have a much better chance of coming to the fore, even though the emotional burden has not been worked through. With an agreement, a commitment to put it into action, and a sense of ownership of the solution, as well as of being in control, there is a much better chance that things will move in the right direction.

The specific study of sixteen family group conferences demonstrates positive views of the process, and bodes well for the future, which was bleak up to the point of having the family group conference. It also highlights the fact that carers can be meaningful and equal partners to major decision-making.

g. User researchers and risk taking

Encouraging mental health service users and enabling them to become researchers is a risk taking strategy insofar as the following may happen:

- Users may drop out during the training due to personal crisis or disillusionment with research.
- The research work with other users may lead them to reliving painful experiences.
- They may be less good researchers than fully qualified ones, and hence carry out their research work at a lower level of quality than qualified researchers.

As against these risks stands the wish to redress the serious imbalance in our knowledge about mental ill health and mental health services, namely the fact that most of our current knowledge has not included users' perspectives, and if anything had invalidated their perspectives.

Furthermore, the perceiving of users as researchers can be contemplated only by those believing in the Strengths approach, in the ability of users to learn to the sophisticated level most research tasks require within a short time, and in the validity of users' views.

It therefore becomes a challenge – or a risk taking strategy – to develop good enough support networks and project methods to reduce the risks outlined above, while enhancing users' abilities to be good researchers (Castillo, 2002, Ramon, 2003).

### **SUMMARY**

Risk is an emotive issue in mental health practice, as well as for the lay public in all European societies. The emotive layer relates to the fear that the unpredictable behaviour of people experiencing mental illness will lead to physical hurt of others primarily, but also of themselves, pushing into the direction of risk avoidance.

Nevertheless, I hope a convincing case has been made in this chapter as to:

- why risk taking is inherent in mental health social work;
- why it should be focused upon and developed further;
- the fact that it can be carried out while taking into account risk avoidance issues, but without letting the latter dictate the agenda.

Social workers are better equipped than other mental health professionals to do so, as they have the initial values, the knowledge and the skills required, and a history of doing so successfully. Yet the retrenchment into risk avoidance, canonised by the role of the ASW, has meant that very little of their time and energy is spent on this necessary element of good mental health work, despite the historical legacy.

I would like to argue that it is high time to re-claim this ground, at a time when the mental health system has become more open to preventive and promoting mental health work, and at the time when the role of social workers in mental health is re-examined. As we know well, this shift would require the commitment of social workers first and foremost, as other professions are keen to claim this ground for themselves.

### REFERENCES

- Barnes, M./Bowl, R./Fisher, M. (1990): Sectioned; London: Routledge.
- Beck, U. (1992): The Risk Society; London: Sage.
- Bowers, L. (2002): Dangerous and Severe Personality Disorder Response and Role of the *Psychiatric Team*; London: Routledge.
- Brandon, D./Atherton, K./Brandon, A. (1996): *Handbook of Care Planning; London*: Positive Publications.
- Brandon, D. (1997): *The Trick of Being Ordinary;* Cambridge: Anglia Polytechnic University.

Brangwyn, G. (1990): *Constructing a Crisis-Focused Social Service*; in: Ramon, S. (ed.) Psychiatry in Transition, London: Pluto Press.

- Brophy, L./Campbell, J./Healy, B. (2003): Dilemmas in the Case Manager's Role Implementing Involuntary Treatment in the Community; in: Special Edition on Psychiatry, Psychology and Law of the Journal of the Australian and New Zealand Association of Psychiatry, Psychology and Law.
- Brown, G.W., Harris. T. (1978): *The Social Origins* of Depression; London: Tavistock Publications.
- Bunce, D./West, M.A. (1996): Stress Management and Innovation Interventions at Work; in: *Human Relations*, 49, 209-232.
- Campbell, J. et al (2001): The Management and Supervision of Approved Social Workers Aspects of Law, Policy and Practice; in: *The Journal of Social Welfare and Family Law*.
- Castillo, H. (2003): Personality Disorder: Trauma or Temperament?; London: Jessica Kingsley.
- Department of Health (1999): *The National Service Framework for Mental Health*; London: HMSO.
- Diggins, M. (2000): Innovation as a Professional Way of Life the Building Bridges Project for Parents-Users of Mental Health Services and their Children; in: Ramon, S. (ed.) *A Stakeholder Approach to Innovation in Mental Health Services*, Brighton: Pavilion.
- Essex Social Services Department (2002): *Essex Family Group Conference in Mental Health*, Essex County Publications.
- Ferguson, H. (2001): Social Work, Individualisation and Life Politics; in: British Journal of Social Work, 31, 41-55.
- Foster, N. (2001): Involuntary Outpatient Commitment: Managing Mental Health Risks the New Hampshire Way; in: *Journal of Mental Health and Learning Disabilities Care*, 4, 11.
- Foster, N. (2002): Control, Citizenship and "Risk" in Mental Health: Perspectives from UK, US and Australia; *Paper presented at the Socrates European Conference on Risk and Mental Health*, University College Cork, May 9<sup>th</sup> 2002.
- Furedi, F. (1997): Culture of Fear: Risk-Taking and the Morality of Low Expectation; London: Cassell.
- Giddens, A. (1991): Modernity and Self-identity; Cambridge: Polity Press.
- Healy, B./Brophy, L. (2002): Law, Psychiatry and Social Work; in: Swain, P.A. *The Shadow of the Law: The Legal Context of Social Work Practice*, New South Wales: The Federation Press.

- Hennelly, R. (1990): Mental Health Resource Centres; in: Ramon, S. (ed.) *Psychiatry in Transition*; London: Pluto Press.
- Hiday, V.A./Scheid-Cook, T.L. (1989): A Follow-up of Chronic Patients Committed to Outpatient Treatment; in: *Hospital and Community Psychiatry*, 40, 4, 52-58.
- Home Office and the Department of Health (1999): *Managing Dangerous People with Severe Personality Disorder: Proposal for Policy Development*; London: HMSO.
- Hugman, G./Pearson, G. (1995): The Silent Partners; London: National Schizophrenia Fellowship.
- Knight, C. (178): Neighbourhood Support Groups; London: Family Welfare Association.
- May, R./Hartley, J./Knight, T. (2003): Making the Personal Political; in the *Psychologist*, 16, 4, 182-183.
- McDermott, R. (ed.) (1975): *Self Determination in Social Work*; London: Routledge and Kegan Paul.
- Mental Health Media (2003): What Women Want Mainstreaming Women's Mental Health; Mental Health Media, *www.mhmedia.com*.
- Rose, N. (2000): *Powers of Freedom Reframing Political Thought*; Cambridge: Cambridge University Press.
- Ramon, S. (ed.) (1991): Beyond Community Care Normalisation and Integration Work; Basingstoke: Mind Macmillan.
- Ramon, S. /Savio, M. (2000): A Scandal of the 80ws and 90s Media Representations of Mental Illness Issues in Britain and Italy; in: Ramon, S. (ed) A Stakeholder's Approach to Innovation in Mental Health Services: A Reader for the 21<sup>st</sup> Century, Brighton: Pavilion Publishing.
- Ramon, S./Castillo, H. Morant, N. (2001): Experiencing Personality Disorder A Participative Research; in: *International Journal of Social Psychiatry*, 41, 4, 1-15.
- Ramon, S. (2003a) (ed): Users Researching Health and Social Care An Empowering Agenda?; Birmingham: Venture Press.
- Ramon, S. (2003b): Evaluating Three Turning Point Projects A Multi-stakeholder Perspective; Unpublished Report, July.
- Rapaport, J. (2003): The Ghost of the Nearest Relative under the Mental Health Act 1983 Past, Present, and Future; in: *Journal of Mental Health Law*, 51-65.
- Rapp, C.A. (1998): The Strengths Model Case Management with People Suffering from Severe and Persistent Mental Illness; Oxford University Press.
- Rose, N. (2000): *Powers of Freedom Reframing Political Thought*; Cambridge: Cambridge University Press.
- Saleebey, D. (ed.) (1992): The Strengths Approach in Social Work Practice; Longman.
- Sawyer, P. (1975): The Right to Fail; in: McDermott, R. (ed.) *Self Determination in Social Work*, London: Routledge and Kegan Paul.
- Schafer, T. (2003): Researching Empowerment; in: Ramon, S. (ed.) Users Researching Health and Social Care: An Empowering Agenda, Birmingham: Venture Press.
- Shera, W./Healy, B./Aviram, U./Ramon, S. (2002): Mental Health Policy and Practice A Multi-country Comparison; in: *Journal of Health and Mental Health Social Work*, 35, 1-2, 547-575.
- Thomas, C.S./Stone, K./Osborn, M./Thomas, P./Fisher, M. (1993): Psychiatric Morbidity and Compulsory Admission among UK-Born Europeans, Afro-Caribbeans and Asians in Central Manchester; in: *British Journal of Psychiatry*, 163, 91-99.

- Wainwright, T. (1992): The Changing Perspective of a Resettlement Team; in: Ramon, S. (ed.) *Psychiatric Hospital Closure Myths and Realities*, London: Chapman Hall.
- Williams, C. (2002): A *Survey of Approved Social Work Services*; Durham: Northern Centre for Mental Health.
- Wolfensberger, W. (1983): The Principle of Social Role Valorisation; in: *Mental Retardation*, 21,6, 234-9.

Chapter 12

# **DECISION MAKING AT THE END OF LIFE**\*

## Kathryn A McCall and Ann Marie Rice

Bedford Hospital, Fort William, Scotland

### ABSTRACT

Palliative care is becoming an increasingly accepted approach to care when cure is no longer possible. Decision making during the palliative phase of any illness can be complex, involving integration of knowledge, legal & ethical considerations, & excellent communication skills for healthcare professionals, patients and, where appropriate, their family/carers. This decision making process may become even more challenging as the focus of care changes from uncertainty in terms of survival or time to progression of disease to the often highly charged emotional terminal stage of the illness.

Patient participation is increasingly purported as a goal in end of life care (Clover et al 2004), and research suggests that most patients prefer a collaborative role in decision making (Pattison 2004). However there is a paucity of information about how those nearing the end of life can actively engage with healthcare professionals & family members to achieve this in such diverse areas as personal care (Clarke & Seymour 1999); symptom management; hydration & nutrition; advance directives (Scanlon 2003); compliance with medical advice; and place of end of life care (McCall & Rice 2004).

The patient's right to self-determination is evidenced in the increasing use and possible questionable legality of advance directives, both written & documented oral statements, and in the ongoing current debate regarding physician-assisted suicide (Scanlon 2003). Traditional decision making roles are changing and the paternalistic medical model, which has previously led much end of life care, is having to be modified, challenging professional caregivers to help the decision making process to be more collaborative.

This chapter aims to explore the psychology of decision making in end-of-life care where competency to make decisions may be threatened by changes in cognitive ability due to advancing disease, the unique psychosocial stressors related to terminal illness, or altered judgement related to the use of high dose drugs in symptom management.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

### INTRODUCTION

Throughout the western world the philosophy of palliative care has become an increasingly accepted approach to caring for patients who have advancing life-limiting illness. Advances in medical science & technology combined with the primary goal of medical care being to restore or maintain health as far as possible, result in difficulties defining when active care ends & palliative care begins. This also often blurs the transition to end of life care which should focus on appropriate treatment and compassionate care (Scanlon 2003). In addition issues such as advance directives and assisted suicide further challenge society to define the rights of dying (Hermsen & ten Have 2005). De Haes & Tuenissen (2005) succinctly describe palliative care as being in three phases – diagnosis of disease recurrence and life-prolonging as opposed to curative treatment; the phase of treatment directed at symptom control only; and the terminal phase directed at end of life care. For the purpose of this chapter the term "end of life" refers to the latter period of that often-lengthy time frame, more usually the final three months of life.

Decision-making is generally accepted as part of everyday activities, a process that is often unconscious and dynamic. An individual's right to be autonomous and to be able to decide about his or her own life is a commonly accepted ethical principle (Beauchamp & Childress 2001). However when dealing with a life-limiting illness, that decision making process may change as the process of choosing one alternative over another becomes more complex. The individual nearing the end of life may become increasingly dependent on others for care and this may result in the loss of some personal self-determination and to a greater degree loss of independence as illness advances. In specific care situations, caregivers may not be willing to let the patients make decisions (Savulescu 1995) or decisions may not accord with the views of the healthcare professional (Spriggs 1998).

This chapter aims to explore the challenges to decision making from the perspective of the patient, family/carers and health care professionals. The chapter will also explore the issue of physician-assisted suicide and the role of advance directives as a means of decision-making at the end life when the individual is no longer "competent" to make his/her own decisions.

The science of medical decision-making assumes that choices are made on a rational basis following the sharing of information between physician and patient (Charles et al 1997). The paternalistic approach that looked at patient involvement in decision making in matters medical as passive participants has now given way to a concept that stresses their active participation, promoting autonomy and self-determination, in all matters of care. The fundamental principles of palliative care deem that honest and open discussion takes place regarding the benefits & risks of any treatment option and this is considered alongside individual patient wishes (Ferrell et al 2003). This process implies that options are offered, their expected outcomes explored, both are understood, deliberation takes place and a shared decision based on individual preferences is communicated and agreed. However, in reality this has shown to raise more paradoxes as expected decisions based on quality of life are often over-ruled in favour of survival at all costs from both the perspective of the health care professional and the patient themselves. It would seem that the giving of treatment, and accepting of it, is one way to avoid discussing the inevitability of advancing disease (de Haes & Koedoot 2003). Often end of life care discussions are delayed until a patient or family is in

crisis or the patient is too close to death to participate in such discussions. This may be due to avoidance by health care professionals not wishing to engender a loss of hope in patients and their families, or avoidance by both patient and carer in an attempt to protect the other from an emotional and difficult subject.

Decision-making during the palliative phase of any illness can be complex often because competency to make such decisions may be threatened by changes in cognitive ability due to advancing disease or altered judgment related to the use of high dose drugs in symptom management. It remains none-the-less of immense importance. Intrusive treatments such as palliative chemotherapy can be offered to reduce tumour bulk or manage symptoms but uncertain or absent gains in terms of survival outcomes or quality of life have to be weighed against the burdens of such treatments. Other considerations include legal and ethical issues and how information about treatment, non-treatment, with-drawing treatment and other care options are communicated to help patients, and where appropriate, their families & carers, be involved in the decision making process for advance care planning. When the focus of care changes from uncertainty in terms of survival or time to progression of disease to the often highly charged emotional terminal phase of illness, the process of decision-making becomes ever more challenging. Although the choices made by patients in the palliative phase of their illness can appear to be uncomplicated on the surface, the context of uncertainty and unpredictability of this phase increase the underlying complexity of the process and often a process of deliberation and trade-offs is required. Bottorff et al (1998) describe this process as striving for a balance at the end of life and consists of three distinct phases – weighing things up, communicating choice and then living with that choice. Studies into the decision making process are divided as to how much participation patients wish to have in this end of life discussion. Pattison (2004) suggests that most patients prefer a collaborative role in the process whilst Waterworth & Luker (1990) suggest that more often patients do not wish to be involved in decision making regarding end of life care. These issues will be discussed in greater depth further in the chapter.

It is difficult to explore the concept of patient decision making on its own as so often there are many players involved – healthcare professionals (doctors/nurses), family, carers – as well as personalities, social networks, availability of different forms of care, and cultural values (Sahlberg-Bloom et al 2000). Whilst it has been already stated that the patient has a right to be autonomous and decide how they want their end of life care to be, that principle also applies to those close to and caring for that individual. Depending on the influence of these players, patient decision-making may be promoted or impeded.

The next section in this chapter will try to differentiate the different challenges for those nearing the end of life and those involved in their care.

### **DECISION MAKING CHALLENGES TO PATIENTS**

Patient participation is increasingly purported as a goal in end of life care (Clover et al 2004) but conversely it is reported that as consumers of healthcare, this patient group are often actively excluded from participation and have few opportunities in which to negotiate their real needs or desires (Small & Rhodes 2000). A number of studies (Clarke & Seymour 1999; Steinhauser et al 2000; Scanlon 2003; McCall & Rice 2004; Heyland et al 2006) have

looked at factors important to patients when approaching the end of life. Good quality end of life care should involve adequate symptom management, helping patients achieve a sense of control, relief of the feeling of being a burden and the strengthening of relationships with loved ones, the avoidance of inappropriate prolongation of dying and the usefulness of advance directives. Interestingly Heyland et al's (2006) study, although showing consistency with many of the factors mentioned above, found that patient's desire to be involved in decisions about the care or treatments received was one of the least frequently rated elements in importance in end of life care (ranked 23<sup>rd</sup> out of a possible 28 factors). There also appears to be inconsistency in what constitutes a good death to patients themselves. Steinhauser et al (2000)'s study suggests that 48% of patients wished all available treatments used, no matter what benefit, whilst 64% in the same study wishing to be prepared for death not connected to machines.

Exploring the literature around patient decision-making reveals that patients use different approaches to decisions they need to make and this depended more on their current context and environment than actual individual personality. Participation in decision making varies in a continuum ranging from total passivity to extreme activity (Sainio et al 2001) but these can be categorised simply as:-

- (a) The wait & see approach a passive approach to decision making the patient waiting for healthcare professionals to let them know what was going to happen; patients stating that they wished to gather more information; or simply procrastinating about a particular issue at that time.
- (b) The quiet accepting approach a passive approach to decision making the patient waiting for health professionals to do what needed done, believing the professional to be the "expert". There is evidence that sometimes this approach was not adopted passively but through fear of challenging the experts and the repercussions of that (Early et al 2000). At times health professionals are seen to encourage patients to play this passive role (Clover et al 2004).
- (c) The active accepting approach a passive role similar to that above but includes some active checking out by the patient so helping the understanding as to why that particular decision was so made, but not necessarily challenging the decision.
- (d) The coerced approach a submissive/compliant approach patients accepted decisions but did not understand why a particular stance was being taken and felt unable to argue. This approach seems to involve a degree of power relations between health professionals and patients, but may also include an element where the communication was not always truthful.
- (e) The negotiated approach a collaborative approach this approach often relies on a close and trusting relationship with the healthcare professional, and involves discussion around particular issues with compromises made on both the side of the individual and the professional.
- (f) The adamant approach an (expert) patient led approach where patients feel able to challenge and/or manipulate decision making to ensure a result they feel happy with.

One of the most important preconditions for participation in decision-making is information (Brennan 1997) where the patient has access to accurate and adequate information about realistic care options, although it is not always clear how much, what kind

and how to give this information. In itself the provision of information is not enough, patients must understand the information given and excessive information does not necessarily equate with increased decision making ability (Sainio et al 2001). Hagerty et al (2004) conclude in their study that the majority of patients with advanced disease want detailed prognostic information but prefer to negotiate the extent, format and timing of the information they receive. It would seem that patient participation could often be seen to be dependent on the negotiating skills between the healthcare professional and the individual patient. It can be seen that if the professional makes decisions without negotiation the patient will often develop a passive approach and conversely, if encouraged to participate, will adopt a more active role in the process. However, Clover et al (2004) conclude their study by stating that even if patient's have been independent decision makers in other life issues does not mean they wish to be involved in decisions about their own health care treatment at the end of life.

An Australian study by Butow et al (1997) found that as disease progressed, patients condition deteriorated, and end of life became more imminent, those patients wished less and less involvement in the decision making process. It could be that in leaving decisions regarding care to the health carer the patient may be choosing to avoid information related to their progressive disease state so avoiding the inevitable negative outcomes – the use of denial helping to cope with fear or anxiety (Leydon et al 2000). Discussing patient involvement in decision making as a "right" may not be acceptable to all. Patients who choose not to be involved in the decision making process or opting for the healthcare professional to act as the "expert" may in fact be actively making that decision as opposed to being passive participants in their own care. Leydon et al's (2000) study demonstrated that with advancing disease often came the difficulty of processing too much information. The consequences of decisions taken meant they then had to bear the responsibility and possible disappointment associated with that for themselves and their families i.e. by avoiding decision making, patients could avoid feeling either guilty or sorry for the consequences of their decision.

Healthcare professional's negotiating skills are the key in empowering patients to choose participation, or the equally valid choice of non-participation, in the decision making process at end of life. Singer et al (1999) agrees that if this level of care is achieved the anxiety often surrounding death will decrease eventually resulting in a more peaceful end of life. Drought & Koenig (2002) remind us that death remains, regardless of attempts to talk it through, plan or formulate it, a difficult concept to accept, approach or embrace. They state that the key for patient's to be involved in decision making around end of life care requires their ability to understand and accept that in fact there is no choice in the matter – in advanced disease, dying is not a "choice".

### **DECISION MAKING CHALLENGES TO FAMILY AND CARERS**

The philosophy of palliative care embraces the needs of both patients and those close to them but where resources are limited this may not be feasible, prompting tension between the philosophy and reality (Morris & Thomas 2001). Even when the resources are there, taking account of the needs of both patient and family caregiver can raise difficult ethical issues when considering the decisions that may need to be considered regarding care provision at the

end of life such as the use of artificial feeding. It is recognised that patients and families in crisis, or under conditions of stress, need assistance with decision making but little is known about the types of assistance in decision making that individuals with cancer and their families want (Lewis et al 1997). As more potentially life-sustaining interventions are available, more protracted states of ill health can result and the need for family involvement and surrogate decision-making, increases with decisions often becoming more complex (Meeker 2004). This combined with the shift from accepted paternalism in medical decision making to increasing self-determination and autonomy has increased pressure on close family members at an already difficult time as end of life approaches. Beyond the complexity of decision-making based on medical facts is the huge emotional burden these decisions may also carry (Weissman 2004). Families may experience feelings of guilt, anger, helplessness and loss at a time when asked to be involved in making decisions regarding end of life care of their loved one. However, the evidence would appear to question whether relatives can accurately predict preferences for end of life care for the patients they represent (Hare et al 1992; Singer et al 1999; Suhl et al 1994; Ditto et al 2001). Conversely Sulmasy et al (1998) state that prior discussion about end of life issues between patients and their family members most consistently improved the accuracy of decisions taken by family members.

Communication and connection appear to be the main components in the involvement of close family/carers in the decision-making processes around end of life care. This allows decisions to be negotiated and a sharing of the emotional and moral burden of decisions around advance care planning between the patient, the health care professional and the family which is in contrast to the idea that decisions should only be shared or discussed between health professional and patient. There may however be barriers to that communication process from the family caregivers themselves. Caregivers' ability to absorb, process and retain information given may be compromised by depression, fear, anxiety or sleep deprivation (Harris 1998). Caregivers when confronted with practical, emotional and existential issues may choose to ignore them or avoid talking about them making development of supportive strategies in the decision making process difficult or even impossible (Hudson et al 2004).

Sahlberg-Blom et al (2000) in a study looking at patient participation in decision making at the end of life from the perspective of the close relative found that patients' resignation to their illness was often coupled with family members taking over, so impeding any chance of patient involvement in the process. It was also found that the personality of both the patient and carer dictated the family involvement in decision-making, as did new roles within the relationship, a change in the manner of communication and evolving patterns of dependency. The stress levels experienced by carers where the patient chose a passive role in the decision making process were much higher than in those situations where there was collaboration between patient and carer. Good communication between key family members is again highlighted as being essential in ensuring involvement by all in the decision making process. Drought & Koenig (2002) comment that families need to look for and review cues from the patient's life to help validate decisions and provide the assurance that they are helping in the process of decision making. It is also vital that families feel able to work with the healthcare professional, are listened to, have easy access to information and are comfortable with the care their loved ones are receiving.

Work by Aranda & Peerson (2001) questioned how much caregivers were actually involved in decision making when end of life care was required and how much of the provision of home care was actually taken for granted by healthcare professionals. This view is supported by Yates (1999) who's study explored how much family members were coerced into caring roles they neither wanted nor felt capable of undertaking, but who felt powerless in expressing these feelings. It may also be possible that feelings of guilt, shame or worry about being judged if they stated they were unable to undertake the care needed allowed this coercion. Decisions therefore seemed to be a mix of desire, obligation and the influences of the situation individuals found themselves in within the family unit. As with patient participation in decision-making, family participation very much depended on the influence and support of the healthcare professionals involved. Aranda & Peerson (2001) present two different decision types in the results of their study – the natural/normal and the automatic/implicit.

- The natural/normal decision type was characterised within marital/close relationships caring for the dying spouse/partner was part of the natural continuum of a life together, desired with no sense of burden or obligation.
- The automatic/implicit decision type was also part of a continuum of ongoing care but the desire to care was often absent and the decision to care was less consciously made and often from a strong sense of responsibility.

In Heyland et al's (2006) study questioning family caregivers on what was important to them at the end of life, as with patient participation in decision making in the same study, family members did not rate decision making highly, ranking this element 11<sup>th</sup> out of a possible 15 options. What was ranked as extremely important was honest and open communication with healthcare professionals. Findings from this study indicate that far from increasing decision making power, autonomy as a result of tools such as advanced directives, can raise further paradoxes in the decision making process by family members. It would seem that quality of care at the end of life has more to do with trust, enhanced relationships within the family unit and improved communication with professionals. However many caregivers have reported a multiplicity of health professionals involved in palliative care which could potentially hinder the development of trusting relationships between caregiver and professional (Hudson et al 2004).

In contrast it is reported that no matter how good that trust or enhanced communication is, when asked about their experiences of involvement in the decision making process of a loved one, family members described a complex and difficult process characterised by struggle and uncertainty, feelings which remained long after their loved one had died (Meeker 2004).

## DECISION MAKING CHALLENGES IN THE ROLE OF THE HEALTH PROFESSIONAL

As previously discussed, attitudes to patient involvement in decision making have changed, challenging health care professionals to help the decision making process be more collaborative. Health care professionals need to acknowledge that the patients are "experts" in their own lives, need to share this with individual patients so empowering them to effectively participate in decisions regarding their own care (Clover et al 2004). For palliative care patients approaching the terminal phase of their illness, planning for end-of-life decisions should start at an early stage and patients, at least in Western countries, expect their physicians to initiate that conversation (Voltz et al 1998). The difficult process of assessing treatment preferences and judging the futility of intervention in advancing disease has become a major preoccupation with healthcare professionals who care for those nearing the end of life. The primary goal of all medical care is to restore or maintain the patient's health as far as possible, to maximise benefit and minimise harm (Hinkka et al 2002). The decision to withdraw or withhold medical treatment is much more difficult than the decision to commence or continue treatment, although from an ethical and legal point of view, any treatment that does not provide net benefit to the patient, may be withheld or withdrawn and the treatment goal switched to the alleviation of symptoms (BMA 1999). This will be discussed in detail further in the chapter.

Clinical decisions that occur near the end of life fall into two broad categories – decisions to use potentially life prolonging treatments for emergency situations e.g. cardiac/respiratory arrest; and those for situations that are non-emergency and involve treatment with an emphasis on the quality of life (Weissman 2004). Decisions' regarding treatments and comfort measures are often deemed to be easier where the disease follows a more definite path, e.g. cancer, as opposed to chronic cardiac disease where the disease trajectory and functional decline is often protracted over many years. One of the main challenges to the health care professional is recognising and acknowledging the terminal phase of any illness, generally accepted as the last three months and seen as heralding the dying process (Teno et al 2001). A number of prognostic indicators have been developed but they appear to be rarely used in clinical practice. Drought & Koenig (2002) comment that prognostic models cannot accommodate human relationships between physician and patient or the meanings of a terminal diagnosis on an individual, which may be the core reason that physicians resist their use and this will impact greatly on the patient-choice theory. Discussions around end of life care and the appropriateness of treatment options are not only dependant on the recognition that death is approaching, but also rely on the acknowledgement of goals the patient may wish to achieve. The recognition of treatable clinical depression, and the awareness of how decisions are made and by whom within the patient's family, an understanding of the patient's competence and his/her culture and belief system will also influence these discussions (Weissman 2004). Judgements about competency are rarely straightforward and involve assessments of emotional readiness as well as cognitive capacity (Seale 2006). Health professional decisions are also influenced by their own gender, culture, spiritual beliefs and personal values (Roter et al 2002).

Discussions around end of life care are complex and physicians are often involved in balancing a number of competing demands for patients, relatives, colleagues and ultimately legal requirements. Heyland et al (2000) report that the health professional influences the endof-life decision making process in many ways – initiating the discussion, framing the options and in holding or withholding information. Even when the patient wished a more active or shared involvement in the process, this was found only to occur when the professional initiated the process and valued that participation. Clover et al (2004) suggest that patient participation in end of life decision-making is dependent on the quality of relationship with the health care professional and also the skill and willingness of the health professional in engaging discussion around end of life care. Failure in achieving this may effectively remove patient's choice in any decisions.

Verhaak et al's (2000) study indicated that patients and carers were often only informed of one treatment option when being offered palliative radiotherapy in advanced disease. Other options or the possibility of abstaining from further treatments were not discussed although, on the basis of that discussion between oncologist and patient, decisions were made. This study also indicated the limited number of questions asked of and from the patients and their carers, a passivity that may be explained by patients feeling they had limited choice and were eager to grasp what could have been seen as a last chance of treating their disease. This theme of treating at all costs is further explored by de Haes & Koedoot (2003) who reported that often palliative chemotherapy was given with very little evidence of survival gain as it provided the clinicians with a sense of control in doing something actively in dealing with the patient's disease and the preferred option to watchful waiting or providing supportive care. Some physicians may have difficulty with such issues as the withdrawal of hydration or feeding, fearing that this may negatively affect the patient's comfort as death approaches, despite the lack of research evidence to support this (Weissman 2004). Even when the option of no treatment was mentioned it was often not explained further or referred to very negatively, seen as the clinician avoiding giving unwelcome news.

Nurses and allied health professionals are equally being actively encouraged to promote the inclusion of patients in decision-making about their care at the end of life (Clover et al 2004). Nurses' approaches to facilitating patients in decision making has been described as a four phase process (Bottorff et al 2000):

- Getting to know the patient critical in allowing the nurse to determine the role patients want to play in decisions about their care;
- Enhancing the patient's opportunities to make a choice tailoring opportunities that maximise patient participation without forcing participation on those reluctant to make choices or those who chose to relinquish control to others.
- Developing an openness to accept the patient's choice this may involve negotiation especially when patient safety is felt to be compromised.
- Respecting choice and making it happen even when at times this may be at odds with the choice the nurses' themselves would have made for the patients.

Although nurses generally support and value patient autonomy, in practice the nature of the hierarchical and bureaucratic nature of the organisational systems within the healthcare system often makes control strategies easier to embrace than collaborative strategies. It is also accepted that even if end-of-life decisions and care are familiar to nurses, this does not diminish the fact that they can be controversial, perplexing and distressing for all concerned (Scanlon 1996). Nursing tends to be the discipline that is most involved with cancer patients from the time of diagnosis to the time of death but there is evidence that suggests that relationships between physicians and nurses is often fraught with conflict which may lead to poor communication with patients, their families and other team members (Rodney 1994)

One other important consideration is that healthcare professionals by virtue of their training, may lack the appropriate skills to confidentially assess and evaluate both patients and caregivers needs, or indeed to deal with death and dying issues. Moreover, patients and

caregivers/family members may choose not to discuss difficulties, real or potential, following a palliative diagnosis, making it difficult for needs to be assessed by the clinician whether they are medical or nursing personnel (Hudson et al 2004).

### LEGAL AND ETHICAL ISSUES IN END OF LIFE DECISION MAKING

All health care is governed by four ethical principles, beneficence, non-maleficence, autonomy and justice (Beauchamp and Childress, 1998) and all decision-making at the end of life must take cognisance of these.

### BENEFICENCE

Beneficence means do harm, although in medical terms this is often very difficult as many of the treatments that are prescribed may do some harm, in the short and even sometimes in the longer term (Shaw 2002). The responsibility of the physician is always to seek a balance between the potential benefits and risks of any intervention being offered, particularly at the end of life when quality of life should be the paramount consideration.

### **NON-MALEFICENCE**

Non – maleficence means to do no harm either physical or psychological and charges all health care professionals' to have respect for the individual and his own assessment of what counts as harm (Gillon 1985). This principles is embedded in many professionals code of conduct.

#### AUTONOMY

The right of patients to make decisions about their own care falls under the ethical principle of autonomy, which is the capacity to make choices based on rational reflection on what we feel will promote our own best interest. Acting autonomously means acting in a way that promotes that aim. However, most health care professionals find a tension between the ethical principals of autonomy and beneficence when faced with the challenge of a patient making a decision regarding his or her care or treatment that the health care professional perceives to be harmful (Woodward 1998). However while respecting the patients right to refuse any and all treatment we have a responsibility to ensure that the patient is not suffering from "weakness of the will" because he has a negative or irrational belief in the outcome of the treatment. Health care professionals have a duty to sensitively explore the individual's reason for refusing treatment, as all that may be required is an explanation.

## JUSTICE

Finally there is the principle of justice this refers to the individual's right to fair treatment and privacy. It is this principle that demands that informed consent is sought prior to any intervention.

Every adult has the right to decide whether or not he will accept medical treatment, even if a refusal may risk permanent injury to his health or even lead to premature death, whether the reasons for the refusal appear rational or irrational, are unknown or even non-existent (Randall & Downie, 1999). In order for the individual to exercise his autonomous right to make decisions "in his own best interest' the patient's cognitive ability to make that decision must be established. This can be particularly difficult when cognitive ability may be impaired as a result of disease or treatment. In other words the patient must be deemed to be competent. This however, does not mean that they need to be rational. A competent individual may make an irrational decision. There should always be a presumption of competence and there is no requirement for the patient to persuade or convince the health care practitioner or the caregiver of their reasons. It is of no importance whether or not the healthcare professional agrees with the patient's decision, what is important is that the patient is competent to make it.

Competence to make a decision about his or her care cannot rely on having first-hand experience of all possible outcomes. No patient will have complete and perfect understanding of his or her illness, and of all the available treatments with their associated benefits and risks (Randall & Downie, 1999). The fact that a patient may change their mind later does not mean their present decision can be ignored. The code of practice regarding the assessment of competence states that 'incapacity must be judged in relation to particular matters, and not as an "all or nothing" generalisation.' Competence is specific not global and is related to a specific decision or setting (Buchannan & Brock 1989). In other words competence must be assessed in relation to the decision being made. Any assessment of competence must establish the following:

- The patient must be capable of making and communicating their choice
- They must understand the nature of what is being asked and why
- They must have the memory and ability to retain information
- They should have an awareness of any potential alternatives
- They must have knowledge of the risks and benefits of their chosen course of action
- They must be aware of their right to, and how to, refuse, as well as having full awareness of the consequences of refusal
- Whether they have ever expressed their wishes relevant to the issue when greater capacity existed
- That they are expressing views consistent with previously preferred moral, cultural, family, and experiential background. (DOH 2005)

Making decisions for patients who no longer have the capacity to make their own decisions presents a number of challenges as other people, such as family and health care professionals, must make decisions on the patient's behalf. These decision makers may have different goals and values that can lead to conflict unless managed sensitively by the

physician (Karlawish et al, 1999). If the patient is deemed to be not competent to make decisions regarding his or her treatment then the physician must make decisions in the patient's best interest. However, there should always be full discussion with family members prior to any decision being made. It has also been suggested that in this situation the physician would be advised to consult with other members of the health care team, who may bring a different perspective to the process (Arie 1996).

### WITHHOLDING AND WITHDRAWING TREATMENT

There are now clear guidelines governing the withholding and or withdrawing of treatment life prolonging treatment in the US, Europe and in the UK. In the UK the British Medical Association and the General Medical Council are the points of reference. What they all emphasize is that the primary goal of treatment is to restore or maintain health. However, if treatment fails or the adverse effects outweigh any benefit then there is no longer any justification for continuing that treatment. Treatment that is futile may be ethically and morally withheld or withdrawn and there is no legal or moral difference between withholding and withdrawing treatment (Simonds 2003). These decisions however can be difficult to make even when ethically and legally correct however, as previously discussed, some physicians may have difficulty in making such decisions (Weissman 2004). For family members a decision to withhold or withdraw treatment signals the onset of the death of a loved one and in some situations loved ones may seek to apply emotional pressure on physicians to continue with treatment. However, to do so when treatment is indeed futile is to raise false hope in family members and in some situations in the patient themselves. The healthcare professional in this situation has a responsibility to sensitively explain the situation giving assurances that all efforts will be made to alleviate any suffering.

### **ADVANCE DIRECTIVES**

The patient's right to self-determination is evidenced in the increasing use and possible questionable legality of advance directives, both written & documented oral statements, and in the ongoing current debate regarding physician-assisted suicide (Scanlon 2003). It is important to remember however, that whereas patients have no moral or legal rights to demand treatment, they do have moral and legal rights to refuse any treatment for any or even for no reason. The legality of valid advance directives regarding refusal of treatment has now been established in common or case law and physicians are expected to comply with them (Simonds 2003). Although living wills or advance directives indicating treatment preferences may be taken into consideration in a discussion assessing an individuals "best interests" they are not legally binding (DOH 2005). The Mental Capacity Act (2005) states that an advance directive refusing treatment is only binding if the following criteria are met.

• The individual making the directive was 18 or older and had been deemed competent to make his or her own decisions.

- It specifies the specific treatment to be refused and the circumstances in which it is to be refused. This can be expressed in lay terms.
- The individual making the directive has not withdrawn the decision at a time when her or she had the capacity to do so.
- The individual making the directive has not appointed, after the directive was made, an attorney to make the specified decision.
- The individual making the directive has not done anything clearly inconsistent with the directive remaining a fixed decision.

Although advance decisions can be either oral or in writing, an advance refusal only applies where it is in writing, signed and witnessed and contains a statement that it is to apply even where life is at risk (DOH 2005). Advance directives represent a means by which patients can continue to exercise their right to be autonomous in decisions regarding aspects of their care when they are no longer competent to do so (Fazel et al 1999). However it has been suggested that competent people may not always be well placed to make decisions regarding their future incompetent selves, as it is difficult for healthy people to envisage the whole range of situations that may occur (Hope 1992). Evidence suggests that patients do change their mind about aspects of their care as disease progresses, for example patients who initially indicated a wish to die at home found that as they became increasingly dependent and less able to self care they changed their mind and requested hospital or hospice care at the end of life. This was particularly true were the main carer was a son or daughter with patients expressing a reluctance to have a "role reversal", preferring to be cared for by professional carers (McCall & Rice, 2004). This then means that health care professionals and family members need to be aware that decisions may change over time and should review decisions as disease progresses and the end of life approaches.

# **Assisted Suicide**

No chapter on decision making at the end of life would be complete without some discussion of assisted suicide and euthanasia. This is a highly controversial topic but one that continues to gain momentum. Assisted suicide is the practice whereby a person, other than the patient, knowingly provides a means to the patient whereby they can take their own life (Matzo et al, 2004). Euthanasia is a deliberate action by another that leads directly to the individual's death (Gordon & Singer, 1995). Despite medical advances in treatment and the management of distressing symptoms, it is clear that these have not always fostered the dignity of the individual, quality of life, personal control or the provision of appropriate care (Scanlon 1996). A request for assisted death may be an indication of a build up of unrelieved physical symptoms or distress that may be of a spiritual or psychological nature whereby the patient feels there is no other relief and this should trigger a full assessment of the patient (Matzo et al 2004). It is possible that for some patients a request for assisted death may be their first expression of unrelieved suffering. Scanlon (1996) suggests that this should serve as a wakeup call to health care providers as this may characterize the inhumanity and inadequacies in the care of the dying and the critically ill. However, it is increasingly evident

that some individuals demand the right to determine the time and the manner of their own death, which represents the ultimate claim to self-determination (Scanlon 2003).

Proponents of assisted suicide and euthanasia argue that such requests from patients deemed competent to make the decision, should be permitted under the principle of autonomy and the right to self-determination. The principle of beneficence is also used to support this argument as these are acts that will relieve suffering. Opponents on the other hand argue human life is sacred and that the societal risk of abuse where vulnerable or frail individuals are concerned outweighs the potential for individual benefit (Gordon & Singer, 1995). What is clear is that this is an issue that will continue to be a cause of debate and discussion when discussing the end of life in patients with terminal disease. However, despite the arguments for euthanasia and assisted suicide, they remain illegal in most countries with the exception of the Netherlands and Belgium.

# CONCLUSION

Decision making at the end of life is complex and can be highly emotive. Traditional decision making roles are changing and the paternalistic medical model which previously led much end-of-life care, is having to be modified, challenging professional caregivers to make the decision making process a more collaborative one. It must also be acknowledged that decisions made while the patient is well may change as disease progresses and the patient becomes increasingly dependant. Health care professionals and family members must remain sensitive to this and facilitate changes, or discussions about possible choices, when necessary. Health care professionals must provide sufficient information about the disease and associated treatments so that patients can make an informed decision regarding their care but this must be given in the appropriate form and at the appropriate time. Ultimately, a patient has a moral and legal right to act and make decisions in what they perceive to be their best interest, and all those involved in providing care must respect this. This is particularly so when caring for those who are approaching the end of life.

### References

- Aranda S. & Peerson A. (2001) Caregiving in advanced cancer: Lay decision making. Journal of Palliative Care 17(4) p270-276
- Arie T (1996) Caring for Older People: Some Legal Aspects of Mental Capacity. *BMJ 313*: 156-158
- Beauchamp T.L., Childress J.F.(2001) *Principles of biomedical ethics* 5<sup>th</sup> Ed. Oxford University Press: New York
- Bottorff J.L, Steele R, Davies B, Garossino C, Porterfield P, Shaw M. (1998) Striving for balance: palliative care patients' experiences of making everyday choices. *Journal of Palliative Care 14*(1) p7-17
- Bottorff J.L, Steele R, Davies B, Porterfield P, Garossino C, Shaw M. (2000) Facilitating dayto-day decision making in palliative care. *Cancer Nursing* 23(2) p141-150

- Brennan M. A (1997) A concept analysis of consent. *Journal of Advanced Nursing 25* p477-484
- British Medical Association. (1999) Withholding & withdrawing life-prolonging medical treatment. BMJ Books: London
- Buchanan A, Brock D W (1989) *Deciding for Others: The Ethics of Surrogate Decision Making.* Cambridge: Cambridge University Press
- Butow P.N, MacLean M, Dunn S.M, Tattersall M.H.N, Boyer M.J (1997) The dynamics of change: cancer patients preferences for information, involvement and support. *Annuls of Oncology* 8 p857-863
- Clarke D & Seymour J (1999) *Reflections on Palliative Care*. Open University Press: Buckingham
- Clover A. Browne J. McErlain P. Vandenberg B. (2004) Patient approaches to clinical conversations in the palliative care setting. *Journal of Advanced Nursing* 48(4) p333-41
- Charles C, Gafni A, Whelan T (1997) Shared decision making in the medical encounter: what does it mean? (or it takes at least two to tango). *Social Science Medicine* 44 p681-692
- De Haes H, Koedoot N (2003) Patient centred decision making in palliative cancer treatment: a world of paradoxes. *Patient Education & Counselling 50* p 43-49
- De Haes H, Teunissen S (2005) Communication in palliative care: a review of recent literature. *Current opinion in Oncology* 17(4) p 345-350
- Ditto P.H; Danks J.H; Smucker W.D; Bookwala J; Coppola K.M; Dresser R. (2001) Advance directives as acts of communication. *Archives of Internal Medicine 161* p421-430
- Drought T, Koenig B (2002) "Choice" in end of life decision making; researching fact or fiction? *The Gerontologist* 42 p114-127
- Early B.P, Smith E.D, Todd L, Beem T. (2000) The needs and supportive networks of the dying: an assessment instrument and mapping procedure for hospice patients. *American Journal of Hospice and Palliative Care 17*(2) p87-96
- Fazel S, Hope T, Jacoby R (1999) Assessment of Competence to Complete Advance Directives. BMJ 318: 493-497
- Ferrell B, Chu D, Wagman L, Juarez G, Borneman T, Cullinane C, McCahill L (2003) Patient and surgeon decision making regarding surgery for advanced cancer. Oncology Nursing Forum 30(6) p106-114
- Gillon R (1985) Philosophical Medical Ethics. John Wiley & Sons, Chichester.
- Gordon M, Singer P A (1995) Decisions and care at the end of life. *The Lancet 346* (8968) 163-166
- Hare J; Pratt C & Nelson C (1992) Agreement between patients and their self-selected surrogates on difficult medical decisions. *Archives of Internal Medicine 152* p1049-1054
- Hagerty R.G, Butow P, Ellis P, Lobb E, Pendlebury S, Leighl N, Goldstein D, Lo S.K, Tattersall M.H. (2004) Cancer patient preferences for and predictors of prognostic information among patients with incurable metastatic cancer. *Journal of Clinical Oncology 22*(9) p1721-30
- Harris K (1998) The informational needs of patients with cancer and their families. *Cancer Practitioner 6* p39-46
- Hermsen M, ten Have H. (2005) Decision making in palliative care practice and the need for moral deliberation: a qualitative study. *Patient Education & Counselling 56*(3) p 268-275

- Heyland D.K, Tranmer J, Feldman-Stewart D (2000) End of life decision making in the seriously ill hospitalised patient: an organised framework & results of a preliminary study. *Journal of Palliative Care 16*(Supplement) p31-39
- Heyland D.K, Dodek P, Rocker G, Groll D, Gafni A, Pichora D, Shortt S, Tranmer J, Lazar N, Kutsogiannis J, (2006) What matters most in end of life care: perceptions of seriously ill patients and their family members. *Canadian Medical Association Journal 174*(5) p627-33
- Hinkka H, Kosunen E, Lammi U-K, Metsanoja R, Puustelli A, Kellokumpu-Lehtinen P (2002) Decision making in terminal care: a survey of Finnish doctors' treatment decisions in end of life scenarios involving a terminal cancer and a terminal dementia patient. *Palliative Medicine 16* p195-204
- Hope R A (1992) Advance Directives about Medical Treatment. BMJ 304:398
- Hudson P, Aranda S, Kristjanson L (2004) Meeting the supportive needs of family caregivers in palliative care: challenges for health professionals. *Journal of Palliative Medicine* 7(1) p19-25
- Karlawish J H T, Quill T, Meier E (1999) A Consensus Based Approach to Providing Palliative Care to Patients who Lack Decision Making Capacity. Annals of Internal Medicine 130 (10) 835-840
- Lewis M, Pearson V, Corcoran-Perry S, Narayan S (1997) Decision making by elderly patients with cancer and their caregivers. *Cancer Nursing 20*(6) p389-397
- Leydon G.M, Boulton M., Moynihan C, Joes A, Mossman J, Boudioni M, McPherson K (2000) Cancer patient's information needs and information seeking behaviour: an indepth interview study. *British Medical Journal 320* p 909-913
- Matzo L M, Sherman D W, Nelson-Marten P, Rhome A, Grant M. (2004) Ethical and legal Issues in End of Life care: Content of the end of life Nursing Education Consortium Curriculum and Teaching Strategies. *Journal for Nurses in Staff Development 20*(2) 59-66
- McCall K & Rice A (2004) What influences decisions around the place of care for terminally ill cancer patients? *International Journal of Palliative Nursing 11*(10) pp541-547
- Meeker M.A (2004) Family surrogate decision making at the end of life: seeing them through with care & respect. *Qualitative Health Research 14*(2) p204-225
- Morris S & Thomas C (2001) The carer's place in the cancer situation: where does the carer stand in the medical setting? *European Journal of Cancer Care 10* p87-95
- Pattison (2004) Integration of critical care & palliative care at end of life. *British Journal of Nursing 13*(3) pp 132-138
- Randall F, Downie R S (1999) *Palliative Care Ethics: a Companion for all Specialties (2ed)*. Oxford University Press, Oxford
- Rodney P (1994) A nursing perspective on life-prolonging treatment. *Journal of Palliative Care 10* p40-44
- Roter D L, Hall J A, Aoki Y (2002) Physician gender effects in medical communication: a meta-analytic review. *Journal of the American Medical Association 282* p583-589
- Sahlberg-Bloom E, Ternestedt B, Johansson J. (2000) Patient participation in decision making at the end of life as seen by a close relative. *Journal of Nursing Ethics* 7(4) p296-301
- Sainio C, Lauri S, Eriksson E (2001) Cancer patient's views & experiences of participation in care and decision-making. *Nursing Ethics* 8(2) p97-113

- Savulescu J (1995) Rational non-interventional paternalism; why doctors ought to make judgements of what is best for the patients. *Journal of Medical Ethics 3* p329-44
- Scanlon C (1996) End of life Decisions: the role of the nurse. Seminars in Perioperative Nursing 5(2) p92-97
- Scanlon C (2003) Ethical concerns in end-of-life care; when questions about advance directives and the withdrawal of life-sustaining interventions arise, how should decisions be made? *American Journal of Nursing 103*(1) p48-55
- Seale C (2006) Characteristics of end of life decisions: survey of UK medical practitioners. *Palliative Medicine 20* p653-659
- Shaw A B (2002) Two challenges to the double effect doctrine: euthanasia and abortion. *Journal of Medical Ethics* 28: 102-104
- Simonds A K (2003) Ethics and Decision Making in End Stage Lung Disease. *Thorax 58* (3) 272-277
- Singer P.A, Martin D.K, Kelner M (1999) Quality end of life care: Patient's perspectives. Journal of the American Medical Association 281 p163-168
- Small N. Rhodes P. (2000) Too ill to Talk? Routledge: London
- Spriggs M (1998) Autonomy in the face of a devastating diagnosis. *Journal of Medical Ethics* 24 p123-26
- Steinhauser K. Christakis N, Clipp E, McNeilly E, MacIntyre L, Tulsky J (2000) Factors considered important at the end of life by patients, family, physicians, and other care providers. *Journal of the American Medical Association 284* p 2476-2482
- Suhl J; Simons P; Reedy T; Garrick T (1994) Myth of substituted judgement: surrogate decision-making regarding life support is unreliable. Archives of Internal Medicine 154 p90-96
- Sulmasy D; Terry P.B; Weisman C.S; Miller D.J; Stallings R.Y; Vettese M.A (1998) The accuracy of substituted judgements in patients with terminal diagnoses. *Archives of Internal Medicine 128* p 621-629
- Teno J M, Weitzen S, Fennell M L, Mor V (2001) Dying trajectory in the last year of life: does the cancer trajectory fit other diseases? *Journal of Palliative Medicine 4* p457-464
- Department of Health (2005) The Mental Capacity Act 2005: London: DoH
- Verhaak C.M, Kraaimatt F.W, Staps A.C.J, Van Daal W.A.J. (2000) Informed consent in palliative radiotherapy: participation of patients and proxies in treatment decisions. *Patient Education and Counselling 41* p63-71
- Voltz R, Akabayashi A, Reese C, Sass H-M (1998) End of life decisions and advance directives in palliative care: a cross-cultural survey of patients and healthcare professionals. *Journal of Pain & Symptom Management 16*(3) p153-162
- Waterworth S & Luker K.A (1990) Reluctant collaborators: do patients want to be involved in decisions concerning care? *Journal of Advanced Nursing* 28(5) p1046-1052
- Weissman D. (2004) Decision making at a time of crisis near the end of life. *Journal of the American Medical Association 292*(14) 1738-43
- Woodward V M (1998) Caring, patient autonomy and the stigma of paternalism. *Journal of* Advanced Nursing 28 (5) 1046-1052
- Yates P. (1999) Family coping: issues and challenges for cancer nursing. *Cancer Nursing* 22(1) p63-71

Chapter 13

# **DECISION-MAKING AT THE END OF LIFE: CROSS CULTURAL PERSPECTIVES\***

# H. Russell Searight

Forest Park Family Medicine Residency, St Louis, MO Department of Community and Family Medicine St. Louis University School of Medicine, St. Louis, MO, USA

# ABSTRACT

Increased cultural diversity in the U.S. has raised challenges for health care law, ethics, and clinical decision making. In the United States, medical ethics emphasizes individual autonomy by encouraging patients to obtain all relevant information about their medical condition from health care providers and to use that information to choose treatment options. The federal 1990 Patient Self-Determination Act (PSDA) mandated health care institutions to discuss advance directives (AD) and to encourage patients to establish an AD to take effect if the patient is no longer able to communicate their medical decisions. Some ethnic communities did not find either type of AD, the living will or durable power of attorney (DPOA), to be consistent with prevailing cultural norms. Asian- and Hispanic- Americans were less likely than White European- or African-Americans to believe that patients should be directly informed about a terminal illness such as metastatic cancer. In these ethnic groups, families often received information about the illness, its prognosis and treatment options rather than the patient, themselves. These ethnic groups often saw informing the patient of serious illness and requiring the patient to make independent decisions as burdensome and isolating. The increased discussion of control over end of life care also highlights a legacy of distrust of the healthcare system among many African-American patients. From the perspective of decision-making models, patients and families appear to rely on arguments influenced by personal narratives to make health care decisions, while physicians are likely to rely on probability and clinical experience as a guide for developing their own recommendations. A more culturally sensitive approach would be to encourage patients to choose if they

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Medicine and Health Care*, edited by Thomas E. Lynch, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

want information about their condition, decisional control, and who they would prefer to have made their healthcare decisions.

### INTRODUCTION

The U.S. is increasingly becoming a multi-ethnic society. In 2004, almost 30% of U.S. adults and 40% of children self-identified as Black, Hispanic, Asian, American Indian, Alaska Native, Native Hawaiian or other Pacific Islander (National Center for Health Statistics, 2005). Specifically, while 65% of the U.S population is of White European background, the remainder are divided among the following ethnic groups: African-American 13%; Hispanic 13%; Asian-Pacific Islander 4.5%; American Indian/Alaska native 1.5%; with 2.5% self-identifying as bi-ethnic (United States Census Bureau, 2001).

This increased diversity challenges many of the institutionalized values and traditions of the dominant White European culture. These diverse beliefs and values include issues such as views of personal health, the role of medical care and the process and content of patients' health care decisions. Cultural values play a particularly prominent role in decision-making about terminal illness and the role of life support. The implementation of the 1990 Patient Self-Determination Act, focusing on advance directives and end-of-life care, highlighted many of these cultural differences.

# **PATIENT SELF-DETERMINATION ACT (PSDA)**

The Patient Self-Determination Act (PSDA), passed in 1990, applies to any U.S. Health care institution receiving federal funds. The PSDA requires that health care institutions provide written information about patient's rights to make their own medical decisions, refuse treatment, and establish written advance directives (AD). Several other goals of the PSDA were to increase use of ADs as well as to encourage documentation of ADs in the medical record. Patients admitted to hospitals without ADs are encouraged to develop these personal medical care guidelines which take effect if they cannot communicate their wishes.

Implemented on December 1, 1991, the PSDA reflected several ethical and legal trends at the time. First, the growing consumer movement has injected greater equality into the physician-patient relationship. Second, ethicists and health policy makers were raising concern about the use of increasingly sophisticated medical technology to maintain life --- even in the face of physical suffering and poor quality of life.

Historically, this dilemma was bought to public attention through the *Cruzan* case. Nancy Cruzan, a young woman who had sustained serious central nervous system damage resulting from an auto accident, was unresponsive to stimuli and could not communicate (Pietsch & Braun, 2000). She was maintained alive through life support for several years until her parents requested that tube feeding and intravenous fluids be stopped. Since Ms. Cruzan had not left any record of her wishes regarding the issue, the parents' request was turned down by the state. Eventually, based upon testimony of several friends, life support was terminated.

Finally, economic analyses indicated that the last few weeks of an adult's life are often associated with expensive healthcare costs. While these end-of life interventions are costly and may prolong life, the patient's quality of life during these final weeks is frequently poor. (Eersek, Barnes, Blackhall,& Koenig, 1999).

There are two types of ADs commonly used that take effect if a patient cannot communicate their wishes. A living will is a written, legal document indicating the types of interventions (e.g., CPR, medication for an infection) that the patient would desire if they could directly communicate their wishes. The second form of AD, a durable power of attorney (DPOA), assigns decision-making responsibilities to another person ---typically a family member and takes effect if the patient becomes cognitively incapacitated or unable to communicate.

# **IMPLEMENTING THE PSDA**

Despite the increased awareness of ADs in the U.S., the PSDA has not achieved the goals of substantially increasing the number of individuals with ADs. Overall, despite generally broad awareness of ADs, only 15-20% of the U.S. population has formalized a health care directive. During the early years of implementing the PSDA, several issues arose---particularly when discussing ADs with individuals from diverse ethnic backgrounds.

Discussions of ADs inherently include direct statements about serious illness and current or projected conditions of one's own death. In some cultures, direct discussion of these issues is seen as inappropriate and as conveying pessimism and hopelessness. For example, when attempting to implement the PSDA in some Navajo Native American communities, health care workers experienced considerable resistance. Further investigation found that among the Navajo, speaking about potentially adverse health outcomes was seen as increasing the likelihood of these consequences actually occurring (Caresse & Rhodes, 1995).

A second conflict emerged around the locus of decision-making. In developing instructions for a living will or appointing a durable power of attorney, patients are extending their autonomy into a future condition in which they can no longer directly express their wishes. Some cultures may not accept the pre-eminent value given to autonomy in U.S. medical ethics (Beauchamp & Childress, 1995). Several qualitative studies suggest that some ethnic communities find the emphasis on individual rights to be inappropriately isolating (Searight & Gafford, 2005). Relationships may present an equally and at times, more compelling, framework for bioethics than individual rights (Callahan, 1996). Anthropological studies indicate that not all patients expect or want full disclosure of their condition and may prefer that their families receive and manage this information. (Gordon, 1994).

The important role of the family also extends to proxy decision-making. Descriptive studies suggest that many patients may be uncomfortable with denoting a single family member as their health care proxy. Assigning this role to a specific family member may be burdensome to the individual designated or seen as isolating the individual from their kinship network and inadvertently promoting divisiveness. Members of some cultural communities, such as Asian-Americans, describe written ADs as unnecessary because the family network will automatically make decisions on the patient's behalf.

Many of these concerns received empirical support from a large study comparing White European-, African-, Hispanic, and Korean-Americans' views of end-of-life care. Blackhall and colleagues (1995) found that when compared with White European-Americans, Koreanand Mexican-Americans were more like to believe that a patient should not be directly informed of a metastatic cancer diagnosis. Additionally, respondents from these two ethnic groups were more likely to believe that family members, rather than the patient, should make decisions about life support. While Mexican- and European -Americans demonstrated greater awareness of ADs than Korean- or African-Americans, rates of AD completion were still low—40% for the elderly sample of European- and 22% for the Mexican- Americans who indicated they knew about ADs. (Blackhall, Murphy, Frank, Michel & Azen, 1995).

From carefully reviewing Blackhall et al.'s (1995) study as well as other quantitative comparisons of ethnic groups' desires for life support, hospice care, and willingness to donate organs, it is impossible to make blanket generalizations about end-of-life decisions for specific cultures. The data do indicate that there are trends within many ethnic communities about specific preferences. However, in general, when compared with White European-Americans, other cultural groups demonstrate greater variability regarding procedures for informing patients and types of health care decisions. For example, when comparing Koreanand European-Americans' preferences for directly informing patients about a metastatic cancer diagnosis, 47% of Korean-Americans indicated that patients should be informed compared with 87% for Whites. (Blackhall, et al., 1995). While ethnic groups' medical decision-making preferences often differ from the dominant culture, it is not possible or appropriate to use this information to categorize individual members of ethnic minority groups. However, both clinicians and policy makers should be aware of the greater variability of preferences and opinions among ethnic minorities as well as trends that may be more common to a specific culture (Searight & Gafford, 2005). A further qualification is that ethnic groups are not homogeneous. "Hispanic" may refer to individuals who trace their heritage to Mexico, Central, or South America. Some studies include Cuban immigrants as well as Puerto Ricans in this category as well. Customs, values and dialect may be strongly influenced by nationality. Even European Americans may differ according to their national background and religion-there are suggestions that current views---particularly around issues such as disclosure of diagnosis and the extent to which physicians dominate end of life choices differs by nationality (Miccinesi, Fischer, Paci, et al., 2005).

However, for organizational purposes, the following discussion will examine end-of-life preferences by ethnicity with a focus on four U.S. ethnic groups: African-, Hispanic-, Asianand White European- Americans. For each group, representative research findings on each culture's views of end-of-life care will be reviewed with a second section describing values and beliefs that may account for these treatment preferences.

# **AFRICAN AMERICANS**

### Research Findings

In a recent study by Torke et al. (2005)., fewer than 40% of the seriously ill hospitalized African American patients he interviewed had discussed their wished for end of life care with family members and none of them had an advance directive. In Blackhall et al's (1995) comparison, African-American respondents were similar to Whites in their belief that patients should be directly informed of their condition and make decisions about their care.

Other investigations have highlighted an important difference between Whites and African-American patients regarding life support and perceived futility of therapy among terminally ill patients. African American patients have consistently been found to desire more aggressive treatment at the end of life. This pattern is evident for elderly patients as well as for infants in neonatal intensive care. For example, among infants in a neonatal intensive care unit, 62% of African American parents agreed to limit life sustaining medical care as compared with 80% of Whites (Moseley, Church, Hempel, et al., 2004). This is a particularly relevant issue since African-American preterm infants are four times more likely to die than comparable White infants.

Similarly African-Americans patients and physicians are less supportive of physician assisted suicide and active euthanasia (Mebane, Oman, Kroonen, & Goldstein, 1999) and more likely than White physicians to recommend aggressive therapy for patients with brain damage and terminal illness (Tulsky, Cassileth, & Bennett, 1997). Among elderly nursing home residents with severe cognitive impairment, feeding tube were found to be four times more common among African American compared with White patients (Gessert, Curry & Robinson, 2001). African American patients have been found to be far less accepting of Do Not Resuscitate Orders (DNR) and are only half as likely to discontinue dialysis treatment (Leggat, Swatz, & Port, 1997) African-American patients with colon cancer were more likely than White patients to desire artificial nutrition, mechanical ventilation and cardiopulmonary resuscitation (CPR) (Mckinley, Garrett, Evans, & Davis, 1996).

## **Cultural Issues**

Among African- Americans, spirituality is an important factor in ascribing meaning to illness, coping with disease, and making health care decisions (Johnson, Elbert-Avila, & Tulsky, 2005). Compared with Whites of European background, African-Americans are more likely to regularly attend religious services, report more frequent use of prayer, and indicate a higher level of personal spirituality (Johnson, Elbert-Avila, & Tulsky, 2005). African Americans with higher religiosity were less likely to have an AD than those with lower levels of religious involvement. This high religiosity may emerge as a view that prayer is more helpful in treating illness than medical technology. For example, African- American cancer patients reported that prayer was more helpful than medication (Johnson, Elbert-Avila, & Tulsky, 2005).

Torke et al. (2005) have described seemingly contradictory views that may contribute to the lower rate of AD completion as well as requests for more aggressive treatment at the end of life among many African-Americans. On one hand, a belief in religious miracles leads to aggressive care ("Don't let those doctors talk you into giving up on daddy. Its' God that does the healing.") including doing everything possible to maintain life even in cases of unresponsive, terminally ill patients. Torke et al's qualitative interviews found that 'Hope of a Cure" was a common theme as in this patient with cirrhosis:

I could possibly have some kind of miraculous turnaround or just a liver come through or whatever...Keep me going until we find us a liver or until I wake up, snap out of it, inner strength, a bolt from heaven, whatever (Torke et al., 2005;p396):

At the same time, there is a view that death is beyond human interventions and ultimately up to God; ("I'll go when He says its time."). This perspective may contribute to a reluctance to discuss end-of life- care: "I'm trying to project the future now. That ain't my place. That's God's business so let that be "(Torke et al., 2005, p. 597). This contradiction emerges as a duty to fight the disease with all that medical technology has to offer while simultaneously recognizing that God will determine the time of death (Torke, et al., 2005). Results of qualitative interviews suggest that while there is greater comfort with oral communication, writing a formal document, such as a living will, may be offensive to many African-Americans

In the U.S., African Americans have had a long history of oppression. It has been suggested that the development of personal spirituality emerged from this history of discrimination and struggle. Spirituality provided a way of making sense of the experience as well as a comfort when recognizing that even Gods' people, the Israelites, were eventually liberated from bondage.

This perspective is reflected by many African American spiritual such as "go down Moses" also known as "Let my people Go" (Kling, 2004)::

When Israel was in Egypt's land,

O let my people go! Oppressed so hard they could not stand, O let my people go! (Kling, 2004; p. 211)

Disease as an oppressor and the ability to tolerate suffering with God's help has historical meaning in the context of the African American community's experience in the United States.

Another factor contributing to suspicion of White dominated healthcare institutions is a history of exploitative and deceptive research practices with African-American patients. During the slavery era, Marion Sims, the father of modern gynecology, developed surgical procedures on slave women who were operated upon without anesthesia. The best known medical exploitation of African American patients is the Tuskegee syphilis study in which African-American men, infected with syphilis, were followed for 40 years and never informed of the availability of penicillin when it was demonstrated to be an effective treatment. Even though it may not be always referred to by name and details may be omitted, the Tuskegee syphilis study is well know in the African-American community (Caplan, 2001).

## **ASIAN AMERICANS**

### Research Background

Historically, patients of Asian background have been particularly likely to see sharing serious diagnostic information with patients as inappropriate. In Blackhall et al..'s (1995) survey, 47% of Korean-American indicated that information about a terminal cancer diagnosis should be withheld from patients While norms are changing, recent data suggest that family-based decision making is more common in contemporary Japan than in the United

States. Japanese medical residents, while more likely to disclose a cancer diagnosis directly to a patient than their U.S. counterparts, were still more likely to inform the family instead of the patient. Additionally while 76% of Japanese medical residents indicated that they had deceived a patient about their diagnosis at the request of the family, only 18% of U.S. residents reported the same behavior (Gabbay, Matsurma, Etzioni et al., 2005). Of interest, about half of the Japanese residents expressed some degree of guilt for withholding this information from the patient---a finding that suggests that there is a movement towards greater patient autonomy in Japan. In the U.S., there does appear to be gradient of acceptance of ADs and DPOAS depending on the degree of acculturation. Among East Asian Indian communities in the U.S., rates of AD completion are much lower than for Whites---9% reported having a living will while none had a DPOA (Doorenboos & Nies, 2003).

### **Cultural Issues**

Asian culture de-emphasizes individual identity in favor of a view of the person as part of a complex social matrix. These social groups include the family-of-origin, fellow employees in the workplace, and in countries, such as Japan, even the entire country. Group decision making is common and in cases of serious illness, families take on the decision-making role. (Gabbay et al., 2005). Patients, themselves, particularly elderly family members, are viewed with respect and deference. Direct disclosure of serious illness would seem both disrespectful as well as cruel. Additionally, in these tight knit families, there is a strong theme of responsibility to one another. Views of illness and death are often bi-directional –the elderly person may feel a duty to remain alive for their adult children's well-being.

Many East Asian immigrants from India are Hindus. Hinduism emphasizes the role of karma as a determining factor in life and illness. Under the law of karma, significant life events are destined to happen. These events are predetermined by actions earlier in ones' life or in a previous incarnation. Current suffering is likely a debt that should be paid before the next incarnation (Yeo & Hikoyeda, 2000). Karma's role in end of life care has led to differing perspectives on life support and advance directives. On one hand, patients should not make the determination about the exact time of death by encouraging the use of life support to artificially prolong their lives. Conversely, others see no value in ADs since the physician has become "the tool" of karma—enacting what has been predetermined. (Doorenbos & Nies, 2003). Regardless, there is a tradition in India and other East Asian countries of familycentered rather than patient-centered decision making. In Doorenboos and Nies' (2003) survey of American East Asian Hindus, two thirds of their respondents indicated that the family would make decisions about end of life care---rendering written ADs irrelevant. In this sample, the use of ADs was low but found to be inversely related to the importance of Hindu religious beliefs ("I believe that suffering is caused by bad actions or deeds done in this life or a past life") and practices (" It is important to have 10 to 14 days of mourning") (Doorenboos & Nies, 2003).

# **HISPANIC AMERICANS**

### Research Findings

In Blackhall et al.'s (1995) study, Mexican- Americans were less likely than Whites and African- Americans to say that patients should be directly informed of a metastatic cancer diagnosis. Research on Hispanic Americans and end-of-life care is limited. However elderly Mexican-American respondents who were more acculturated and with incomes over \$ 10,000 were more likely to favor directly informing patients of serious illness. Rates of AD completion also appear to be lower as compared with African- and Europeans. Among participants in a comprehensive elder care program that included AD completion, 40% of Hispanic elders had completed ADs as compared with 80% of White- and African-Americans (Eleazer, Hornung, Egbert et al., 1996).

Hispanic Americans have also indicated less willingness to stop life support and to be maintained alive regardless of their illness severity and course (Caralis, et al. 1993). Nearly 80% of Mexican-Americans indicated agreement with the position, "Life-sustaining machines should never be stopped even if the patient appears to be dying because there is always hope of a miracle "while only 29% of European Americans endorsed this position (Blackhall, Frank, Murphy et al., 1999) While Hispanic patients may have a stronger preference for family-based decision-making than Whites (Blackhall, et al., 1995), they also are also more likely to accept a physician's judgment when there was disagreement between the family and the physician compared with Whites or African-Americans. Nearly half of Hispanics interviewed would defer to the physician's judgment compared with approximately one-third of African- and European-Americans (Caralis, et al., 1993; Talamntes, Gonzales, & Braun, 2000).

### Cultural Issues

Because there is a tradition of family-based decision-making, Hispanic patients may see ADs as unnecessary (Morrison, Zayas, Mulvihill, Baskin, & Meier, 1998). Even establishing a DPOA may be incongruent since it lists one person as the decision-maker rather than the collective extended family unit. "Familisimo," a common Hispanic value, emphasizes the family's welfare over individual well-being (Talamantes, Gonzales, & Braun, 2000). When making medical decisions for a seriously ill family member, there is often a process of group decision-making in which the patient's relatives work to achieve some consensus on a plan. The greater tendency to defer to a physician's judgment may reflect the value of "jeraquismo;" emphasizing respect for authority (Talamentes, Gonzales, & Braun, 2000). Blackhall, et al. (1999) note that that the seeming contradiction between wanting aggressive treatment while deferring to the physician is explainable by the positive regard for physicians as respected authorities. When a physician asks, do you want CPR, tube feeding, etc, Mexican Americans assumed that since the physician was raising this possibility, it was their professional recommendation and should be followed. Rather than recognizing the physician's responsibility to describe all possible options, families concluded that the physician must think aggressive measures would help the patient or they wouldn't be brought up: "To a Mexican-American patient or family member, the question, "Should we do CPR?" may be interpreted as meaning that the physician is recommending CPR when in fact the opposite is likely to be true..." (Blackhall, et al., 1999; p.1787).

# **PATIENTS OF EUROPEAN BACKGROUND**

### Research Findings

White Europeans are typically the standard against which other ethnic groups are compared. In Blackhall et al's (1995) study, 69% of European Americans reported that a patient should be told of a terminal cancer diagnosis, with 65% indicating that the patient should make decisions about life support. In contrast to African- and Hispanic –Americans, religious affiliation was associated with significant differences within the European American group. Among Protestants, 81% indicated that patients should be informed of their diagnosis and prognosis and 73% stated that the patient should be the primary decision maker. Comparable figures for non-Protestants were 61% and 59%, respectively. Jewish participants were least likely to state that patients should be informed of their condition and prognosis (52%) (Blackhall et al., 1995).

When examining perceptions of end-of-life decision-making in European countries, an interesting pattern emerges. In examining the use of life sustaining measures in contemporary Europe, the level of decisional control given to families and patients is significantly diminished in comparison to contemporary U.S. practices. In comparing Swedish, German, and Russian physicians' responses to patient vignettes, the Swedish physicians chose less aggressive treatment than the other two nationalities (Richter, Eiserman, & Zgonnikova, 2001). The German and Swedish physicians were more likely to comply with a patient's AD that included a "do not resuscitate" (DNR) order than Russian doctors. Among the Russian physicians, half of them would perform CPR regardless of the patient's wishes not to be resuscitated.

Other studies suggest differences in perceptions about the source of the decision. In the Netherlands, the physician assumes more of the decisional authority in relation to the patient's family. However, a key difference is that rather than making a decision to initiate and then stop artificial nutrition and hydration as in the US, these interventions are simply not initiated in Dutch health care institutions. For example in 1995, 23% of Dutch nursing home deaths were associated with a clear decision to forego artificial nutrition and hydration (Pasman, et al., 2004). Of interest, there were differing perceptions about who actually made the final decision. Nearly all (97%) of the physicians said that they, primarily, had made the decision to forego life sustaining measures while only slightly more than half (57%) of the families said that the physician had made the decision. About a third of the families said that they had made the final decision alone while a comparable proportion of physicians indicated that they, alone, had made the final decision (Pasman, et al., 2004).

### Cultural Background

Control and autonomy are central values for Whites of European background. ADs are likely to be seen as particularly valuable since they are a vehicle for extending autonomy even when the patient is no longer able to articulate their desires. The written AD or 'living will" as well as specific, detailed instructions for the DPOA permits continued independent direction of one's own care. Depending on others is feared and seen as an imposition as well as a burden to family members (Blackhall et al., 1999)—a theme very evident in the following request: "If I would go into any kind of coma that would last anything more than 24 h, I would say hey, that's about it …because I know the burden it sets on other people's head" (Blackhall et all., 1999, p. 1787). When cognitive functioning becomes impaired, Whites are less likely to se themselves as 'whole' and as more likely to anticipate a loss of identity (Blackhall et al., 1995). As one respondent said: ""[I don't] see what the purpose of life is if you're not mentally alive" (Blackhall et al., 1999, p. 1786).

In many European societies, physicians still are the dominant decision-maker. This may, in part, stem from limited resources but also may reflect a communitarian ethic that contrasts with the individualistic focus in U.S. health care ethics. A communitarian ethic, while perhaps not as dominant as in Asian or Hispanic communities, appears to have somewhat more influence in contemporary Europe than in the U.S. Values are not personally held but are in the public domain (Kuczewski & McCruden, 2001) and there is a greater emphasis on responsibility for the health care of the society at large than holding the individual as paramount. However, physicians in many of these societies, perhaps through their communication with patients' families, leave the family with an experience of having made a decision that the physician follows (Cassell, 2005).

# **DECISIONAL PROCESSES AT THE END-OF-LIFE**

In attempting to explain patient and family decision-making processes contributing to these differences in desires for information about one's medical condition, views of appropriate decision-makers, and wishes for palliative versus aggressive care, rational forecasting models are of minimal value While physicians rely on a combination of local norms, personally-held moral values, probability (i.e., patients in this condition have a 1 out of 100 likelihood of survival), and cognitively-held case prototypes derived from previous clinical experience, the key variables for patients and their families are less clear-cut. For families and patients, terminal illness is a situation in which cultural values, personally held moral arguments, perceived duties to other family members and narrative history are all likely to play a role.

As Hardman and Ayton (2004) note, when precise information is absent or a decision is unique and without historical precedent, usual heuristics such as probability are often not employed. but are replaced by a reliance on arguments as a basis for decisions: "Unlike the thinking tasks that participants receive in psychological experiments, such real-life situations involve different groups of people who lack hard facts but who are reasoning from different and conflicting assumptions" (Hardman & Ayton, 2004; p. 164). In qualitative investigations that probe the reasons behind decisions to seek more aggressive care or refuse to withdraw

life support, lay people are less likely to rely on probability reasoning (e.g., only 1% of patients with cancer at this stage survive) increasingly common among younger physicians educated in evidence-based medicine.

.Instead, patients and families often develop a "story" to assist in constructing arguments. This narrative approach is similar to the structure of causal models that jurors employ (Hardman & Ayton, 2004; Pennington & Hastie, 1986).

Another characteristic of ambiguous situations---where there are no clear rules to predict the outcome--- is that preexisting beliefs are held in the face of challenging, even disconfirming, information. Much of the research in the area has involved a moral issue, the death penalty. An early study by Lord, Ross and Lepper (1979) found that when persons who were pro and con on the death penalty issue were given a reading that was in opposition to their position, there was no attitudinal change. However, if given articles representing both sides of the issue, their preexisting position became stronger. In the face of arguments against their position, more arguments were generated than in response to material that was congruent with their beliefs. These arguments focused primarily on refuting the presented opposing belief. (Edwards & Smith, 1996). This process may be particularly salient to the African American experience with a White dominated medical system. The strong elements of distrust as well as the narrative accounts of the abuse of African Americans in the service of medical science are likely to lead to pronounced distrust of physicians and healthcare institutions. This historical experience, often transmitted orally in the African American community, may lead to a discounting of the medical evidence and prognostic information provided by the physician.

# **CONCLUSION AND IMPLICATIONS**

Current legal and ethical norms in the U.S have been described as a mandate requiring patients to make autonomous medical decisions. This demand may include a perceived *duty* to draw up an advance directive even when cultural values and traditions see ADs as irrelevant or even harmful. True autonomy should extend to the right *not* to have an advance directive (Volker, 2005). From the perspective of many cultures, autonomy is seen as isolating and even cruel, when foisted upon seriously ill patients. Additionally, suffering may not be seen as "unnecessary and pointless" by many ethnic communities. The choice to terminate suffering by withdrawing life support may be interpreted as inconsistent with spiritual beliefs that find meaning in suffering and view any life, regardless of its perceived quality, as valuable.

In order to respond to the diversity of values in the U.S., policies such as the PSDA should be reformulated to allow a broader range of choices. At admission to a hospital and/or prior to undergoing diagnostic tests, patients could be asked if they want to be informed of test results, their diagnosis and, if known, prognosis with and without treatment. Health care professionals may also ask patients if they would like someone else to be informed of their condition and treatment options instead of or in addition to, themselves. These policy changes would support the broader array of decision-making preferences that exist in our increasingly multicultural society.

### REFERENCES

- Beauchamp, T., & Childress, J. L. (1995). Principles of Biomedical Ethics (Fourth). New York: Oxford University Press
- Blackhall, L. J., Frank, G., Murphy, S. T., Michel, V, Palmer, J. M., & Azen, S. P. (1999). Ethnicity and attitudes towards life sustaining technology. *Social Science and Medicine*, 48, 1779-1789.
- Blackhall, L. J., Murphy, S. T., Frank, G., Michel, V, & Azen, S. (1995). Ethnicity and attitudes towards patient autonomy. *Journal of the American Medical Association*, 274, 820-825.

Callahan. (1994).

- Caplan, A. L. (2001). Twenty years after: The legacy of the Tuskegee Syphilis Study. In W. Teays & L. M. Purdy (Eds.), *Bioethics, justice, and health care* (pp. 231-235). Belmont CA: Wadsworth-Thomsom Learning.
- Caresse, J. A., & Rhodes, L. A. (1995). Western bioethics on the Navajo reservation: Benefit or harm ? Journal of the American Medical Association, 274, 826-829.
- Cassell, J. (2005). Life and Death in Intensive Care.
- Doorenbos, A. Z., & Nies, M. A. (2003). The use of advance directives in apopulation of Asian Indian Hindus. *Journal of Transcultural Nursing*, 14, 17-24.
- Edwards, K., & Smith, E.E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology*, 71(1), 5-24.
- Eleazer, GP., Hornung, C. A., Egbert, CB., & et al. (1996). The relationship between ethnicity and advance health directives in a frail older population. *Journal of the American Geriatrics Society*, 44, 938-943.
- Gabbay, B.B., Matsumura, S., Etzioni, S., & et al. (2005). Negotiating end-of-life decision making: A comparison of Japanese and U.S. residents' approaches. *Academic Medicine*, 80, 617-621.
- Ganz, F.D., Benbenishty, J., Hersch, M., Fischer, A., Gurman, G., & Sprung, C. L. (2006). The impact of regional culture on intensive care end of life decision making: An Israeli perspective from the ETHICUS study. *Journal of Medical Ethics*, 32, 196-199.
- Gessert, C. E., Curry, N. M., & Robinson, A. (2001). Ethnicity and end-of-life care: The use of feeding tubes. *Ethnicity and Disease*, 11, 97-106.
- Hardman, P. (2004). Arguments and Decisions. *Psychological Invesigations of C ompetence in Decision Making*. Cambridge, United Kingdom: Cambridge University Press.
- Johnson, K.S., Elbert-Avila, K.I., & Tullsky, J. A. (2005). The influence of spiritual beliefs and practices on the treatment preferences of African Americans. *Journal of the American Geriatric Society*, 53, 711-719.
- King, D.W. (2004). The Bible in History. New York: Oxford University Press.
- Kuczewski, M., & McCruden, P. J. (2001). Informed Consent: Does it take a village ? The problem of culture and truth telling. *Cambridge Quarterly of Healthcare Ethics*, 10, 34-46.
- Kurent, J. E. (2003). Case Presentation: Medical decision-making in hopeless situations: The long-lost son. *Journal of Pain and Symptom Management*, 25, 191-192.
- Leggat, JE., Swartz, R. D., & Port, F. K. (1997). Withdrawal from dialysis: A review with an emphasis on the Black experience. Advances in Renal Replacement Therapy, 4, 22-29.

- Lord, C.G., Ross, L., & Lepper, M.R. (1979). Biased and assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal* of Personality and Social Psychology, 37, 2098-2109.
- McKinley, E. D., Garrett, J. M., Evans, A. T., & Davis, M. (1996). Differences in end-of-life decision making among black and white ambulatory care patients. *Journal of General Internal Medicine*, 11, 651-656.
- Mebane, E. W., Oman, R F., Kroonen, L. T., & Goldstein, M. K. (1999). The influence of physican race, age, and gender on physician attitides toward advance care directives and preferences for end-of-life decision-making. *Journal of the American Geriatircs Society*, 47, 579-591.
- Miccinesi, G., Fischer, S., Paci, E., Onwuteaka-Phillipsen, B. D., Cartwright, C, & et al. (2005). Physicians' attitudes towards end-of-life decisions: A comparison between seven countries. *Social Science and Medicine*, 60, 1961-1974.
- Morrison, R. S., Zayas, L. H., Mulvihill, M., Baskin, S. A., & Meier, D. E. (1998). Barriers to completion of healthcare proxy forms: A qualitative analysis of ethnic differences. *The Journal of Clinical Ethics*, 9, 118-126.
- Moseley, K. L., Church, A., Hempel, B., Yuan, H., Goold, S. D., & Freed, G. L. (2004). Endof-life choices for African American and White infants in a neonatal intensive care unit *Journal of the National Medical Association*, 933-937.
- National Center for Health Statistics. (2005). *Health, United States, 2005 with chartbook on trends in the health of Americans*. Washington DC.: U. S.Government Printing Office.
- Pennington, N., & Hastie, R. (1988). Evidence evaluation in complex decision making. Journal of Personality and Social Psychology, 51, 242-268.
- Phipps, E., True, G., Harris, D., & et al. (2003). Journal of Clinical Oncology. *Approaching* the end of life: Attitudes, preferences and behvaiors of African American and White patientsand thier family caregivers, 21, 549-554
- Pietsch, J. H., & Braun, K. I. (2000). Autonomy, advance directives, and the patient selfdetermination act. In KL. Braun, J. H. Pietsch & P. I. Blanchette (Eds.), *Cultural issues in end-of-life decision-making* (pp. 37-54). Thousand Oaks, CA: Sage.
- Richter, J., Eisemann, E., & Zgonnikova, E. (2001). Doctors' authoritarianism in end-of-life treatment decisions. A comparison between Russia, Sweden, and Germany. *Journal of Medical Ethics*, 27, 186-191.
- Searight, H. R., & Gafford, J. (2005). Effect of culture on end-of-life decision making. *Quarterly Newsletter of the American Academy of Hospice and Palliative Medicine*, 6(4), 1-5.
- Talamantes, M. A., Gomez, C., & Braun, K. L. (2000). Advance directives and end-of-life care: The Hispanic perspective. In K. L. Braun, J. H. Pietsch & P. L. Blanchette (Eds.), *Cultural issues in end-of-life decision-making* (pp. 83-100). Thousand Oaks, CA: Sage.
- Torke, A.M., Garas, N. S., Sexson, W., & Branch, W.T. (2005). Medical care at the end of life: Views of African American patients in an urban hospital. *Journal of Palliative Medicine*, 8, 593-602.
- True, G., Phipps, E., Braitmen, L.E., & et al. Treatment preferences and advance care planning at end of life: The role of ethnicity and spiritual coping in cancer patients. *Annals of Behavioral Medicine*, *30*, 174-179.
- Tullsky, J. A., Cassileth, B. R., & Bennett, C. L. (1997). The effect of ethnicity on ICU use and DNR orders in hospitalized AIDs patients. *Journal of Clinical Ethics*, *8*, 150-157.

- United States Census Bureau. (2001) U. S. Census 2000. Washington D.C.: United States Census Bureau.
- Volker, D. L. (2005). Control and end-of-life care: Does ethnicity matter ? *American Journal* of Hospice and Palliative Care, 22, 442-446
- Waters, C. M. (2001). Understanding and supporting African-Americans' perspectives of endof life decision-making. *Qualitative Health Research*, 11, 385-398.
- Yeo, G., & Hikoyeda, N. (2000). Cultural issues in end-of-life decision making among Asians and Pacific Islanders in the United States. In K. L. Braun, J H. Pietsch & P. L. Blanchette (Eds.), *Cultural issues in end-of-life decision-making* (pp. 101-125). Thousand Oaks, CA: Sage.

Chapter 14

# FEELING RESPONSIBLE: DECISION-MAKING IN THE CONTEXT OF GENETIC RISK<sup>\*</sup>

# Holly Etchegary<sup> $\dagger$ </sup>

University of Ottawa, Ottawa, ON, K1H 8M5, Canada.

# ABSTRACT

Decision-making in the context of genetic risk is normally constructed as an opportunity for choice. For example, a person at risk for a genetic disorder could choose to accept or decline a genetic test. Since genetic risk has implications for other family members, however, perceptions of responsibility to important others could constrain some of the choices to be made around genetic risk. Drawing upon semi-structured interviews with at risk people and their family members, this chapter investigates decisions about genetic risk for Huntington disease (HD) - a fatal genetic disorder. In particular, it explores perceptions of genetic responsibility since qualitative data analysis suggested that decisions about genetic risk were often influenced by obligations to other family members. For example, some at risk people felt responsible to determine their genetic risks through testing, particularly for their at risk offspring. Responsibility to current and future partners, to plan for a future that might include HD and to communicate genetic risk to other family members also emerged as important dimensions of genetic responsibility. It is argued that perceptions of genetic responsibility constrain some of the choices of those who live with genetic risk having implications for test decisions, post-test adjustment and family relationships. Theoretically, the findings suggest that decision-making models might include measures of perceptions of responsibility to better account for their influence on decisions. At least in the context of genetic risk, a consideration of perceptions of responsibility to others could provide a richer picture of decision-making.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence concerning this article should be addressed to: Dr. Holly Etchegary, University of Ottawa, 451 Smyth Rd., Rm. 3227, Ottawa, ON, K1H 8M5. Phone: 613-562-5800 X8290; Fax: 613-562-5465; Email: hetchega@uottawa.ca.

Keywords: Huntington disease (HD), genetic decision-making, genetic testing, genetic responsibility, moral obligation

## INTRODUCTION

Ongoing research and new discoveries in molecular genetics, combined with the concomitant availability of predictive genetic testing, provide new choices for managing health (Bell, 1998). Prenatal screening (e.g., for phenylketonuria; PKU) and carrier testing (e.g., for cystic fibrosis) are often carried out in the context of reproductive decision making. Predictive genetic testing, on the other hand, is unique in that it purports to ascertain an individual's own risk of developing a genetic disease in his/her lifetime. By highlighting opportunities for health decision-making, the discourse of the new genetics presents the information derived from genetic tests as empowering (Petersen & Bunton, 2002). Genetic testing does have the potential to improve disease outcomes by providing targeted risk information for risk-management decisions. For example, a woman carrying a BRCA 1 or 2 mutation (the so-called breast cancer genes) could begin an intensive program of surveillance, including breast self exam, clinical breast exam and mammography. She also has the option of removing her breasts or ovaries as a risk-reducing measure (Rebbeck, Levin, Eisen, Snyder, Watson, Cannon-Albright, et al., 1999). Thus, as it is currently offered, genetic testing is based on a biomedical model that constructs the test candidate as independent and autonomous, rationally weighing the pros and cons of these various risk management options (d'Agincourt-Canning, 2006).

In practice, however, decisions around genetic risk are not taken in a completely autonomous or independent manner. Perceptions of responsibility driven by moral obligation to kin constrain some of the choices to be made in the context of genetic risk (Hallowell, 1999; Hallowell, Foster, Eeles, Ardern-Jones, Murday, & Watson, 2003). This chapter will demonstrate that an important feature of the lived experience of genetic risk is its construction as a responsibility, heavily infused with moral undertones of obligation to other family members. It is further suggested that even when family members are not themselves at risk (e.g., a spouse) or have tested negative, genetic risk continues to evoke feelings of responsibility, albeit toward different issues (e.g., caregiving obligations). The data are drawn from a qualitative interview study comprised of individuals with a family history of Huntington disease (HD).

### Huntington Disease (HD)

HD has an autosomal dominant transmission: It affects both men and women and, at birth, each child of a HD parent has a 50:50 chance of inheriting the disease and is said to be at risk for the disorder. Therefore, inheriting the mutation that causes HD is a random event that can be compared to the flip of a coin. *Not* inheriting the mutation, also a random event, effectively eliminates HD in subsequent generations. The identification of the specific mutation for HD in 1993 (Huntington Disease Collaborative Research Group, 1993) allowed testing by direct mutation analysis. As such, a simple blood test can now provide an

individual's HD risk information without the cooperation of other family members (required in the older linkage analysis procedure).

### Symptoms

HD is a progressive, neuro-degenerative disease whose symptoms normally include a movement disorder, personality changes and intellectual decline (Huntington Society of Canada [HSC], 2002). Therefore, the disease can encompass affect, behavior and cognition, leading to significant morbidity and early mortality in most cases. Uncontrollable jerking movements, typically called chorea, affect the trunk, limbs and face of a person with HD (SuttonBrown & Suchowersky, 2003), although as many as one fifth affected with HD suffer from muscle rigidity instead (Cox, 1999). Initial motor symptoms become worse as the disease progresses and can cause difficulty with walking, speaking and swallowing (HSC, 2002). People with HD often have abnormal gait and slurred speech, for instance. Choking is a real possibility as the disease progresses and impairs normal swallowing. Cognitive impairments can include recall difficulty, attention deficits, difficulty in decision making and eventual dementia. Emotional impairment can include personality changes such as impulsiveness, disinhibition and aggression. Other emotional symptoms can include depression, irritability and obsessive-compulsive behavior. However, symptoms and their severity vary from individual to individual. For example, some people with HD suffer from severe involuntary movements, while others are more likely to suffer cognitive or emotional impairments.

## Origin

HD was the first serious, autosomal dominant disorder for which predictive testing with DNA markers became available (Harper, Lim, & Craufurd, 2000). The genetic defect is confined to a small sequence of DNA on the short arm of chromosome 4. The gene is known as IT15, and its protein product is referred to as huntingtin. Huntingtin is thought to play an important role in the normal functioning of the nervous system; however, it is currently unknown exactly how the defective gene leads to nerve cell damage in the brain. The gene is composed of three DNA bases – cytosine, adenine and guanine (CAG) - known as trinucleotide repeats. The number of repeats is significant since HD is caused by a CAG repeat expansion.

Notably, the mutation that causes HD is almost 100% penetrant. This means that individuals who carry the mutated HD gene *will* manifest the disease in their lifetime, excepting death from some other cause before the disease manifests. The normal allele contains 26 or fewer CAG repeats, with the most common length between 17 and 19 (Potter, Spector, & Prior, 2004). Conversely, HD appears fully penetrant (i.e., the disease will manifest) when the CAG repeat expansion is equal to or greater than 40 (ACMG/ASHG, 1998). An allele size of 36-39 CAG repeats, however, can be equivocal. People bearing these repeat expansions may or may not develop symptoms of HD, and about one percent of people tested fall into this reduced penetrance category (HSC, 2002).

Therefore, some inherent uncertainty remains for individuals undergoing predictive genetic testing for HD, especially when CAG repeats fall in the 36-39 range. Further, while some alleles are 100% penetrant (e.g., 40 or more CAG repeats), even these individuals at risk for HD do not know when the illness will strike, which symptoms they are most likely to manifest or how severe those symptoms will be.

### Timing and Prevalence

Age of onset for HD is usually between 30 and 45, although it can appear in young children or in adults as old as 70 (HSC, 2002). Age of onset is associated with the length of the CAG repeat, such that the longer the repeat, the lower the age of onset (Potter et al., 2004). Regardless of age of onset, HD is a progressive disease. Complications from the disease (e.g., aspiration, infection or heart failure) typically cause death 10-30 years after disease onset (SuttonBrown & Suchowersky, 2003). There is currently no cure for HD, and limited options for treatment.

The prevalence of HD is about five to eight per 100,000 in Europe and North America; it is less common in non-European ethic groups (SuttonBrown & Suchowersky, 2003). The exact prevalence of HD is difficult to specify for a number of reasons (HSC, 2002). For example, HD is often misdiagnosed (e.g., as other dementing or personality disorders). Mortality statistics for HD are often inaccurate as cause of death is normally a secondary complication such as infection, aspiration or heart failure. As such, HD might never be recorded on a death certificate. Finally, some families are secretive about HD for fear of discrimination or social rejection. The best estimate suggests that one in every 10,000 Canadians has HD, while approximately five in every 10,000 canadians is touched in some way by HD, whether as an affected individual, friend, family member or caregiver (HSC, 2002).

There is little doubt that HD is a terrible disease, and each child of a HD parent lives at risk for the fatal disorder. How do these at risk individuals make decisions about their genetic risk?

### **Decision Making Models**

Within health and social psychology, a variety of decision making models are used to predict and explain any number of health decisions and behaviors, for example, the Health Belief Model (Jans & Becker, 1984), Protection Motivation Theory (Rogers, 1983), and the Theory of Planned Behavior (Ajzen, 1991). While there are some differences among the theories' constructs (Conner & Norman, 2005), they are similar in some resects. For example, social cognition models rest on the assumptions that 1) cognitions are the primary motivator of health behavior with less emphasis on affective or social influences on decision making, 2) decision-makers are rational, autonomous and make decisions in an independent, deliberative manner, and 3) decisions are largely static (Broadstock & Michie, 2000; Soldan, Street, Gray, Binedell, & Harper, 2000). The models have their roots in subjective expected utility theory (SEU; von Neumann & Morgenstern, 1947) which assumes that people seek to maximize utility and so will make decisions that provide the highest expected utility (Conner & Norman, 2005). Thus, decision makers are assumed to survey the possible decisional alternatives, assess and weigh the likelihood of various outcomes and choose the optimal or the "best" solution (Broadstock & Michie, 2000; Conner & Norman, 2005). There is now considerable evidence that decisions are often not taken in this manner (Broadstock & Michie, 2000; Cox, 2003; Henderson, Macguire, Gray, & Morrison, 2006; Soldan et al., 2000). Indeed, decisions are sometimes made "on-the-spot," using cues from the decision environment or internal reactions to the decision situation (Peters, Lipkus, & Diefenbach, 2006). For example, some genetic test decisions appeared to be taken with little to no

deliberation; rather, the at risk individual perceived no decision to be made (Cox, 2003; Etchegary, 2006). Instead, some at risk for HD simply "had to know" (Etchegary, 2006).

Social cognition models are quite valuable in predicting a number of health behaviors and as such, are useful for the design of behavior-change interventions (Conner & Norman, 2005). However, they are far less useful in explaining the process of decision making. Broadstock and Michie (2000) noted there was a significant lack of research exploring the decision making process in different patient contexts. It has been suggested that social cognition models might not be applicable in the context of genetic risk since they cannot accommodate the dynamic nature of the process of decision making and the gradual adjustment to genetic risk information and potential diagnosis (Soldan et al., 2000). Further, "...decision making about genetic testing, particularly presymptomatic testing, is complex, varies across individuals, and requires application of models of human emotion and behavior rather than purely cognitive decision making" (Soldan et al., 2000, p. 16). A recent study concurred, suggesting that patient decision-making in clinical genetics was a process in which patients' own styles of thinking, not all of them analytic, were the primary influence on decisions (Anderson, 2007). In addition to analytical thinking, ethical, moral, reflective, practical, hypothetical, judgmental, scary and second sight thinking all emerged as types of thinking that influenced the decision-making process. Anderson (2007) suggested that this model of decision-making was a more accurate reflection of patients' actual decision-making processes compared to the traditional cognitive-consequentialist theory of decision-making.

Broadstock and Michie (2000) further suggested that social influences in decision making had been understudied. This is a particularly relevant research area in the context of genetics since genetic risk is, by definition, family risk. Thus, decisions around genetic risk may be constrained by perceived responsibility to other family members. This chapter explores perceptions of responsibility as they relate to genetic risk for Huntington disease (HD). It will show that decisions about genetic risk are subject to social influence, particularly to feelings of responsibility to other family members.

### Responsibility to Determine and Manage Genetic Risk

One of the most frequently cited reasons for, and benefits of, genetic testing is the provision of risk information for other family members. A primary motivation for undergoing BRCA 1 or 2 testing was to provide risk information for other family members, for both unaffected women at risk of hereditary breast-ovarian cancer (HBOC; Hallowell, 1999; Clark, Bluman, Borstelmann, Regan, Winer, Rimer, & Sugg Skinner, 2000; Foster, Watson, Moynihan, Ardern-Jones, & Eeles, 2002; d'Agincourt-Canning, 2001, 2006) and for high-risk women who previously had cancer (Hallowell et al., 2003; Hallowell, Foster, Eeles, Ardern-Jones, & Watson, 2004). However, perceptions of responsibility to children and other relatives are not unique to women in the context of genetic risk. Men with a family history of HBOC underwent testing primarily for the sake of their children (Hallowell, Ardern-Jones, Eeles, Foster, Lucassen, Moynihan, & Watson, 2005; d'Agincourt-Canning, 2006). Responsibility to other family members also prompted testing interest and uptake for inherited colon cancer (Ramsey, Wilson, Spencer, Geidzinska, & Newcomb, 2003; Warner, Curnow, Polglase, & Debinski, 2006). Similarly, those at risk for HD cited providing

information for their children as a primary motivation for testing (Chapman, 2002; Etchegary, 2006; Taylor, 2004).

Taken together, this body of work reveals that many at risk people are motivated to undergo testing in order to provide risk information for other family members. This information may have many perceived uses for family members: 1) to be used in their screening and other risk-management decisions (Hallowell, 1999; Hallowell et al., 2005; Ramsey et al., 2003); 2) to alleviate children's potential anxiety about their risk (Etchegary, 2006; Hallowell et al., 2005); and, 3) to be used in reproductive decisions (Chapman, 2002; Etchegary, 2006). The latter motive was especially salient in test candidates for HD. Many people at risk for HD did not discover their family history until mid-life, well after their own childbearing was complete (Cox, 1999). At risk parents perceived the test as an opportunity to provide their children with risk information they themselves did not have (Cox, 1999; Etchegary, 2006). Those at risk for HD in Chapman's (2002) study, for example, suggested they were tested primarily for their children, providing the next generation with accurate knowledge to be used in the planning of their own families. Using case studies, Downing (2005) recently illustrated the complex ways in which responsibility was interpreted and negotiated in the context of reproductive decision-making in HD families.

The discussion so far has focused largely on responsibility to determine genetic risk. In the context of inherited cancers, at risk people also expressed responsibility to control and manage their genetic risk. For example, Hallowell (1999) showed that women at risk for HBOC felt obligated to control and manage their risk through screening adherence (e.g., breast exams, mammography) and in some cases, prophylactic surgery (Hallowell, Jacobs, Richards, Mackay, & Gore, 2001). Obligation to children, in particular, motivated riskmanagement decisions. Women were also motivated to avoid being a burden to family members or to deprive children of their mothers should they develop cancer (Foster et al., 2002). Feeling responsible not only for their own health, but also the health of other family members, at risk women encouraged family members to undergo testing and to engage in appropriate screening (Hallowell, 1999; Foster et al., 2002). Indeed, some people at risk for HBOC were expressly concerned to provide relatives with information that would allow them to make informed risk-management decisions (Hallowell et al., 2003). Relatives who did not seek testing, particularly those with children, were seen as selfish or acting irresponsibly, both those at risk for HBOC (Foster et al., 2002) and HD (Downing, 2005). This could lead to strained relationships and communication difficulties in families when some members do not want to know their risk (d'Agincourt-Canning, 2006). Rather than acting irresponsibly, some who declined testing did so out of a sense of responsibility to other family members, fearing that a positive test result would do more harm than good (d'Agincourt-Canning, 2006). Thus, test decliners may also be motivated by perceived responsibility to other family members.

Notably, perceptions of responsibility for risk management may be somewhat different in the context of HD, a fatal genetic illness with limited options for treatment. There are currently no effective preventive or screening behaviors that will arrest or impede disease progression. Thus, feelings of responsibility emerge in other areas such as reproduction, marriage decisions or future care-giving obligations.

A growing body of work suggests that the increasing availability of genetic testing may result in at risk people developing a sense of genetic responsibility (Kenea, 1994), an obligation to determine their risk and share the information with other family members. Conversely, genetic responsibility may manifest in declining the test when a positive test result is viewed as too harmful to other family members (d'Agincourt-Canning, 2006). In either context, the self is not experienced as independent, but rather interdependent (Kenea, 1994), and as such, the needs of important others are incorporated into decision-making. Concern has been expressed that such feelings of responsibility to others might constrain the autonomy and choices of some at risk people (Etchegary, 2006; Foster et al., 2002; Hallowell, 1999). Constraint can take the form of relinquishing their right not to know their risk status (Etchegary, 2006; Hallowell, 1999), constraining their risk-management options (Hallowell et al., 2001) or making less informed decisions to undergo testing that could leave people unprepared for a positive test result in the long term (Etchegary, 2006; Foster et al., 2002).

This chapter presents perceptions of responsibility as they relate to genetic risk for HD. Its purpose is to further understanding of how perceptions of responsibility are experienced, not only in the context of test decisions (Etchegary, 2006) or reproduction (Downing, 2005), but also in other domains such as future planning and caregiving. It is hoped the discussion will contribute to the understanding of the moral aspects of genetic decision-making from the perspectives of people living with genetic risk, a perspective underrepresented in discussions of ethics and genetics (d'Agincourt-Canning, 2006). It is also hoped to highlight a variable notably absent in most social cognition models, but which may aid in the prediction of some health behaviors. The study received full ethical approval from the appropriate university's Human Investigation Committee (HIC).

# Метнор

### Participant Recruitment

Participants were recruited from a medical genetics clinic and from HD support groups in an Eastern Canadian province. Three genetic counselors and two representatives of the Huntington Society of Canada (HSC) provided clients with an information package about the research. Clients were invited to contact the researcher if they were interested in participating. Recruiters contacted one person in HD families and asked him or her to inform other family members about the study. While helpful in reducing the time and effort spent on recruitment, this practice likely reduced study participation as the contacted individual may not have informed other family members about the study. It is also possible that contacted individuals may have informed only those other family members who were thought to have similar views about the impact of HD on the family, thereby possibly biasing study findings. During interviews, however, many participants revealed other family members to whom they had talked about the study. They would sometimes suggest that their relative(s) did not feel the same way they did about particular issues (e.g., reproductive choices), thus lessening the possibility that only similar family members were approached.

Genetic counselors and a HSC social worker decided who should be informed about the research. Consistent with the local ethics committee policy, those deemed as too vulnerable to participate (e.g., because of recent family death or cognitive impairment) were not invited to the study. Fourteen families were contacted by the genetics clinic, and three others were informed through HD support groups. The HSC also included an insert about the study in its newsletter, *Horizon*. Three people responded to the insert, and all three were interviewed. In

total, 24 people participated in the study, representing ten different families affected by HD. To my knowledge, there were only two refusals to participate, both owing to deaths in these families at the time of recruitment.

### Participants

In total, 14 participants had undergone genetic testing, resulting in a variety of test outcomes (see Table 1). Six participants declined testing and were at risk for HD. Four family members, not themselves at risk, also completed an interview.

Tested positive	3
Tested negative	5
Tested, intermediate gene	2
Tested, did not receive results*	2
Tested, now affected with HD	2
Family history, never tested (i.e., at risk)	6
Family member, not at risk	4
Total	24

Table 1. Test results or risk status of participants at the time of the interview

\* One result was pending, while the other participant had been tested but subsequently decided not to request the result.

Participants were approximately 46 years of age (*Range* 21 - 73), and for those who had been tested, an average of 6.5 years had passed at the time of the interview (*Range*: two months – 15 years). Three-quarters of the participants were female (n=18). Most participants were married or living with a partner at the time of the interview (n=17), and nearly all had children (n=21). All but two participants completed high school and most went on to complete college diplomas or university degrees, two at the graduate level.

### The Interviews

Interviews were conducted by the author in participants' homes or the researcher's office, with a minority by telephone between January and June, 2004. With permission, all interviews were tape-recorded and transcribed verbatim. All participants, excluding one, were provided with a copy of their interview transcript. One participant wanted only a summary report of research findings. Participants were invited to check the transcript for accuracy and to contact the researchers with any corrections, removals or additions. No participant reported any concerns regarding his or her interview transcript.

Interviews lasted from one to three hours, with the average being about an hour and 15 minutes. They were semi-structured and covered a core set of topics such as family history of HD, daily life with genetic risk and/or illness, test decisions, stigma associated with HD and healthcare concerns (see Appendix). Questions were chosen following a wide reading of the

319

literature in several fields such as health and social psychology, medical sociology and medical and clinical genetics. Questions were discussed with key informants (e.g., representatives from the local HSC chapter and the medical genetics clinic) prior to conducting interviews, thus refining the interview guide. Although questions were not confined to a specific order, all topics were covered in each interview, and participants were actively encouraged to discuss other issues they deemed important. Minimal prompt questions were used as necessary to elicit more detail (e.g., Could you tell me a little more about that? How did you feel about that?). The current analysis focuses on perceptions of responsibility around genetic risk since these figured prominently in participants' reflections on living with (risk for) HD.

## Data Analysis - Interpretative Phenomenological Analysis (IPA)

The research questions were designed to explore the meanings participants gave to their risk and how these were negotiated in daily life. IPA was thought to be an appropriate method given these aims. IPA (Smith, Flowers, & Osborn, 1997; Smith, Jarman, & Osborn, 1999) is a qualitative approach with roots in phenomenology and symbolic interactionism. It is a recent approach to empirical research in psychology and may be particularly valuable in exploring novel issues surrounding the new genetics (Chapman & Smith, 2002). IPA aims to provide a detailed exploration of how people make sense of their experiences. It is recognized, however, that the analyst's own perceptions are needed in order to make sense of the personal world being studied through a process of interpretative activity (Smith et al., 1999).

IPA follows an idiographic approach to analysis, beginning with an initial transcript and slowly working up to more general categorization across all transcripts (Smith et al., 1999). The following process applies to IPA studies with a sample of ten or more. Transcripts are read and re-read several times. With each reading, meaningful groupings of data are identified which collect together emergent themes. In this way, clusters of themes are generated for each transcript. For example, participants talked of their feelings of responsibility to their children, their partners and other family members, as well as their responsibility to plan for their futures and to communicate their risk to other family members. These groupings of data seemed to cluster together in a meaningful way to describe one lived dimension of genetic risk, that of genetic responsibility. Emergent themes in transcripts were fed back to the analysis of subsequent transcripts. In this way, connections across participant accounts were made until a final set of shared themes was identified.

The current sample is somewhat large for an IPA study, although samples of up to 48 have been reported (Brocki & Wearden, 2006). Smith et al. (1999) described the method of analysis for samples of more than ten participants (outlined above), suggesting that early coding was broader than it might be for smaller samples of ten or less. For the current analysis, there was concern to give as full and interesting an interpretation of the data as possible, rather than focus specifically on any one particular group's experiences (e.g., caregivers, test candidates). Including all participants in the analysis allowed for a maximum comparison among participants' experiences and allowed a full exploration and analysis of the entire data set. It is acknowledged, however, that the analysis might be different had each group been analysed separately. For example, had only caregiver narratives been analysed, perceptions of responsibility about test decisions may not have emerged.

Analysis required a constant shifting back and forth between (and within) transcripts to continuously compare the experiences of participants. Silverman (2000) advocated use of the constant comparative method to safeguard validity, but also argued for the "refutability principle" (p. 178) – a process whereby researchers actively sought to refute their initial assumptions about their data. This proved particularly helpful to the current analysis and was accomplished by vigilant attention to (and questioning of) interview transcripts and emerging themes. The process was facilitated by "comprehensive data treatment" (p. 180) where *all* cases of data were used in the analysis (Silverman, 2000). This suggestion fit well with IPA's method of detailed analysis of transcripts. Attention was also paid to participant accounts that seemed to contradict emerging explanations and themes. Deviant case analysis is a "long-established tactic" for improving the quality of explanation in qualitative research (Mays & Pope, 2000, p. 51). As a further check, a detailed summary report of findings/themes was fed back to participants for their review at the end of the study. No errors or misinterpretations were reported by participants.

### RESULTS

### Responsibility to Future Generations

The process of decision-making for genetic testing for HD was discussed fully in a companion article (Etchegary, 2006) and is only briefly reviewed here. It was found that for some people at risk for HD, the decision to be tested was constrained by feelings of responsibility to children. Some parents revealed that they were tested solely to provide risk information for their children. Dennis explained:

But the part is, what about their children and grandchildren? Maybe I never would have gotten tested if my [children] hadn't asked me. And they had the right to know. I didn't want to know. It was for them. (...) It came to the point where I could have said no and walked away, and they could probably have gotten tested anyway. But, why put them through it if they didn't have to? I had no other choice but do it then for the sake of my family. –Dennis, tested positive

Similarly, Jerry suggested his children could use the knowledge gained from his test result in their own reproductive decisions:

I want them to know somehow, for them to know someday, and give them the fair choice of whether, ok, do I have kids when I get married, or whenever? – Jerry, tested positive

Dennis's comment suggested that he compromised his own need of not knowing his genetic risk for the sake of his children. Thus, perceptions of responsibility to the next generation can act to limit the choice of not knowing about one's genetic risk. Serena, for example, declined the genetic test for HD. Regarding her adult children's risk, she said:

Now, in the meantime, my kids, they know about HD. They know and they know that for them, it's up to them. We told them, we talked to them, and we told them if they wanted to go get the test (pause)...I know then if they have it, I have it. If I don't have it, they won't have it.

HE: You're right. If they choose to have the testing, you will know then for yourself. Would you be ok with that?

Oh yes. I would. It's their decision. -Serena, at risk

This narrative, in particular, highlighted a distinct ethical issue that can arise in decisions about genetic testing. Should her children decide to take the genetic test, Serena could be forced to know her own risk status, raising questions about her readiness to cope with a positive test result. More broadly, these narratives about genetic-test decisions demonstrate the construction of the self as interdependent with others (Kenea, 1994). In this way, decisions were not taken in an isolated, independent manner; rather, the needs of important others were considered. As Dennis's and Serena's accounts demonstrate, this interdependent construction of the self can override personal wishes to remain ignorant of one's own risk, raising important ethical questions about test decisions taken in this manner.

The interdependent self could also be seen in relation to family members other than children. Kathleen, for example, explained why she postponed her genetic testing for over a year after discovering HD was a part of her family:

...I just couldn't at the time. I said to myself, I just couldn't because I think I had went through a lot mentally. And I didn't think I could handle another. And I didn't want to put it on my sisters and brothers either. Because, they had to take Mom [diagnosis of HD], then they had to take the news [positive test result] of (names sibling), then take the news of (second sibling), and then, turn around and take the news with me? I thought it was too much for me and for them to handle. –Kathleen, tested, intermediate gene

Two of Kathleen's siblings had tested positive for the HD mutation, and she perceived the knowledge of her own test result as "too much" for her and her siblings to cope with at the time. As such, she postponed her testing for over a year. In this way, Kathleen assumed responsibility for the emotional well-being of her siblings.

Perceptions of responsibility to the next generation might also affect social attitudes toward at risk individuals. Some social others expected those at risk for HD to determine their genetic risk and act morally in accordance with this knowledge. A participant who tested negative, for example, said:

My [relatives], they continue to have children. To me, who in their right mind would do this if you knew you were going to inflict such terrible things on your children? If you don't know, I suppose that's different. If you do know it's in the family, I just can't see it.

Downing (2005) also revealed how an awareness of the family history of HD created expectations in family members that the genetic risk would be determined and factored into reproductive decisions. These examples highlighted the responsibility felt to future

generations, a common finding in the context of testing for HD (Cox, 1999; Downing, 2005; Etchegary, 2006), perhaps owing to the progressive and fatal nature of the illness. However, Hallowell et al. (2005) reported similar findings in a sample of men with a history of HBOC. For example, two male participants suggested that at risk people had a duty to determine their cancer risks and to avoid having children if they were confirmed to be mutation carriers.

Genetic responsibility in the context of HD revolved not only around reproduction, but also, partners, future care-giving obligations and family communication. In what follows, I wish to highlight these other domains of genetic responsibility, thus revealing additional moral aspects of genetic decision-making in this context.

### Responsibility to Partners

Participants with current partners expressed their unease with their partners having to provide extended care for them if (or when) they developed HD:

I made the decision years ago, a couple of years ago, that when I get sick, I'm not going to make my husband have the total responsibility of my health. I don't want him to. I know it's for better or worse, but I don't think he should have to give up his life for me. (...) So, when I start to get sick and can no longer look after myself, I want to go in a home. –Kathleen, tested, intermediate gene

Similarly, Lori spoke about giving her partner the option of 'getting out:'

Because I had already gone through a lot of health problems...And I said, you know, he's just had enough. (...) And I said to him when I found out, 'You know, if you want to go, I totally understand.' I said, 'I totally get it. You've done more than enough.' –Lori, tested, waiting for results<sup>1</sup>

Neither Kathleen's nor Lori's partner indicated he would leave and both suggested they would provide future care for their partners should that be necessary. However, these passages were illustrative of the perceived personal responsibility for one's future care and a desire to avoid burdening one's partner in the future.

Novas and Rose (2000) suggested that the focus on the genetic basis of disease "creates an obligation to act in the present in relation to the potential futures that come into view" (p. 486). Later in the interview, Lori explained how discovering she was at risk for HD prompted her partner to think about their potential future:

But it did a wonderful thing for my marriage, finding out that I was at risk. My husband, he has high blood pressure and there's a lot of heart disease in his family, so he's just gone off to the Happy Heart program next Monday.

HE: Oh, is that right?

<sup>&</sup>lt;sup>1</sup> Lori has since received a negative test result.

Yeah, and you know, trying to get into the exercise and trying to quit smoking and going for all these tests so that he can look after me. And he's said out loud for the first time, you know, that he thinks that maybe he may outlive me – which he never thought ever would happen. –Lori, tested, waiting for results

Thus, the discovery of her risk leads Lori and her husband to discuss future caregiving possibilities. Single participants also expressed a responsibility to future, hypothetical partners:

Before I get married, before I decide to have children, I will definitely test. I don't think it's fair knowing that you might have something like...your husband will be completely – well, if they love you, they have to take care of you right up until you die. And that's 15 years, or about that. That's a long time to watch someone you love die. I think anyone that's going to spend the rest of their life with you has the right to know. –Cheryl, at risk

It is Cheryl's construction of the self as interdependent (Kenea, 1994) that underlies her perception of genetic responsibility:

If I'm by myself, it all comes down to me. But say if I was older and had a husband right now, I'd ask you to talk to him. Because you're affecting other people, and you have to take that into consideration. –Cheryl, at risk

For participants with partners, and even for participants without, there was a perceived responsibility to inform them of the potential future of HD. Similar results were reported in interview and case studies with people at risk for HD (Chapman, 2002; Cox, 1999; Downing, 2005; Taylor, 2004). When the self is constructed as interdependent (Kenea, 1994), decisions around genetic risk are not taken in isolation. Rather, the needs of important others (e.g., children or partners), but also hypothetical others (e.g., potential partners) fuel a sense of genetic responsibility.

When the discovery of HD comes out of the blue, *unaffected* partners of at risk persons are faced with the knowledge of impending disease in their spouse and possibly their children. These narratives too, revealed a sense of responsibility to the partner and to children, although the uncertain future was terrifying:

That's very hard. Thinking about that. I suppose for me, I'm not sure what's going to come. I don't know if I'm going to be looking out to [partner]. I'm afraid that my children may come down with something and I'm going to be the one who's going to be looking out to everybody. That's a very scary thing for me. –Daphne, partner tested positive

Two participants testing positive in the current study recounted marriage breakdowns and eventual divorce. Both suggested their partners left because they did not want to "deal with" HD. Similarly, Tony recounted difficulties early in his marriage while he was caring for his affected father:

My Dad was home with HD, probably a year or two. (...) He was becoming incontinent in his urine and feces. Myself and (names wife) were the only ones home at that time. My

sister in [city], my brother back and forth. When he was home, he was causing all kinds of difficulty and problems in the family. (...) My marriage sort of went south. She couldn't handle the pressure anymore. So that went south for a few weeks. I finally managed to get Dad placed at [care facility]. So, I was back and forth there and taking him on the weekend, and hope to God he doesn't become incontinent with his feces on that weekend. –Tony, tested negative

Despite these difficulties, Tony's marriage remained intact. And recall Lori's narrative above which suggested there could be some positive effects on a marriage subsequent to discovering genetic risk.

## Beyond Partners – Responsibility to other Family Members

There was also evidence in participant narratives of responsibility to other family members, especially affected members. Stacey, for example, perceived a responsibility to help family members who were not lucky enough to escape the family legacy:

The time is going to come when I am going to have to go down and deal with this and help them, and they are going to need a major load of help in the next few years coming up. –Stacey, tested negative

Beyond responsibility to (and felt by) unaffected partners, perceived responsibility for future care was most poignantly expressed by caregivers, especially those caring for their children. Caregivers feared their increasing age and ability to continue to provide care for their affected family member(s), in two cases, an adult child. Their narratives were illustrative of the anxiety invoked by acknowledging they will not be able to provide care indefinitely. Laura was adamant that her affected family members would always be with her:

You know, I don't know where I'm going. I try not to think about where I'm going. I just wait for tomorrow to come. I only know that right now, they'll always be with me. For as long as I can do it. And when I can't do it, I'll have somebody here to do it. But they'll still be with me. At home. I know they are cared for when they're with me. –Laura, caregiver

Laura's entire narrative was reflective of her obligation to her affected family members. Similarly, Marjorie suggested that she feels *she* is the one who needs to care for her child, despite the assurances of her other, unaffected children:

That's the worst I find - [names affected child] is young, and I'm old.

HE: So, your concern is for later on down the road?

Yes. Now, they [unaffected children] tell me not to be concerned. No matter what happens, [affected child] will be taken care of. I know that. But, of course, I think *I* need to do it. –Marjorie, caregiver, emphasis in original

For parents of affected children, there was a strong sense of responsibility to provide athome care. This perception of responsibility could be painful as caregivers faced increasing age and sometimes poor health of their own.

Conversely, at risk children felt responsible for providing future care when their parent eventually becomes affected with HD. The idea of caring for a parent like a child was distressing and frightening:

He's my father, and no matter what happens, I should be there to take care of him. But I just don't think I could change him, and bathe him and change his diaper, to give him the proper care he would need. I don't think I'd be able to do it. The whole bathing them and taking care of them like a baby, I can't picture ever seeing him like that. -Roxanne, at risk

Her potential future identity as caregiver to her parent was incomprehensible to Roxanne:

Growing up, your parents are the ones who take care of you; you're not supposed to have to do that stuff. It shouldn't have to be the tables turned, and you be the one to be changing their diapers. –Roxanne, at risk

Some at risk children also recognized that future care-giving obligations might constrain some of their life choices:

Things aren't the way they used to be either. Kids...there was a time when you moved a block down from your parents and you could always be there for them. I want to be successful, and I don't know if it's always going to be possible for me to be right next to my Mom and be right there for her. So that's scary too. I'm hoping, I mean obviously I want to have all sorts of hope in science that they will come up (pause)... but it's really difficult, you know what I mean? –Cheryl, at risk

These excerpts highlight the fear at risk children can experience at the thought of the future role reversal with their (then) affected parent. At that time, at risk children may have to provide the parenting role. Despite this threat to their current identity and the recognition that future choices may be constrained, at risk children still spoke of their obligation to parents. As Roxanne said, "I *should* be there to take care of him." Thus, the self is experienced as interdependent and the (future) needs of their parents may be incorporated into at risk children's decision-making.

### Responsibility to Plan for the Future

Perceived responsibility to plan for the future was particularly salient in the narratives of tested participants. Jackie said:

If you did have it [the HD mutation], you want to manage the rest of your life - whether it's a living will or do not resuscitate order or anything else. I didn't want to leave that to somebody else to have to decide. –Jackie, tested negative

Similarly, Lori suggested:

The first thing we did was, the day after I found out, I made my will. You know, power of attorney and what I wanted if things turn bad. (...) So for me, the future is having my affairs in order, having my directive, you know, all that stuff done. –Lori, tested, waiting for results

There was a perceived responsibility to have future affairs settled before the manifestation of HD impeded the ability to do so. David noted some decisions that might have been left to the future were taken earlier in light of his having HD, particularly those about finances:

[Partner] is only going to work for 20 years and then she is going to take care of me, so we know she's not going to have enough insurance with her work. So now we have to plan to put money away for her retirement because she was going to need 30 years of retirement after that. So we have to plan for that. –David, affected with HD

### Responsibility to Communicate Genetic Risk

It has already been noted that many participants expressed a responsibility to communicate their risk for HD to current and future partners. Some participants also expressed a responsibility to inform other, extended family members of their potential risk. Julie recalled how she provided a copy of the confirmation letter from the genetics clinic to extended family:

So we sort of picked one person in each family and said that, here is a copy for everybody. I sort of felt that each person needed to have their own copy of the letter, and I felt that that was our responsibility. Some of them hated us for that because it changes your life. –Julie, at risk

This account was notable since it highlighted how rifts occurred in families when some branches did not want to acknowledge the family history of HD. Similarly, Hallowell (1999) found that some women at risk for HBOC were prepared to compromise relatives' rights of not knowing about their genetic risk. Thus, perceptions of genetic responsibility evoked a sense of duty to disseminate genetic-risk information within the family, despite the difficulties involved (Hallowell, 1999).

Other participants spoke about their responsibility to inform children of their risk, replicating research with other HD study populations and women at risk for breast-ovarian cancer (e.g., Forrest Keenan, Simpson, Wilson, van Teijlingen, McKee, Haites, et al., 2005; Hallowell, 1999). In the current research, most participants with children fully endorsed their right to know about their risk and had spoken with their children about the family history of HD. There were only two exceptions to this: One exception involved parents of young children. These participants suggested their children were still too young to handle such news. All confirmed, however, they would talk to their children as soon as they were old enough to understand the family history of HD. There was an overriding perception of responsibility to protect young children from such devastating news. At the same time, however, parents in the

current research were adamant that children should know about the family risk before critical life junctures (e.g., marriage).

In a second exception, a participant waited to inform an adult child of the family history of HD since the child was already coping with a painful life event. Concealing the family history was stressful, and even though Shirley knew the news would be devastating, she perceived a responsibility to inform the child of her own risk:

I was living with this and keeping it so close to my chest. I was so overwrought and overwhelmed (...). She had the right to know. I felt horrible having to tell her, but I knew I couldn't not tell her. –Shirley, caregiver

This passage may inform the debate surrounding whether clinicians' have the right (or the duty) to share genetic-risk information with other family members beyond the test candidate. In the UK, for example, the Nuffield Council on Bioethics (1993) recommended that medical confidentiality might be broken and clinicians could intervene in situations where a test candidate does not want to inform other family members for whom the test result may have implications. Richards (1998) argued against this practice on the grounds that test candidates might have valid reasons for not passing on genetic information to kin, and intervening could have unpredictable consequences for family relationships. In Shirley's situation, the at risk child was coping with the death of her unaffected parent. Shirley explained that she needed to give the child at least a little time to cope with this painful event before revealing the confirmation of HD in the family, fearing the child's ability to cope with the genetic information. Similar judgments about family members' coping abilities and a desire to protect relatives from distress have been reported as motivators of nondisclosure (Clarke, Richards, Kerzin-Storrar, Halliday, Young, Simpson, et al., 2005; d'Agincourt-Canning, 2001). Such a specific situation as Shirley's reminds us that genetic risk (and the decisions it provokes) is experienced within the context of a full life. Clinicians may have no way of knowing what other life events are happening during the testing process, nor can they anticipate the consequences for family relationships if they were to intervene (Richards, 1998).

## CONCLUSION

Health is constructed as a matter of personal responsibility in the current risk society (Beck, 1992), so much so that illness can be perceived as a moral failing (Galvin, 2002). Individual responsibility for health is espoused not only for voluntary health risks (e.g., smoking), but is also apparent in discourses about the new genetics (Petersen & Bunton, 2002).

Participants in the current research described themselves as having a responsibility to their families, not only to determine their genetic risk, but also to provide knowledge for future generations and to avoiding burdening current and future partners. Study participants also noted their responsibility to plan for their futures and to communicate their genetic risk to children and other family members. These findings suggest that some people at risk for HD did not construct the self as independent, rationally making choices that affected only

themselves. Rather, in the context of decision-making about genetic risk, the self was constructed as interdependent (Kenea, 1994), and the needs of important others were recognized as an integral part of the self. In this way, some at risk people were empowered by genetic testing since it allowed them to construct themselves as moral agents, taking responsibility not only for their own, but also others', genetic risk (d'Agincourt-Canning, 2001; Hallowell et al., 2003).

However, perceptions of genetic responsibility raise moral aspects of decision-making that have the potential to be anxiety-provoking. Novas and Rose (2000) suggested:

...the new genetics also links up with contemporary practices of identity. It operates in a political and ethical field in which individuals are increasingly obligated to formulate life strategies, to seek to maximize their life chances, to take actions or refrain from actions in order to increase the quality of their lives, and to act prudently in relation to themselves and to others (p. 487).

Perceptions of what constitutes acting "prudently" in relation to others raise questions about whether genetic-test decisions are purely autonomous and voluntary. This is important since ethical principles that guide genetic counseling suggest that test decisions should be completely voluntary (Nuffield Council of Bioethics, 1993). However, findings of the current research and others (Cox, 1999, 2003; Hallowell et al., 2003; 2005) suggest that family obligations and perceptions of responsibility to kin may undermine an individualistic conception of autonomy in genetic-test decisions (Hallowell et al., 2005). In contrast, Hallowell et al. (2005) suggested that decision-making about genetic testing might be seen as a family affair, rather than an individual choice. As such, rather than trying to eliminate family influences per se, it will be important to help test candidates "acknowledge the extent to which their actions are born out of a sense of responsibility or explicitly influenced by their relatives or indeed, others factors" (p. 215).

Similarly, I am not suggesting that test decisions motivated by responsibility to other family members are *necessarily* involuntary or in some way not autonomous. It is worth noting that none of the participants in the current research suggested they were coerced in any way to accept or decline the genetic test for HD. As d'Agincourt-Canning (2006) noted:

To view all obligations or responsibilities as constraints on a person's autonomy is to neglect the social and moral fabric of people's lives. What people seek from testing will be moderated by self-interest as well as social and family relationships, including responsibilities toward others (p. 114).

That said, however, perceptions of responsibility to others do suggest it is important that the *personal* implications of genetic testing or risk-management decisions be thoroughly explored prior to testing. Left unexplored, it could be that some at risk people will be unprepared for the potential distressing effects of a positive test result (Foster et al., 2002). This might be particularly important in the context of a fatal illness such as HD. If, as some participants suggested, they determined their risk status out of perceived moral duty to kin, one wonders what effect a positive test result will have on their psychological well-being. How do these people cope with unwanted genetic knowledge? An important question for future research is whether a different kind or amount of psychological support is needed for these individuals. Longitudinal studies that measure genetic responsibility and its effect on subsequent coping and psychological well-being would help address this important clinical question.

When the meaning of genetic risk resides in responsibility to others and to self, there are difficult life choices to be negotiated. Younger participants who had not yet married worried about how or when to tell a future partner about their family history of HD. Married participants worried about being a burden to their spouse. Partners and parents of at risk participants and those affected with HD worried about their ability to provide extended, long-term care. Children of at risk or affected parents worried about future caregiving obligations.

Partners of HD test candidates have received relatively little research attention in predictive testing programs (Quaid & Wesson, 1995; Tibben, Timman, Bannink, & Duivenvoorden, 1997). Quaid and Wesson (1995) reported that spouses were more depressed than their at risk partners at baseline. Spouses were also found to have the most difficulty coping with the impending threat of HD (Sobel & Cowan, 2000; Tibben, Frets, van de Kamp, Niermeijer, Vegter-van der Vlis, Roos, van Ommen, Duivenvoorden, & Verhage, 1993). The difficulty was especially pronounced in partners with children (Tibben et al., 1997). As Daphne suggested, partners of those at risk for HD can perceive the future as uncertain and frightening. Changing roles in the relationship (e.g., from spouse to potential caregiver) can lead to marital distress. As in the current study, marriage breakdowns have been reported following genetic testing for HD (Codori & Brandt, 1994; Decruyenaere, Evers-Kiebooms, Cloostermans, Boogaerts, Demyttenaere, Dom, et al., 2004; Sobel & Cowan, 2000). Recall, however, that Lori's discovery of her genetic risk motivated her partner to take better care of himself (e.g., smoking cessation) in preparation for his potential caregiving role.

Conflicting findings have been reported in this area with most studies finding no adverse effects on the marital relationship, regardless of test outcome (Richards, 2004). In fact, Richards and Williams (2004) observed an increase in the level of marital adjustment in carrier couples. It is clear that more research is needed about the impact of (risk for) HD on close relationships. For example, relatively little is known about the extent to which genetic risk affects choice of marriage partner (Richards, 1998). Anecdotally, a collection of personal accounts revealed that marriage normally proceeded when a potential partner was told about the family history of HD (Marteau & Richards, 1996). However, when the information was revealed subsequent to the marriage, partner resentment and negative effects on the marriage were recounted (Marteau & Richards, 1996; Richards, 2004). These findings raise questions about whether spouses of at risk (or affected) partners may need psychological or practical support, not only during the testing process, but as the age of onset approaches in their partner. Future research that explicitly focuses on partners of those at risk for (or affected with) HD could explore these issues.

Study findings also suggest that parents and children may have differing perceptions of responsibility. Parents worried about having passed their risk to their children, and felt responsible to determine their own risk in order to provide at risk offspring with accurate risk knowledge. Children of at risk parents were uneasy about the possibility of having to care for their parent in the future. It is suggested that future caregiving obligations threatened children's current identity as their parent's child, to be cared *for*, rather than the one providing care. Becoming the caregiver for a parent represented a role reversal that generated worry in the present. Thus, consistent with the notion of an interdependent self (Kenea, 1994), children of at risk parents incorporated the (future) needs of their parents into their own decision-

making, despite the threat to their identity. Genetic risk is family risk, and the subsequent construction of the self as interdependent may drive perceptions of responsibility and moral duty to kin in relation to genetic risk.

Study findings also raise questions about the experience of caring for a person affected with HD. Caregivers in the current research revealed perceptions of responsibility to care for their affected relative(s). However, they reluctantly admitted they would not be able provide care indefinitely. Limited research has explored the practical and psychological consequences of caring for a person affected with HD (but see Dawson, Kristjanson, Toye, & Flett, 2004 for an exception). The caregiving role in the context of HD is particularly stressful, not least because of the variable nature of the illness and the unpredictability in symptoms and their severity (Dawson et al., 2004). Findings of the current study suggested that feelings of caregiving responsibility were stressful, for both current caregivers and for those anticipating future caregiving obligations. Indeed, Skirton and Glendinning (1997) found that almost 20% of caregivers of people with HD suffered from a stress-related illness. These findings all suggest that carers of people with HD may need both practical and psychological support. It has been suggested that carer education and counseling be tailored to the needs of this special population (Dawson et al., 2004). Future research that takes HD caregivers as its focus would assist in determining exactly what type of support is needed by this special population. Such research would also likely assist public health agencies or other healthcare organizations in planning and policy decisions regarding home-support programs for families affected by HD.

Study findings challenge the value and current use in clinical genetics of an overly individualistic, analytic model of decision-making (Anderson, 2007). Such a model fails to properly acknowledge social influence on health decisions. Thus, study findings may highlight an important, but neglected, variable for social cognition models of health decision-making: perceived responsibility to important others. There is no reason to believe that perceptions of responsibility are relevant only for decisions about genetic risk. As Cox (2003) noted:

Choices are always hedged in by constraints; we are not free to decide upon just any course of action nor are we ever positioned in such a way that we can see what the full range of choices might consist of. As mothers and daughters, fathers and sons, sisters, brothers, aunts, uncles, cousins, spouses, life partners and friends, we exist in and through our social and familial ties with others (p. 262).

Thus, obligation to other family members could also be important in a variety of other health decisions such as smoking cessation, dietary choices, exercise behavior or safe sex practices. These health behaviors have been well studied with social cognition models, although the proportion of variance accounted for in health behaviors is typically low to moderate (Conner & Norman, 2005). It is suggested that future research using social cognition models to predict these sorts of health behaviors include measures of perceived responsibility to others. It is quite possible the addition of this variable could improve the prediction of health behaviors or decisions, thereby providing another avenue for behavior-change interventions.

#### Study Limitations

The cross-sectional design precluded an examination of how perceptions of responsibility might change over time. Arguably, perceptions of responsibility could lessen or heighten with the unfolding of family life (e.g., with the birth of a child or as the age of onset approaches). Prospective, longitudinal designs would assist in charting changing perceptions of responsibility and their effects on such domains as test decisions, risk-management decisions and coping. The data must be viewed with caution since participants were mostly female and highly educated, largely excluding the voices of males and at risk people with less education. Thus, the research is not generalizability is not the goal of research utilizing IPA. Rather, the goal is to capture how individuals perceive and respond to their experiences, highlighting the value of each particular case. It is notable, however, that some of the moral aspects of genetic decision-making revealed in the current study correspond with studies of other risk populations, notably HBOC and inherited colon cancer (Hallowell, 1999; Hallowell et al., 2003, 2005; Warner et al., 2006).

Despite these limitations, the current study is useful in that it exposes responsibility as a central factor in decisions around genetic risk, including reproductive and marriage choices, test decisions and family communication. As such, it revealed several moral aspects of genetic decision making from the perspective of those living with genetic risk. Future research might now explore the long-term effects of perceptions of genetic responsibility, particularly on psychological well-being and coping.

## ACKNOWLEDGMENTS

I am grateful to participants for their patience and trust. Without their generosity, such research would simply not be possible. I also thank representatives from the medical genetics clinic and the local branch of the Huntington Society of Canada for their help with recruitment. The research presented in this chapter was generously supported by a doctoral fellowship to the author from the Newfoundland and Labrador Centre for Applied Health Research and the Department of Psychology, Memorial University, both in St. John's, NL. During the writing of this article, Holly Etchegary, PhD was supported by an interdisciplinary postdoctoral fellowship from the Gap-Santé research group and a Canadian Institutes of Health Research Interdisciplinary Capacity Enhancement Team grant (GeneSens), both at the University of Ottawa. The chapter is an expanded version of the article, "*They had the right to know*." Genetic risk and perceptions of responsibility, H. Etchegary & K. Fowler, *In press* (March 2007), Psychology & Health, a Taylor & Francis journal. Psychology & Health is available online at http://journalsonline.tandf.co.uk/.

#### REFERENCES

ACMG/ASHG. American College of Medical Genetics/American Society of Human Genetics, Huntington disease Genetic Testing Working Group (1998). ACMG/ASHG

statement: Laboratory guidelines for Huntington disease genetic testing. *American Journal of Human Genetics*, 62, 1243-1247.

- Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior & Human Decision Processes, 50, 179-211.
- Anderson, G. (2007). Patient decision-making for clinical genetics. *Nursing Inquiry*, 14(1) 13-22.

Beck, U. (1992). Risk Society. London: Sage.

- Beeson, D. (1997). Nuance, complexity, and context: Qualitative methods in genetic counseling research. *Journal of Genetic Counseling*, *6*, 21-43.
- Bell, J. (1998). The new genetics in clinical practice. British Medical Journal, 316, 618-620.
- Broadstock, M., & Michie, S. (2000). Processes of patient decision making: Theoretical and methodological issues. *Psychology & Health*, 15, 191-204.
- Brocki, J., & Wearden, A. (2006). A critical evaluation of the use of Interpretative Phenomenological Analysis (IPA) in health psychology. *Psychology & Health*, 21 (1) 87-108.
- Chapman, E. (2002). Ethical dilemmas in testing for late onset conditions: Reactions to testing and perceived impact on other family members. *Journal of Genetic Counseling*, 11 (5) 351-367.
- Chapman, E., & Smith, J. (2002). Interpretative phenomenological analysis and the new genetics. *Journal of Health Psychology*, 7 (2) 125-130.
- Clark, S., Bluman, L., Borstelmann, N., Regan, K., Winer, E., Rimer, B., & Sugg Skinner, C. (2000). Patient motivation, satisfaction and coping in genetic counseling and testing for BRCA1 and BRCA2. *Journal of Genetic Counseling*, 9 (3) 219-235.
- Clarke, A., Richards, M., Kerzin-Storrar, L., Halliday, J., Young, M., Simpson, S. et al. (2005). Genetic professionals' reports of nondisclosure of genetic risk information within families. *European Journal of Human Genetics*, 13, 556-562.
- Codori, A., & Brandt, J. (1994). Psychological costs and benefits of predictive testing for Huntington's disease. *American Journal of Medical Genetics*, 54, 174-184.
- Conner, M., & Norman, P. (2005). Predicting health behavior: A social cognition approach. In M. Conner & P. Norman (Eds.), *Predicting health behavior* (pp. 1-27), Berkshire: Open University Press.
- Cox, S. (1999). "It's not a secret but..." Predictive testing and patterns of communication about genetic information in families at risk for Huntington disease. *Unpublished doctoral dissertation*, The University of British Columbia.
- Cox, S. (2003). Stories in decisions: How at-risk individuals decide to request predictive testing for Huntington disease. *Qualitative Sociology*, *26*, 257-280.
- d'Agincourt-Canning, L. (2001). Experiences of genetic risk: Disclosure and the gendering of responsibility. *Bioethics*, 15 (3) 231-247.
- d'Agincourt-Canning, L. (2006). Genetic testing for hereditary breast and ovarian cancer: Responsibility and choice. *Qualitative Health Research, 16* (1) 97-118.
- Dawson, S., Kristjanson, L., Toye, C., & Flett, P. (2004). Living with Huntington's disease: Need for supportive care. *Nursing and Health Sciences*, 6, 123-130.
- Decruyenaere, M., Evers-Kiebooms, G., Cloostermans, T., Boogaerts, A., Demyttenaere, K., Dom, R., & Fryns, J. (2004). Predictive testing for Huntington's disease: Relationship with partners after testing. *Clinical Genetics*, 65, 24-31.

- Downing, C. (2005). Negotiating responsibility: Case studies of reproductive decisionmaking and prenatal genetic testing in families facing Huntington disease. *Journal of Genetic Counseling*, 14 (3) 219-234.
- Duisterhof, M., Trijsburg, R., Niermeijer, M., Roos, R., & Tibben, A. (2001). Psychological studies in Huntington's disease: Making up the balance. *Journal of Medical Genetics*, 38, 852-861.
- Etchegary, H. (2006). Genetic testing for Huntington disease How is the decision taken? *Genetic Testing*, 10 (1) 60-67.
- Forrest Kennan, K., Simpson, S., Wilson, B., van Teijlingen, E., McKee, L., Haites, N., & Matthews, E. (2005). 'It's their blood not mine:' Who's responsible for (not) telling relatives about genetic risk? *Health, Risk & Society, 7* (3) 209-226.
- Foster, C., Watson, M., Moyhnihan, C., Ardern-Jones, A., & Eeles, R. (2002). Genetic testing for breast and ovarian cancer predisposition: Cancer burden and responsibility. *Journal of Health Psychology*, 7 (4) 469-484.
- Galvin, R. (2002). Disturbing notions of chronic illness and individual responsibility: Towards a genealogy of morals. *Health*, 6 (2) 107-137.
- Hallowell, N. (1999). Doing the right thing: Genetic risk and responsibility. Sociology of Health & Illness, 21 (5) 597-621.
- Hallowell, N., Jacobs, I., Richards, M., Mackay, J., & Gore, M. (2001). Surveillance or surgery? A description of the factors that influence high risk premenopausal women's decisions about prophylactic oophorectomy. *Journal of Medical Genetics*, 38, 683-691.
- Hallowell, N., Foster, C., Eeles, R., Ardern-Jones, A., Murday, V., & Watson, M. (2003). Balancing autonomy and responsibility: The ethics of generating and disclosing genetic information. *Journal of Medical Ethics*, 29 (2) 74-83.
- Hallowell, N., Foster, C., Eeles, R., Ardern-Jones, A., & Watson, M. (2004). Accommodating risk: Responses to BRCA1/2 genetic testing of women who have had cancer. *Social Science & Medicine*, 59, 553-565.
- Hallowell, N., Ardern-Jones, A., Eeles, R., Foster, C., Lucassen, A., Moynihan, C., & Watson, M. (2005). Men's decision-making about predictive BRCA1/2 testing: The role of family. *Journal of Genetic Counseling*, 14 (3) 207-217.
- Harper, P. (1996). Huntington's disease. London: Sunders Publishing.
- Harper, P.S., Lim. C., & Craufurd, D. (2000). Ten years of presymptomatic testing for Huntington's disease: The experience of the UK Huntington's Disease Prediction Consortium. *Journal of Medical Genetics*, 37, 567-571.
- Henderson, B., Maguire, B., Gray, J., & Morrison, V. (2006). How people make decisions about predictive genetic testing: An analogue study. *Psychology & Health, 21* (4) 513-539.
- Huntington's Disease Collaborative Research Group (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*, *72*, 971-983.
- Huntington Society of Canada (2002). Genetic testing for Huntington disease. ON: Canada.
- Janz, N., & Becker, M. (1984). The health belief model: A decade later. *Health Education Quarterly*, 11, 1-47.
- Kenea, R. (1994). The Human Genome Project: Creator of the potentially sick, potentially vulnerable and potentially stigmatized? In I. Robinson (Ed.), *Life and death under high technology medicine* (pp. 49-64). London: Manchester University Press.

- Marteau, T., & Richards, M.P.M. (Eds.). (1996). *The troubled helix: Social and psychological implications of the new human genetics*. Cambridge: Cambridge University Press.
- Mays, N., & Pope, C. (2000). Assessing quality in qualitative research. *British Medical Journal*, 320, 50-52.
- Novas, C., & Rose, N. (2000). Genetic risk and the birth of the somatic individual. *Economy* & *Society*, *29* (4) 485-513.
- Nuffield Council on Bioethics. (1993). *Genetic screening: Ethical issues*. Available online at http://www.nuffieldbioethics.org
- Peters, E., Lipkus, I., & Diefenbach, M. (2006). The functions of affect in health communications and in the construction of health preferences. *Journal of Communication*, 56, S140-S162.
- Petersen, A., & Bunton, R. (2002). *The new genetics and the public's health*. London: Routledge.
- Potter, N., Spector, E., & Prior, T. (2004). Technical standards and guidelines for Huntington Disease testing. *Genetics in Medicine*, 6 (1) 61-65.
- Quaid, K., & Wesson, M. (1995). Exploration of the effects of predictive testing for Huntington disease on intimate relationships. *American Journal of Medical Genetics*, 57, 46-51.
- Ramsey, S., Wilson, S., Spencer, A., Geidzinska, A., & Newcomb, P. (2003). Attitudes towards genetic screening for predisposition to colon cancer among cancer patients, their relatives and members of the community. *Community Genetics*, 6, 29-36.
- Rebbeck, T., Levin, A., Eisen, A., Snyder, C., Watson, P., Cannon-Albright, L. et al. (1999). Breast cancer risk after bilateral prophylactic oophorectomy in BRCA1 mutation carriers. *Journal of National Cancer Institute*, 91, 1457-1479.
- Richards, M. (1998). Annotation: Genetic research, family life and clinical practice. Journal of Child Psychology & Psychiatry, 39 (3) 291-305.
- Richards, F. (2004). Couples' experiences of predictive testing and living with the risk or reality of Huntington disease: A qualitative study. *American Journal of Medical Genetics*, 126A, 170-182.
- Richards, F., & Williams, K. (2004). Impact on couple relationships of predictive testing for Huntington disease: A longitudinal study. *American Journal of Medical Genetics*, 126A, 161-169.
- Rogers, R. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protective motivation. In J. Cacioppo & R. Petty (Eds.), Social Psychophysiology: A source book (pp. 153-176). New York: Guilford Press.
- Silverman, D. (2000). Doing qualitative research: A practical handbook. London: Sage.
- Skirton, H., & Glendinning, N. (1997). Using research to develop care for patients with Huntington's disease. *British Journal of Nursing*, 6 (2) 83-90.
- Smith, J.A., Flowers, P., & Osborn, M. (1997). Interpretative phenomenological analysis and the psychology of health and illness. In L. Yardley (Ed.), *Material discourses of health* and illness (pp. 68-91). London: Routledge.
- Smith, J.A., Jarman, M., & Osborn, M. (1999). Doing interpretative phenomenological analysis. In M. Murray & K. Chamberlain (Eds.), *Qualitative health psychology: Theories and methods* (pp. 218-240). London: Sage.
- Sobel, S., & Cowan, D. (2000). Impact of genetic testing for Huntington disease on the family system. American Journal of Medical Genetics, 90, 49-59.

- Soldan, J., Street, E., Gray, J., Binedell, J., & Harper, P. (2000). Psychological model for presymptomatic test interviews: Lessons learned from Huntington disease. *Journal of Genetic Counseling*, 9 (1) 15-31.
- SuttonBrown, M., & Suchowersky, O. (2003). Clinical and research advances in Huntington's disease. *Canadian Journal Neurological Science*, 30 (Supp. 1), S45-S52.
- Taylor, S. (2004). Predictive genetic test decisions for Huntington's disease: Context, appraisal and new moral imperatives. *Social Science & Medicine*, *58* (1) 137-149.
- Tibben, A., Frets, P., van de Kamp, J., Niermeijer, M., Vegter-van der Vlis, M., Roos, R., van Ommen, G., Duivenvoorden, H., & Verhage, F. (1993). Presymptomatic DNA-testing for Huntington disease: Pretest attitudes and expectations of applicants and their partners in the Dutch program. *American Journal of Medical Genetics*, 48 (1) 10-16.
- Tibben, A., Timman, R., Bannink, E.C., & Duivenvoorden, H.J. (1997). Three year follow-up after presymptomatic testing for Huntington's disease in tested individuals and partners. *Health Psychology*, *16*, 20-35.
- Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior* (2<sup>nd</sup> ed.). Princeton, NJ: Princeton University Press.
- Warner, B., Curnow, L., Polglase, A., & Debinski, H. (2006). Factors influencing uptake of genetic testing for colorectal cancer risk in an Australian Jewish population. *Journal of Genetic Counseling*, 14 (5) 387-394.

## APPENDIX

Interview topic guide

- 1. History of personal experience with Huntington disease (HD)
  - How/when did you first learn of your family history of HD?
  - What was your childhood experience of HD?
- 2. Risk awareness/test decisions
  - When you hear the term 'at risk' for HD, what comes to mind?
  - Is or was being 'at risk' a part of your self-identity?
  - Does (or did) being at risk impact your daily life? If so, how?
  - Did you have the genetic test? Tell me about your test experience.
  - Did you NOT want to know your genetic risk? Is it ok for someone to NOT want to have the genetic test?
- 3. Stigma about HD
  - Do you feel any stigma within your family as someone who does (or might) have this disease?
  - What does society think of those who have (or might have) HD?
  - What do you perceive the response of others to be once they know you carry (or might have carried) the HD gene?
  - Do you feel responsible for your own health?

- Do you think society holds you accountable for your own health? For example, do you think society would expect those with a family history of HD to have a genetic test?
- Have you ever encountered any form of discrimination or prejudice because of HD? Or, if you received a negative test result, did you ever encounter any form of discrimination before you were declared gene-negative? [Prompt: Did anyone ever treat you badly or unfairly because of your family history of HD?]
- 4. Communication about the disease
  - Do you talk about HD within the family (immediate and extended)?
  - Do you talk about it with others outside the family (e.g., friends, co-workers?)
  - If so, under what circumstances? If not, why not?
- 5. Healthcare needs
  - What are your concerns, if any, regarding your own healthcare?
  - Do you have adequate insurance coverage?
  - Is there anything you need (e.g., health information or support) which is currently unavailable to you?
  - Do you have any suggestions for medical professionals who work with HD families?
  - Is there anything else you think we should talk about?

Chapter 15

# MAKING REPRODUCTIVE DECISIONS IN THE FACE OF A LATE-ONSET GENETIC DISORDER, HUNTINGTON'S DISEASE: AN EVALUATION OF NATURALISTIC DECISION-MAKING INITIATIVES<sup>\*</sup>

## Claudia Downing<sup> $\dagger$ </sup>

Research Fellow, Centre for Family Research, University of Cambridge, Free School Lane, Cambridge, England CB2 3RF

## ABSTRACT

This chapter makes two contributions to the psychology of decision making. It draws on empirical work to document how family members make and live with reproductive decisions as they become aware of their risk for a serious late-onset genetic disorder, Huntington's disease (HD). Decision-making involves negotiating two dimensions of reproductive risk – that for any child which might be born and the uncertainty that arises about the at-risk parent's ability to sustain a parenting role should he or she become symptomatic. A detailed account is given of how the model of responsibility was generated from their narratives. This model encapsulates what families find important when making reproductive decisions. It demonstrates that how people make decisions can become as important as what they decide. Examples are given to show how the model provides a framework for comparing how people deal with each dimension of risk, to compare different people's decision-making and to illuminate decision-making in the face of change. This includes how they address new options generated by recent developments in molecular genetics which can resolve uncertainty about, or avoid these

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Medicine and Health Care*, edited by Thomas E. Lynch published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: Centre for Family Research, University of Cambridge, Free School Lane, Cambridge, England CB2 3RF cd10008@cam.ac.uk

risks. A detailed account is also given of the process of their decision-making. These findings are used to evaluate claims being made by naturalistic decision-making (NDM) initiatives to account for decision-making in the real world. An outline of this initiative prefaces the research findings. Beach's (1990) image theory, used to illustrate the NDM approach, emphasises the role that values play in forming, negotiating, implementing and living with decisions that arise in everyday lives. Participants' accounts support this claim - revealing how values such as responsibility become established and contribute to decision-making. This and other findings lend support to the NDM claims. The chapter concludes with suggestions about how the model might be used more generally to further our understanding of the psychology of decision-making in the face of risk.

### INTRODUCTION

This chapter is about decision-making in the face of genetic risk. It presents a model of responsibility that encapsulates what families find important when making reproductive decisions in the growing awareness of their risk for a late-onset genetic disorder, Huntington's disease (HD). It will be argued that this model facilitates our understanding of how those living with genetic risk plan their lives in the age of the new genetics. These comprise advances in molecular genetics making it possible to offer genetic testing – both prenatal and predictive – to those identified as at-risk from their family history for an increasing number of disorders. Testing establishes whether mutations, which are likely to result in these disorders, are present in genes. It differs from other forms of medical testing in that any body tissue can be used and it can be performed at any stage of life from conception. Results may be sought to inform personal decision making rather than medical care.

Concerns about one value, responsibility, have been consistently cited as motivating the new decisions that people faced about predictive testing (Cox, 2002, 2003; Kenen, 1994, 1996; Kay, E. & Kingston H.,2002; Hallowell, 1999; Robertson, 2000; Smith, Stephenson and Quarrel, 1999; Taylor, 2004). Decision-making is instigated and constrained by social and biological connections to others such as biological family members who share similar risks, future generations, and partners (Hallowell, 1999; Burgess and d'Agincourt-Canning, 2001; Downing, 2002, 2005). Responsibility is not enacted from a single script: those at-risk for HD report aiding their adult children's reproductive decision-making as an important consideration both when choosing to have or not have predictive testing (Smith, Stephenson and Quarrel, 1999). This finding suggests how new decisions become part of ones people have always faced in their everyday lives - such as those around reproduction. It also shows that these decisions can continue to be made without recourse to clinical settings and the additional information that genetic testing provides.

Reproductive decision-making in the face of genetic risk has been an under-researched area. A review of 547 studies published between 1986 and 1996 categorised as being about informed decision-making yielded only seven that dealt with genetic factors, and none which addressed late onset disorders or examined how reproductive decision-making is experienced as new options to clarify risk status have become available (Bekker et al, 1999:32). The indepth qualitative study of reproductive decision-making in the face of late-onset genetic risk whose findings are reported in this chapter was undertaken to address this lacunae.

The chapter also addresses suggestions that adopting a decision theoretic approach could enhance our understanding of how people make reproductive decisions under these conditions (Huys, Evers-Kiebooms and d'Yedwalle, 1992; Shiloh, 1996; Broadstock and Michie, 2000). The introduction notes problems inherent in traditional approaches used to study decisionmaking in the face of risk and suggests why the naturalistic decision-making (NDM) initiative offers to become a more informative resource. The study findings will be used to test their claim to account for personal decision-making. In preparation for this evaluation the introduction presents key tenets of NDM, reasons for focusing on one NDM model, Beach's (1990) image theory, what this comprises, what was previously known about perception of genetic risk, reproductive decision-making in the face of genetic risk and what is different about late-onset risks.

#### Naturalistic Decision Making

The NDM initiative sets out to address limitations inherent in the approaches, such as health psychology and normative decision theory, more traditionally deployed to study decision-making in the face of risk. These approaches were initially constrained by an underlying assumption that people lived their lives in a planned and rational fashion, and that, once informed about risks would engage in preventative behaviour. This assumption continues to drive research agendas, as can be seen from the following comment concluding a systematic review of 547 studies published between 1986 and 1996 categorised as being about informed decision-making (Bekker et al, 1999: 22). The authors' key recommendation was that:

Genetics and prenatal diagnosis ... is a good place to begin [further studies] because the nature of the information base, with relatively simple clearly-defined options with known risks corresponds closely to the terms that decision theories use

Another limitation has been these approaches' tendency to focus on the individual. Individual factors *do* shape decision-making: holding internal or external attributions of causality (Wallston, 1992) being risk averse or seeking (Lopes, 1987) and decision averse or confronting (Beattie et al, 1994) contributes much to how people assess problems confronting them, constrains their ability to deal with them, and to negotiate with others. But, as Kahneman (1991: 145), in a personal reflection on the field of the psychological study of judgement and decision-making points out, it is essential to acknowledge that 'significant decisions are made in a social and emotional context'. The value of a multiplicity of studies over the years has been to document social constraints in which risk is actually experienced and acted on. Health psychology has identified other factors, such as perceived social support and efficacy, which can mitigate these constraints. Efforts have been made to incorporate these factors in subsequent health psychology models such as the protective motivation theory (Rogers, 1994) and self regulation theory (Leventhal, Nerenz and Steele, 1994).

Considerable overlaps occur between evolving health psychology models and decision theories. Theoretical overlaps occur in concepts, such as risk perception, perceived benefits of actions, beliefs about self-efficacy and response-efficacy and perceived control, used to explain behaviour in the face of risk. Correspondences exist in findings about how people perceive and react to the presentation of risk information. For example, laboratory based normative decision theory studies have yielded a taxonomy of heuristics and biases that shape the way people process risk information when making decisions (Kahneman and Tversky, 1984) which have been replicated in health settings and used to inform clinical practice. Both approaches have come to depict decision-making as a process, with a preliminary period of cognitive restructuring preceding choice. The health belief model now postulates that taking preventative measures is a function of considering perceived susceptibility to the risk, benefits of taking actions such as genetic testing and barriers to taking this action whilst the protection motivation theory emphasises the role played by beliefs about magnitude and likelihood of threat, and the efficacy of envisaged responses.

Three characteristics distinguish NDM initiatives from previous decision-theoretic approaches: their focus on decision-making in real life contexts; their recognition of the competence that people achieve when making decisions under difficult conditions; and, being open to a range of disciplines and perspectives thought to contribute to a richer understanding of decision-making. These characteristics are considered in more detail below.

#### Decision-Making in Real Life Contexts

NDM initiatives have concentrated on two aspects of real life: decision-making in field settings which may result in disaster situations,<sup>2</sup> and personal decisions that people see as important in their private lives including divorce (Willen and Montgomery, 2000), and childbearing (Willen, 1994). Personal decisions, which can entail negotiating risks, are characterised by:

... deep and long-term significance to the subject. By making the decision in question the person is attempting to make his or her life turn out differently (Willen, 1994: 24:1).

Decision-making frequently involves, or has to consider, significant others, precisely because people are embedded in social networks. Far less empirical work has been conducted into personal decision-making. Montgomery (1999) remarked on the absence of studies which in addition to risk explore 'hot' psychological factors such as "social influences ... emotions, interests and values".

NDM's initial remit was to identify what decision-making encompasses (Cannon-Bowers, Salas and Pruitt, 1996). The value of resulting working models has been in delineating and broadening the parameters within which decision-making actually occurs rather than in specifying *exactly* how each decision-maker will operate (Lipshitz, 1993).<sup>3</sup> Considerable agreement exists between models about the nature of these parameters. NDM models can be categorised into two groups. The first group, *typology* models, identify contrasting cognitive processing styles, such as intuitive and analytic, which may be employed at various stages of decision-making. The second group, *process* models, suggests what these stages might be. How these approaches coalesce is outlined below.

<sup>&</sup>lt;sup>2</sup> For detailed accounts of NDM approaches see Montgomery and Svenson (1989); Klein (1999); Klein, Orasanu, Calderwood and Zsambok, (1993); Svenson and Maule (1993); Svenson, (1996), Klein (1997); Zsambok and Klein (1997).

<sup>&</sup>lt;sup>3</sup> For further details of these models see Klein et al (1993) and in particular Klein's diagrammatic representation of their communalities in figure 22.1, page 391.

Process models distinguish three broad stages. These are: (i) the preliminary work needed to come to a decision, (ii) implementation of choices, and (iii) living with decisions. Preliminary work involves perception of the situation, which may include evaluation of risks, formulation of goals, and assessment of options open and strategies available to realise these goals. Models differ in how they conceptualise this work: as screening out unacceptable options according to values held (Beach, 1990); as 'the search for a dominant option' (Montgomery, 1993); or, as 'differentiation and consolidation' of chosen options (Svenson, 1991). It is probable that people act in all of these ways, but by limiting their focus, models remain manageable. Implementation is characterized by mental simulation of chosen strategies, realising them, and modifying them in the light of progress decisions about their ability to achieve goals. Progression through these stages may be both linear and recursive. For example, living with decisions involves post decision evaluation, which can then contribute to preliminary work for subsequent decisions. It may also include drawing on decisions made by others.

If decisions are perceived as clear, acceptable, feasible and familiar, passage through these stages becomes unproblematic. Little overt consideration is required, and the processing style employed is described as intuitive. It will be suggested that this frequently characterises reproductive decision-making in general populations.

A different picture of decision-making emerges when decision-makers perceive themselves as confronting unfamiliar decisions or uncertain, ambiguous, redundant or incomplete information - such as genetic risk. Under these circumstances decision-making may necessitate a more analytical processing style. Action feedback loops (which may include previous decisions) and mental simulations may be employed in order to structure and implement decisions and reflect on them. Feedback loops and mental simulations blur distinctions between the stages. As Orasanu and Connolly (1993:19) explain:

Instead of analyzing all facets of a situation, making a decision, and then acting, it appears that in complex realistic situations people think a little, act a little, and then evaluate the outcomes and think and act some more. ... Development of this knowledge [is] ... an integral part of decision making.

Competing and shifting goals may emerge which have to be resolved within a temporal emotional framework of anticipation, hindsight and regret (Kahneman, 1991). Conflict, which may lead to procrastination, is seen as 'the price one pays for the freedom to choose' (Tversky and Shafir, 1992: 358).

#### Recognition of Competence

NDM models emphasize the competence people achieve whilst operating within the constraints inherent in naturalistic settings. As Klein (1999: 1-2) explains:

Instead of trying to show how people do not measure up to ideal strategies for performing tasks, we have been motivated by curiosity about how people do well under difficult conditions.

Competence is conceptualized as feedback loops showing the ability to benefit from past experience. This can take several forms: shaping values (Beach, 1990); recognition of ability to adapt to changing situations and priorities (Payne, Bettman and Johnson, 1993), which may involve the creation of new pathways and perspectives (Willen and Montgomery, 2000); and, developing the expertise to utilise appropriate and disparate sources of power, such as intuition, mental simulation, metaphor and storytelling. Klein (1999: 3) elaborates on how these strategies coalesce:

The power of intuition enables us to size up a situation quickly. The power of mental simulation lets us imagine how a course of action might be carried out. The power of metaphor lets us draw on our experience by suggesting parallels between the current situation and something else we have come across. The power of story telling helps us consolidate our experiences to make them available in the future, either to ourselves or to others.

Competence is also associated with exercising control, in that decisions represent a desire to:

... act upon the world, to make sure that the future does not look like the past. Decision makers go to great lengths to insure that they have the ability to control future events, (Beach and Lipshitz, 1993: 26).

Where decisions involve risks, which can be defined as uncertainty about negative or adverse outcomes (Teigen and Brun, 1997), it may not be possible to realise this objective. Under these circumstances competence may become making decisions that it is possible to live with.

#### Transtheoretical Aspects

The third distinguishing aspect of NDM is that it is open to a variety of theoretical perspectives thought to provide additional insights into decision-making. These include phenomenology (Karlsson, 1987), identity formation (Sloan, 1986; Nelson and Nelson, 1995), social approaches to risk (Wartofsky, 1982; Krimsky and Golding, 1992; Palmlund, 1992) and narrative accounts (Beach, 1997; Cox, 2002, 2003; Klein; 1999). What these perspectives comprise and can contribute is outlined below.

Phenomenology captures how people experience exercising their will, develop a sense of competence, reflect on what they should do and structure decisions as realisable projects with others (Karlsson, 1987). Decision-making shapes identity in that it contributes to the way we are perceived, by others and ourselves. As Nelson and Nelson (1995: 136) explain:

Choices made ... in a very important way determine who we are, who we have been, and who we will be as moral people; this gives the choices their proper weight, and defends us from the kind of self-deception that is sometimes a part of having to decide in morally difficult circumstances.

Social approaches to risk elaborate how identity is shaped. Conceptualising decisionmaking as social drama, consisting of competing stories involving justifications of actions and attributions of blame, identifies a range of roles that are played out when making decisions in relation to risk (Palmlund, 1992). These include risk bearers, risk generators, risk informers and risk arbiters. Questions that can be explored in empirical work include how these roles come to be allocated and assumed in families, the expectations they arouse, and how these expectations can constrain or facilitate decision-making (Downing, 2005).

Answering these questions involves formulating 'looser' definitions which draw on narrative approaches. For example, Beach (personal communication, August 1997) describes decision-making as 'allowing people to tell stories to construct alternative futures and find ways of making the stories they like happen'. This definition connects us to studies such as those of Cox (2002, 2003) which enhance our understanding of the different ways in which people 'progress' from intentions to implementation of predictive genetic testing through examination of this as a 'narrative moment'. Her participants' stories of 'evolving towards it', 'having to know' and 'taking the decision' bear little similarity to conventional notions of decision-making as a rational planned process of decision and implementation. In this way, stories become 'natural experiments, linking a network of causes to effects' (Klein, 1999: 196). Exploration of narratives reveals and challenges prevailing cultural stories, as shown in Sandelowski and Jones' (1996) qualitative study of women's constructions of choice after the detection of fetal abnormalities.

#### Image Theory as an Example of NDM Approaches

Underlying similarities in how NDM models conceptualize decision-making enable their principles to be illustrated in more detail by one particular model, image theory (Beach, 1990). An additional reason for focusing on this model was that it evolved from a normative study of reproductive decision-making in a general population and thus can demonstrate how NDM approaches offer a far richer understanding of decision-makers' perspectives and processes.

The original research involved construction and employment of a decision analysis tool, the Optional Parenthood Questionnaire, to help couples think more rationally about the costs and benefits of having a/another child (Beach, Townes, Campbell and Keating, 1976). Partners were asked to rate a list of pre-determined values - gleaned from birth planning literature - which children could fulfil for parents. Values were narrowly defined as utilitarian benefits and costs. They indicated considerations such as caring for children with health problems which people might take into account when making reproductive decisions. Questions were weighted equally making it impossible to assess, even under hypothetical circumstances, the relative importance people placed on each aspect. These constraints, inherent in a traditional decision-analysis approach, were subsequently removed when values were re-conceptualised more loosely as 'images' in image theory.

#### Processes Identified in Image Theory

Decision-making is depicted as comprising two different but connected processes: screening and choice. Findings that people do not necessarily return to previously rejected options when those in the choice set become unavailable demonstrate that screening is more than a means of reducing cognitive overload, as had previously been suggested by Simon's (1955) model of bounded rationality. According to Simon people reduce problems to a limited number of variables in order to decrease cognitive overload, enabling them to choose

options that seem reasonable, but which are not necessarily optimal. In contrast, Beach (1990) showed that screening limits choice to options that meet decision makers' minimal standards - or, produces an empty choice set, thereby prompting a search for better options. Choice only becomes possible when options exist that are not discordant with the decision makers' morals, beliefs, attitudes and values - henceforth referred to by Beach as *images*.

## Role Played by Images

Images are defined as schema with cognitive and affective components that organize decision-makers' values and knowledge, guiding their decisions and creating the frames within which decision-making occurs. Three types of images form, and have to be reconciled, in order for decision-making to proceed. These are:

- value or principles, relating to how people feel they should behave;
- *trajectory or goals*, relating to what they want their future to be like (which may be different);
- strategic or plans, encompassing ways of securing that image.

Images continue to be drawn on to monitor progress, demonstrating that they are influential at *all* stages of decision-making.

Empirical support exists for Beach's (1990) claim that these considerations shape decision-making. A meta-analysis of involvement studies in relation to attitude formation showed that the form of involvement differed according to whether values, impressions or outcomes were prioritised (Johnson and Eagly, 1989). High value-relevant involvement was associated with selection and bias, whilst high outcome-involvement was associated with a more flexible approach, which allowed decision-makers to respond to social interaction.

Verplanken and Svenson (1997) drew on Svenson's (1991) NDM differentiation and consolidation model to demonstrate that the form of involvement affected how decisions were made. Pre-choice cognitive processing is depicted as differentiating between options in order for one to emerge as preferable and become chosen, and post decision processing as consolidation of this option by accentuating its positive aspects whilst stressing negative aspects of rejected ones. Consolidation corresponds closely to the process of cognitive dissonance (Festinger, 1957) but differs from it in that it commences before decisions are implemented.

Further insights into the relationship between emotional engagement and decisionmaking are provided by Willen and Montgomery (2000). They show that people in troubled marital relationships can only move towards a decision situation after they have changed their form of involvement, termed 'perspective'. Reconfiguring partners' previously positively regarded qualities as negative produces psychologically more 'distant' (or less emotionally involved) perspectives on these relationships. This shift enables them to contemplate divorce. As they distance themselves from their relationship, decision-makers become increasingly emotionally involved with future perspectives. These incorporate possible outcomes of divorce, such as how it will affect them, their life and those they care about and how these concerns can be reconciled with values important to them.

#### Contribution of Narrative Approaches to Image Theory

Beach does not elaborate how images arise, change or become shared. Incorporating narrative approaches to family therapy (Dallos, 1997) and family aspects of health (Nelson and Nelson, 1995) makes it possible to see how this happens. Dallos amalgamates attribution theory (Heider, 1958), personal construct theory (Kelly, 1955) as applied to families (Proctor, 1981). Dallos's (1997: 96) starting point is the personal constructs (reflecting individual values) that people develop to make sense of the world. Personal constructs affect how disclosures and sharing of life stories are presented. They signal what is important (but not necessarily agreed) for family members about their decision-making:

Remembering in families can be seen as a joint, collaborative activity in which partners may remind each other of the facts - what happened, when it occurred, who was there, what was said - but also the meaning of these events, in particular what people's intentions were, how they felt at the time. There may also be a reflexive view about what has been learned from the event, how perceptions of the event differ now, and how in the same circumstances they might now act differently.

The resulting narratives identify stakeholders in decisions; encompass alternative representations of events, predictions or anticipations of other's actions. They are shaped by internal and external attributions (Heider, 1946) including justifications and blame. They contain themes such as responsibility; maintain family myths and link family members together over time. These become shared social realities, indicating 'family' constructs or values, which are more than a collection of individual constructs.

Nelson and Nelson (1995: 130) suggest that narratives from families of birth provide working models to be drawn on and adjusted in subsequent decision-making in families by marriage:

In early adulthood we begin to plot out our life story, using the first chapter our parents wrote with us as a point of departure. That first chapter can now take on meaning as a guide to action, if we integrate the present moment of decision into the overall narrative of our lives, either by ratifying our existing course or by deliberately charting a new one. It is open to us to make our present choices in such a way as to weave our past into our future, consciously and thoughtfully constructing a meaningful moral autobiography.

In this way, narratives come to provide an ethical basis for decision-making:

The narrative ... asserts that the quality of the decision matters as well as who is making it. (Nelson and Nelson, 1995: 107).

Support is therefore shown for claims that additional insights into the decision-making process will be provided by links made to other domains, such as narrative approaches (Beach, 1997). Testing these claims involves deploying qualitative methods that provide access to decision-makers' narratives.

#### Accounting for Shared Decision-Making

Work conducted in corporate contexts led Beach (1993: 91) to state that individual decision-making precedes shared decision-making:

Ultimately every decision is made in the isolation of a single human mind. Even when working in a group, each individual has to make his or her own decision. Through negotiation, the group's decision is derived from its members' private decisions.

He postulates that a process of shared screening, termed negotiation, subsequently occurs. Negotiation only becomes possible when an option exists that is acceptable to at least one individual. Different patterns of negotiation observed by Beach (1996) were:

- *yielding*, defined as passive acceptance of decisions formulated by others;
- *compromise*, defined as the abandonment of an individual's own decision but not requiring wholesale acceptance of someone else's;
- *contending*, defined as each person stating what they think the decision should be, then attempting to persuade others of this until other options are dropped or a majority consensus is formed;
- *problem solving*, defined as looking for a further option that might suit all participants.

His taxonomy provides a starting point to compare joint decision-making in other contexts, such as the family, where individuals may be connected biologically as well as having shared a past and be anticipating a shared future, and in relation to other types of decisions, such as the reproductive ones which form the focus of the next sections and this chapter

## Reproductive Decision-Making in General Populations

General populations comprise those who do not have to contend with reproductive problems such as infertility or the known risk of a genetic disorder. Understanding their experiences helps clarify how reproductive decision-making differs when people face considerations about genetic risk. This section summarises what demographic and exploratory studies reveal about their decision-making.

*Demographic studies* report outcomes of reproductive decision-making, such as family size. They provide valuable contextual information by documenting how this changes over time for different cohorts, but cannot access the decision-making that culminates in these outcomes. As Wexler (1980: 322) - herself at-risk for HD - states, decision-making, being essentially a private matter, requires different approaches. :

We know so pitifully little about the myriad of complex motivations that propel people to propagate their kind, particularly now that birth control methods are giving people such freedom of choice. Counting facts retained and babies born is not enough. What goes into that crucial decision?

Answering this question involves qualitative approaches that seek out and prioritize perspectives of those making decisions. Exploratory interview studies reveal decision-making to be a process, characterised by a multiplicity of decision points rather than a single 'choice' (Busfield and Paddon, 1977; Payne, 1978; Richards, 1978; Dowrick and Grundberg, 1980; Fox, 1982; Currie, 1988; Willen, 1994). As well as deciding whether to have a child, these include with whom, whether to have further children, when to do this, under what conditions, and what action should be taken to achieve these goals. These considerations face everyone including those contending with difficulties, such as infertility or genetic risk even if as will be shown, they come to conceptualize them differently.

These studies also show how people construct and resolve these decision points. Establishing shared social realities forms a significant preliminary step. For example, many of Willen's (1994) respondents framed reproductive decision-making primarily as a relationship question, focusing not so much on *whether* to have children as *with whom*. Reproduction becomes part of the 'natural' pattern of acceptable relationships and, as such, not requiring overt reflection or negotiation, except about what might be the 'right time' to implement previously unexpressed and unquestioned intentions. This is revealed to be a subjective judgment of circumstances rather than a chronological point (Currie, 1988). These observations are consistent with other findings, such as those from an earlier study (Mansfield and Collard, 1988) which found that newly-wed couples did not discuss reproduction, and were evocative of the commonly used expression, 'falling for a baby' which suggest little in the way of conscious planning.

Adopting a NDM perspective enables us to construct a coherent account of these findings. This confirms that decision-making *does* occur but in a more intuitive form, consistent with the values reproduction represents for those concerned. As Beck and Beck-Gernsheim (1995:196) postulate:

A child represents ... the 'natural' side of life ... Motherhood seems to offer the woman an alternative refuge from the working world, where it is imperative to behave responsibly, and soberly, and emotions are generally considered a nuisance. Committing yourself to a child means contradicting the cognitive side of life, and finding a living counterweight to all that souldestroying routine.

Taking a 'natural' stance can thus be interpreted as conforming to existing norms that children will 'happen' within relationships and do not need to be consciously 'decided' or planned. The existence of these norms is corroborated by the experiences of the involuntarily childless who cannot fulfil them (Pfeffer and Woollett, 1983; Mason, 1993; Read, 1995), and the voluntary childless who must resist them (Campbell, 1985; Kiernan, 1989; Bartlett, 1994; Morell, 1994; McAllister and Clarke, 1996).

#### Reproductive Decision-Making in the Face of Genetic Risk

Questions arise about whether the seemingly unquestioning acceptance of norms to have children and associated intuitive processing can be sustained (and how) once decision-makers become aware of genetic risk. In contrast with the 'normal' experience outlined above, reproduction in the face of genetic risk is depicted as involving the recognition of uncertainty about potentially severe and possibly irreversible consequences, and that decision-making has significant implications for the decision-makers, other family members and future generations. Factors thought to influence reproductive decision-making under these circumstances are: the value placed on having children; the perception of genetic risk; and personal experience of the condition (Shiloh, 1996). How these factors are thought to operate is considered below.

#### The Value Placed on Having Children

The experiences of the childless illuminate how the value placed on having children is construed by those for whom reproduction is problematic. Examination of the contrasting standpoints of the infertile and the voluntarily childless reveal how having children becomes imbued in contrasting ways with notions of *loss* and *threats to identity*. Their thinking, outlined below, may also apply to those facing genetic risks.

#### Involuntary Childlessness

Being confronted with infertility creates feelings of abnormality, leading to loss of selfesteem and self-identity, a sense of not being in control of one's life, and an existential anxiety about the purpose of life (Read, 1995). These losses may be experienced differently by men (Mason, 1993) and women (Pfeffer and Woollett, 1983), and hence between partners (Mason, 1993). Problems relating to one partner can lead couples to feel disconnected, or 'out of step' with each other. The supposition that those facing genetic risks experience similar losses and disconnections is borne out in a qualitative study of women carriers of X-linked genetic disorders (Kay and Kingston, 2002). Loss of normal expectations became one of the women's principal concerns.

Hesitations were also noted in various studies about employing strategies, such as artificial insemination by donor (AID), to redress these problems. This phenomenon indicates another dimension: the considerable, and sometimes overriding, value people place on being genetically connected to their children. Not wishing to lose this connection can prevail over the envisaged loss of childlessness.

#### Voluntary Childlessness

Accounts from the voluntarily childless stress what they see themselves as having *gained* by not having children: for example, self-fulfilment, and alternative ways to establish identity (Morell, 1994) and bond with partners - such as working on shared projects (Willen, 1994). In contrast with the infertile, they construe reproduction as leading to the loss of their identity. McAllister and Clarke (1996) found that only a third of women self-categorised as voluntarily childless had initially made a firm decision never to have children and ten percent were still ambivalent. Many of their respondents had regular, and valued, contact with other people's children. These observations suggested that factors other than values held about children were crucial in implementing women's decisions. It emerged that a key aspect was holding negative images of parenthood: they viewed parenthood as sacrificing hard won social lives, financial independence and jobs. Not wanting to experience these losses made them reluctant to become mothers.

#### Perception and Experience of Genetic Risk

The psychometric work conducted by Slovic, Lichstenstein, and Fischhoff, (1984) and Slovic (1987) provides a useful starting point for learning how people perceive risk. Analysis of people's ratings of a variety of risks on a number of given dimensions generated three underlying factors. These are:

- (i) *dread* risk, which relates to perceptions of the uncontrollability of exposure to risk, dread associated with its impact and whether distribution is judged as inequitable;
- (ii) unknown risk which reflects the extent of personal experience and expert knowledge that exists of a risk, how observable it is, and whether effects are considered to be immediate or delayed. The impact of experience of risk on decision-making is likely to be complex. It has been suggested that risks which are known are more likely to be perceived as less threatening (Slovic, 1987; 1992);
- (iii) the number of people likely to be exposed to or affected by a risk.

Their findings (despite not drawing specifically on genetic risk) *are* consistent with prevailing discourses about genetics. For example, Rothman (1998: 233) explains how genetic determinism views an individual's genetic makeup as an uncontrollable accident of fate which has a great impact in that it holds the key to identity, and because it is readily observable to those with expert knowledge:

A story is written into our bones and skin and hair, there to be read by anyone who knows the code.

Reproductive decision-making involves transcending this reductionist stance, in that:

gametes are not private possessions. Biologically they are shared values, realised in sexual union. They have to be combined and merged with the genetic values of others, (Fletcher, 1980: 134).

Awareness of potential and actual genetic connections changes an independent concept of self into an 'interdependent' social one, heightening one's awareness of genetic responsibilities towards others who may also be affected by how one acts on this information (Kenen, 1994, 1996). They include biological family members who share similar risks, future generations, and partners. Their stories portray risk as experienced subjectively and emotionally (Gifford, 1986; Hallowell, 2006), as "threats to people and things they value" (Kates and Kasperson, 1983: 7029).

Thus what risk means and what people are prepared to *do* about it is likely to vary according to their underlying value systems. If, as has been claimed earlier in this chapter, values are formed in social contexts, decision-making about genetic risk will also be shaped by the wider social and cultural framework in which decision-makers live (Douglas, 1985). As Rothman (1998: 39-40) remarks, risks become enmeshed with:

our own personal troubles, our daily lives, our intimate concerns, and the world in which we live, the issues that face us collectively. We live a biography, a personal tale, but we live it in a moment of history, in a collective time.

Rothman's comment reminds us of the need to acknowledge that reproductive decisionmaking differs from other decisions about genetic risk in that it is embedded in a particular history. In addition to their biological inheritance, families have had to contend with a 'social' legacy of opposition to their existence and reproduction. Social legacies, as well as current attitudes, vary between cultures. Within living memory, members of HD families were murdered in Nazi Germany. In other western countries sterilization was advocated, and sometimes enforced in some American states (Davenport and Muncey, 1916; Wexler, 1980; Harper, 1996a; Thom and Jennings, 1996). Reducing the incidence of at-risk births remained as an objective of genetic counselling in Britain for many years (Carter and Evans, 1979; Harper et al, 1979; Thom and Jennings, 1996). These eugenic measures were justified as being the responsible way to promote the wellbeing of future generations. Perception of this legacy is thought to have caused many of those at-risk who wished to have children to remain "doctor shy" (Wexler, 1979: 200) long after more liberal attitudes came to prevail.

In many countries it has subsequently become unacceptable to suggest that those at-risk should refrain from reproduction. It is now generally assumed that people want to feel in charge of their own destinies, are able to exercise freedom of choice in ways that benefit their wellbeing and wish to take an active role in managing their own affairs, including negotiating genetic risk (Conrad and Gabe, 1999). What has *not* changed is an underlying assumption on the part of health care professionals and the general public that, given a choice, people would prefer not to pass on known risks.

This assumption is borne out by prenatal studies in general populations which identify a moral and a 'technological imperative' to make use of risk avoiding possibilities (Tymstra, 1989). It explains why some women terminating an abnormal pregnancy identified through routinely offered prenatal testing report having had "no choice" - when clearly they did - because the alternatives were rejected as unacceptable, or too difficult to justify (Green, Statham and Snowdon, 1992). Pregnancy is changed even for those who decide not to make use of fetal diagnosis. As Green (1990: 38) explains:

mothers who decline to take part ... will always know that they could have had that knowledge and could have acted on it .

Nelson and Nelson (1995: 69-76) argue that expectations arise that *both* parents will act in their child's interests:

"Parents' duties to children, ... arise from their direct responsibility for bringing the children into existence, whether they meant to or not - and for putting them into the world in a condition of extreme vulnerability ... Nor can one parent release the other from this responsibility, as it is the child, and not the other parent, to whom the debt is owed.

Parents are likely to be viewed as irresponsible if they act in a way that is judged by others not to be in their fetus' best interests (Lippman, 1992).

Potential quality of life, becomes a paramount consideration when making decisions about testing and termination for fetuses at-risk for, or affected by, disorders that manifest at birth or in infancy (Kitcher, 1996). The more a disorder is perceived to negatively affect quality of life the more willing people are to consider termination (Evers-Kiebooms et al, 1993; Drake, Reid and Marteau, 1996).

Little is known about how people view these considerations in relation to a life that might be 'normal' until adulthood, but then be affected by a genetic disorder. HD offers a particularly rich context to explore them, in that the clinical picture presents families with two decision-making dilemmas. The first is whether to have children who, in later life, may develop the disease. The second is whether it is justifiable to have children - given that uncertainty arises about the at-risk parent's ability to sustain a parenting role. HD's late-onset makes it likely that people will have had had to make decisions about childbearing, and know of their risk, before developing any overt symptoms. Affecting both men and women these comprises cognitive changes, loss of control over voluntary movements, psychiatric symptoms and marked physical decline over a period of ten to fifteen years leading to inevitable death both men and cognitive changes, loss of control over voluntary movements, psychiatric symptoms and marked physical decline over a period of ten to fifteen years leading to inevitable death (Harper, 1996b). The mutation is highly penetrant, meaning that, providing they live long enough, most people who inherit it will develop HD. If they were to develop the disorder, the current lack of any effective treatment or cure would then threaten their ability to mother or father their children.

Few have utilized genetic testing since its introduction in the late 1980s to resolve these dilemmas (Harper, Lim and Craufurd, 2000; Hayden, 2000; Simpson and Harding, 1993). Low uptake of predictive testing has been attributed to reluctance to make contact with clinical genetics services, reluctance to lose hope, uncertainty about ability to cope given the continuing lack of treatments, and threats to identity. Van der Steenstraten et al (1994) found that non-requesters of predictive testing were more likely to have learnt of their risk before adulthood than requesters. This led them to hypothesise that non-requesters were more likely to have incorporated HD into their identity and have less need to resolve their uncertainty than requesters who perceived their risk as threatening their previously established identity.

Making decisions about prenatal testing concerns both general populations who undergo routine antenatal screening as well as those aware of genetic risk. Decisions about whether or not to employ prenatal and subsequent decisions about terminations may need to be made under severe time constraints. Factors found to be associated with use of prenatal testing for early onset disorders include perceived accuracy of the test, willingness to consider termination, and other family members' attitudes towards termination (Shiloh, 1996).

Values held about abortion can be either absolute or relative, in that people might accept abortion for disorders that affect children's quality of life but not see it as appropriate for HD. If, as has been suggested above, those growing up in affected families see HD as part of their identity, they can be reluctant to consider termination, seeing it as rejection of themselves (Hayes, 1992; Wexler, 1992; Richards, 1993; van de Steenstraten et al, 1994). Those unable to agree (for whatever reason) to termination of high risk fetuses are unlikely even to be offered prenatal testing. However, documented cases of women discharging themselves from hospital at the penultimate moment rather than undergoing a termination (Tolmie et al, 1995) and of continued high risk pregnancies (Bloch and Hayden, 1990) reveal that people cannot always predict how they will react as the decision-making process unfolds.

Learning more about clinical encounters can shed some light on what they can contribute to the decision-making process. For those uncertain about their reproductive decisions - one third of the participants in one study (Wertz, Sorenson and Heeren, 1984) - counselling can be helpful at an early stage of decision-making, as part of (Kessler 1989: 350):

an ongoing process, over time, of evaluation, weighing of options, and of responses to personal and interpersonal factors all of which contribute to reproductive decision making

Jungermann's (1997) advice giving and taking (AGT) model offers an explanation of what happens when counselling occurs at this pre-choice stage of decision-making. He focuses on interactions, such as those which arise between consultants prepared to give advice and clients prepared to consider advice. Advice is defined as the preferences of another person. Four components thought to comprise this process are: description of the problem followed by identification of appropriate options; justification of particular options; particular options being offered as advice by consultants; and clients' evaluation and consideration of this advice culminating in either acceptance or rejection. Willingness to accept advice is linked with factors such as perceived credibility and confidence felt about the person giving advice. Jungermann (1997) suggests that uptake of advice is limited because decision-making is frequently guided by values, about which clients are the experts, rather than facts, such as the probability of a negative genetic outcome, which are the provenance of the consultant.

Other research reveals Jungermann's dichotomy as too simplistic, in that it is not possible for facts to be provided in a neutral way. Comparisons made between professional groups have shown them to evaluate risks differently (Marteau, Drake and Bobrow, 1994) which can then influence how they convey information. It is also necessary to acknowledge the taxonomy of cognitive heuristics and biases that shape the way people process information presented to them (Kahneman and Tversky, 1979; 1984). These findings have implications for clinical practice. As Wexler (1992: 219) comments:

It is so important to explain genetic information both in terms of gain and loss. Telling clients they have one chance in four of having an affected child conveys one psychological message, saying that they have three chances in four to have a normal baby conveys a different one - even though the statistics are the same.

It is likely that consultants and clients prioritize different forms of involvement (Verplanken and Svenson, 1997) which affect how they present and perceive risk and responses to it.

It has also been shown that when and where people are offered these tests influence their uptake. Women are less likely to question them if they are presented in antenatal clinics as 'routine' prenatal care (Richards and Green, 1993; Green, Statham and Snowdon, 1994). Testing for rare late-onset disorders is less likely to be presented in this way, as it is felt that more detailed information will be provided by genetic counselling clinics. This factor may partially explain why uptake of prenatal testing is very low for HD (Hayden, 2000) and practically non-existent for other late-onset disorders such as inherited breast, ovarian and bowel cancers (McAllister, 1999). Very little interest has been expressed in prenatal testing for HD (Mastromauro, Myers and Berkman, 1987) and even fewer have actually used it after it became available (Hayden, 2000).

Genetic counselling has a different impact when it coincides with a later stage of decision-making. A review of prospective studies of genetic counselling concluded that counselling then *confirms or reinforces* decisions already taken (Kessler, 1989). Counselling can also *reduce anxiety* about implementing decisions already taken. This supposition would

account for an observed post counselling increase in births for those who perceived themselves to be at high risk.

Clinical studies, some of which are cited above, report only the views of a small and unrepresentative proportion of the total estimated HD at-risk population willing to make their family history known. Clinical studies' focus on the 'patient' also means they fail to address the social aspects of decision-making. Aggregation of existing data makes it impossible to learn how couples come to see HD as an issue for their reproductive decision-making and decide to contact clinicians.

Studies adopting a family systems approach produce useful insights into how risk about HD is experienced, perceived and communicated. Focusing on interactions within families, they reveal that crucial cognitions and behaviour are shaped long before reproduction is contemplated. Members may collude to 'preselect' one of their members, sometimes on the basis of lay beliefs about gender vulnerability or physical resemblance, as the future sufferer (Kessler, 1988). It produces an illusion of control and alleviates uncertainty about other family members' risk. This provides welcome reassurance if, as Korer and Fitzsimmons (1985) found, the statement that HD 'does not skip generations' is sometimes interpreted as meaning that someone from each generation *will* definitely be affected. Interpreting the same nominal risk quite differently for, and by, each family member, generates contrasting expectations about what comprises responsible decision-making (Downing, 2005).

Knowledge of HD also influences how family members relate with outsiders. Families can become either "HD orientated" (defined as becoming excessively preoccupied with HD) resulting in them leading socially isolated lives, or "independently orientated", with HD being given a less prominent role and families remaining open to contact with others (Korer and Fitzsimmons, 1985). These different perspectives affect how partners are informed about HD. Existing family members may collude to keep information from them. The nature and extent of information divulged, when this occurs, and the impact that information has on decision-making can vary considerably (Barette and Marsden, 1979; Shakespeare, 1992). Wives of HD patients felt vital information had been kept from them until *after* they had had children, and that they would have made different reproductive decisions if they had been fully informed (Hans and Koeppen, 1980). This could explain why Barette and Marsden (1979) found that over sixty percent of their respondents were prepared to tell their *children* not to procreate.

In order to comprehend the minutiae of how HD is experienced it is necessary to access personal accounts. Their evidence, though compelling, is often criticised as being of an anecdotal nature. The onus is placed on the researcher to impose rigor on how they are gathered and analyzed in order to decipher underlying commonalities yet preserve their diversity. The following aspects of the grounded theory approach adopted in the study help meet these requirements. These include grounding the analysis in the data, making analytic journeys explicit and moving *beyond* description of themes to develop an explanatory framework, or model, around the core category to facilitate subsequent evaluation (Hamburg et al, 1994; Boulton and Fitzpatrick, 1997; Creswell, 1998; Reid and Armstrong, 1998; McAllister, 2001). Spelling out the substantive area, or context, allows others to decide whether the framework is relevant in other situations, leading to transferability and the building up of general theories, (Glaser and Strauss, 1966; Seale, 1999). The next sections detail how grounded theory principles were applied to generate the model from the personal accounts gathered in the study.

## ETHICAL APPROVAL

The Local Clinical Research Ethics Committee, the Huntington's Disease Consortium and the Cambridge University Psychology Research Ethics Committee granted ethical approval for the study. All names and some details have been changed to protect confidentiality.

#### Methods

#### Recruitment and Characteristics of Participants

Theoretical sampling (to find those encompassing a range of relevant experiences and address the gaps noted earlier in this chapter) involved seeking out people at various stages of the reproductive process; of different ages and gender; with varied experiences of HD, Huntington's Disease Association (HDA) support groups and clinical genetics services; making a range of choices such as accepting or avoiding passing on the risk; and, of doing this in several ways, including having predictive, prenatal testing and undergoing termination of high-risk pregnancies. Seventy-six participants, comprising two generations, lateral kin and their partners within 17 kinships, were recruited to the study through the HDA, a Regional Clinical Genetics Centre and 'snowballing', or seeking chain referrals from initial respondents. Recruiting through family members offered a broader range of people the opportunity to opt in to the study than a purely direct approach by the researcher would have done. This technique provided access to people *not* in contact with either the clinical genetics centres or the HDA, and who would not have otherwise known about the study.

Characteristics of participants are shown in Table 1 below. The older generation had completed their reproductive decision-making before genetic testing was introduced. They could still play an active part in the younger generation's decision-making in the following ways: ascertaining their status through predictive or diagnostic testing; providing samples for linkage or prenatal testing; or by being supportive or critical of their children's decisionmaking. The younger generation, who formed the main focus for the study, were those making reproductive decisions prior to and at the time of interview. Some of them had made decisions before as well as after the introduction of genetic testing. Including siblings and cousins who face similar risks enabled the study to explore whether lateral kin provide guidelines for each other and how different responses are tolerated

Total N = 76 (generation $1 = 26$ , generation $2 = 50$ )				
Category	Details	Gen 1	Gen 2	
Gender	Male	9	21	
	Female	17	29	
Age	range	46-70 yrs	20-49 yrs	
relationship	Married	23	38	
	Single	0	8	
	Separated/ Divorced	1	3	
	Widowed	2	1	

Table 1. Sample Characteristics at Time of Interview Tatel N = 76 (concertion 1 = 26, concertion 2 = 50)

Category	Details	Gen 1	Gen 2
Reproduction	No children	0	25
	One child	4	13
	More than one child	22	12
	Intending to have child/ more children	0	23
RISK	At risk	4	20
	Partner at risk	1	14
	Affected	2	0
	Partner affected	12	0
GENETIC TESTING	Tested positive	0	2
	Partner tested positive	0	2
	Tested negative	1	6
	Partner tested negative	1	5
	Used prenatal testing	0	4 couples
CONTACT	No contact with Clinical	15	18
	Genetics		
	No contact with HDA	11	10

#### Table 1. (Continued)

Data Collection and Analysis

## **Preparatory Work**

Qualitative studies often start off with broad questions, such as in this study about how reproductive decision-making is experienced. They are also initially imprecise about sample size and type as well as the form results might take. This means that traditional 'pilot studies' are unlikely to be useful. Rather what is needed is a period of careful preparation in order to approach sensitive issues of interest (Lee, 1993). Preparatory work involves familiarizing oneself with the relevant literature. A multidisciplinary approach resulted in that the thinking and methodology of psychology was augmented with that of related disciplines, such as family therapy, sociology, social anthropology, medicine and genetics. Adopting a social anthropological approach helped accessing populations and becoming sensitized to their ways of thinking. As Yow (1994: 156) points out, it is helpful to first locate "the places where people (of interest) go to talk":

For the hospital history, I found it was the employee cafeteria. For the history of the women's cooperative art gallery, hanging around after gallery meetings and going to shows enabled me to learn a lot.

Accordingly, in addition to reading about HD, contact was made with people at the following locations: a regional clinical genetics centre, HDA support groups and organisations such as Sue Ryder homes and the Crossroads service providing care for those affected. Another reason for contacting the regional clinical genetics department was to establish how many HD families they knew to exist in the vicinity. Concerns had been expressed that it might be difficult to recruit enough families. It was reassuring (from a research point of view) to establish that, despite HD being thought of as a rare disorder, approximately 250 families had been seen there over the last twenty years. Contact with

families had varied and information about family structure was sometimes incomplete. Being given permission by clinicians to examine these records gave me an invaluable understanding of the work of clinical genetics departments in pre- and post-testing eras. I learnt that it was not unusual for clinicians to present an optimistic picture to those at the beginning of their reproductive life, setting up expectations that genetic testing would become available one day. This had lead them to encourage the use of temporary rather than permanent methods of birth control. It also became apparent that people had faced considerable practical difficulties in realising certain options, such as prenatal testing, which were first introduced in only one or two centres. Being in the department and sitting in on consultations informed me of current work, including preparation for the predictive testing programme. Contact with a second regional clinical genetics department refined these ideas.

Visits to Sue Ryder care homes and spending time with the local Crossroads family support service enabled me to learn what happens when families can no longer cope unaided. Regular attendance at one HDA support group's meeting, and joining in social activities with families over a period of a year and a half - which overlapped with the commencement of data collection - enabled me to 'hear the talk' outside medical settings. These meetings varied from informal chat to structured discussion of particular issues, such as reproduction. With the group's permission, I tape recorded some of these more focused sessions in order to reflect on them.

#### **Key Points Emerging from Preparatory Work**

The preparatory work enabled me to show the families that I had made some effort to understand their perspectives before we met for the interviews. It facilitated drawing up a list of core questions and phrasing them in terms reflecting participants' language. Hearing of the difficulties people experienced when talking about HD within families indicated that it would be advisable to interview each person separately. Stories told of affected parenting led me to realise that what are normally regarded as non-threatening questions such as early childhood memories (Yow, 1994), and which are therefore often asked first in interviews, might need to be approached later in the interview.

During the group discussions I first heard people prefacing their stories with comments about knowing or not knowing about HD, or of knowing but not really understanding what it could mean. I was to return to these observations when I subsequently identified 'awareness' as the causal condition for my model.

#### Subsequent Data Collection and Analysis

Taking part in the study involved completing a short questionnaire to gather background information and enable participants to state what they felt the study should address. They were subsequently individually interviewed at a place of their choice, normally their home. The semi-structured in-depth interviews were tape-recorded with participants' consent. A check list ensured that participants had covered key areas such as their reproductive decision-making and future plans, their own childhood, how they had become aware of HD, attitudes to and experiences of genetic testing and termination, what they felt the future held for those at-risk and decision-making by other family members. Interviews took between one and a half to six hours and were transcribed in full. Approximately six months later participants completed a follow up questionnaire or spoke with the researcher to report any developments

and comment on their experience of taking part in the study. Management of the data and analysis was aided by NUD\*IST, qualitative analysis software (Richards and Richards, 1994).

Decision-making has traditionally been plotted as flow charts, or decision trees which show how people apply decision criteria to choice alternatives and come to decision outcomes. This study was guided by a combination of ethnographic decision tree modelling (Gladwin, 1989) and mental model constructions (Bostrom, Fischhoff and Granger Morgan, 1992) instead. These approaches were thought more able to encompass the variety of ways that decision-making is actually structured and to meet grounded theory guidelines.

Ethnographic decision tree modelling employs qualitative methods to learn how decision makers perceive decisions and structure the decision-making process. Questions asked by the interviewer vary for each participant, but the claim is that it is possible to incorporate these into decision trees which encompass individual variations, interrelated sequential decisions and feedback loops that the decision-making process involves. Content analysis of participants' narratives distinguished key concepts. The method of constant comparison then revealed the different meanings these concepts hold for people. Decisions were then represented by creating simple mental models linking these concepts, using an approach originally developed to compare lay and expert perceptions of environmental risk (Bostrom, Fischhoff and Granger Morgan, 1992). Models were then compared with those from other individuals, such as partners, involved in the decision-making process to search for overlaps, contradictions and omissions. These were refined into the model of responsibility presented below which encapsulates what families facing a late-onset genetic disorder, HD, find important when make reproductive decisions in the face of changing options being presented by developments in genetics and reproductive technology. The detailed analysis that generated the model will be presented next as part of the Results section.

## RESULTS

This section provides an overview of the model, discusses the analysis that generated it, and then outlines how it provides a framework to account for variations in what people do. A subsequent section discusses what emerged about the process of decision-making.

#### Overview of the Model of Responsibility

The grounded theory analysis was facilitated by applying Strauss and Corbin's (1998) paradigm model, which it was felt could accommodate the process of reproductive decisionmaking. In the following summary the components of their paradigm are placed in parenthesis and the form these take in the study model are italicised. The 'causal condition' is one that provides the context in which the phenomena of interest occur. Comparisons of aware and unaware decisions identify this to be *awareness* of the risks for HD. A key aspect of the model is that it encapsulates the two dimensions of reproductive risk identified above: that which may be passed on to any children born, and that of not being able to sustain a parenting role. These are the *child-centred* and the *parenting* risks. The 'consequence' of becoming aware of these risks is that *reproductive decision-making becomes defined as*  *problematic.* This prompts people to *redefine* certain elements of their situation - such as the nature of the risks, themselves, fertility, and relationships. The form this 'strategy' takes reflects 'modifying factors' such as the values that people hold, their concept of the future and their perception of social support. The 'consequence' of redefining is to enable people to tell a different story when subsequently *accepting*, *modifying* or *avoiding* one or both of these risks. Concerns about one value, responsibility, dominate these stories - identifying *responsibility* as the 'core concept' for the model. A key aspect of the model is that it demonstrates that *how* people are perceived to make decisions can become as important as *what* they decide.

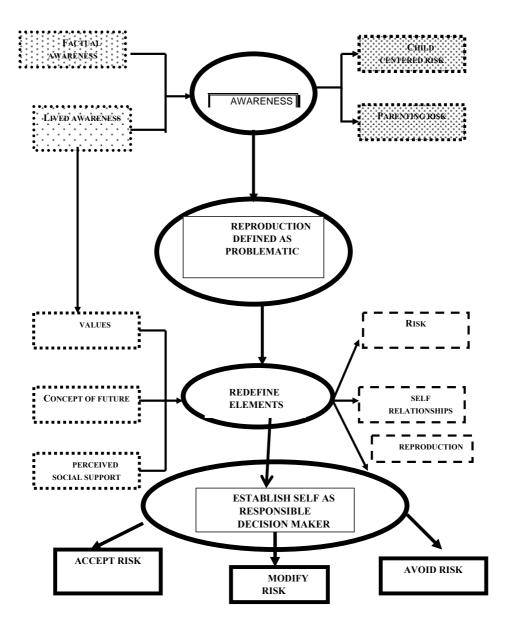


Figure 1: Model of Responsibility

#### Concepts Associated with the Model of Responsibility

Key aspects of the analysis underlying the model are:

- responsibility
- the nature of awareness
- the impact on decision-making of growing awareness of the risks that HD presents for reproduction;
- coping strategies such as redefining
- the role and nature of modifying factors such as individuals' values, perceptions of the future and social support

#### Responsibility

Responsibility is shown to be both a private and a public concern. Fundamentally people want and need to be seen as responsible decision-makers, by themselves and others. What this involves becomes apparent from exploring how people account for their decision-making in the face of growing awareness of HD.

#### Awareness

Awareness refers to the subjective perception of the risks that HD presents for reproduction, which both reflects and influences how people engage with them. Comparing new examples with previously coded extracts generated sub-codes of *factual* and *experienced*, or lived, dimensions. Factual awareness of HD corresponds to textbook or medical models, incorporating (in participants' words) just the "blatant black and white facts" such as diagnosis, prognosis and how genetic mechanisms work. In contrast, lived awareness feels "sort of arbitrary", encompassing "anything that happened" as well as emotional reactions. Becoming fully aware necessitates knowing in both factual and lived dimensions and integrating these perspectives.

#### Factors Facilitating or Limiting Awareness

Awareness is a *temporal* process in that it accumulates over time. It can also arise retrospectively, as hindsight. The ongoing analysis clarified links between responsibility and increasing awareness. One woman's attempt to reassess her mother-in-law's response to her first pregnancy shows how hindsight shifts the focus of responsible decision-making to others and introduces a contrasting facet, notions of blame.

When you think back, how, how sort of gob-smacked she WAS when I told her I was expecting. We'd been married four years, so, you know, a different reaction altogether to what my mother was - my mother was over the moon. ... Looking back now, did she know something? And, was she in a quandary whether or say, or not?

Her reflections identify a second factor, *communication*, which impinges on awareness. All of the study families - including those who could communicate freely about most subjects - reported difficulty in talking about HD, supporting the initial fieldwork memos.

Further analysis revealed that barriers to communication took different forms in each generation. The older generation spoke of those affected being hidden away by their families. Their reluctance to make the condition known was not altogether surprising in view of the 'social' legacy of opposition to their existence and reproduction noted above. Another explanation to emerge was there was less incentive for the previous generation to consider the outcome of their decisions because they held different expectations about life.

Then, if you got to fifty you were lucky ... You might have died of pneumonia, or TB, or half of the children that had Huntington's could have been stillborn anyway. ... It's only NOW that you expect to have longer quality of life that it's become such a problem. That you feel you're entitled, you know, you sort of say, "Oh, I'm not even going to get fifty years."

Older generation participants who *had* known about HD reported information being linked in uncompromising terms to reproduction. They were told either not to have children or to go away and "get on with things". In the latter case this could be interpreted as not questioning widely held expectations that marriage would lead to having children.

The younger generation reported that their parents had waited divulge relevant but limited information until their teenage years or when they had formed serious relationships. Now that HD is cited on the school science curriculum families report feeling obliged to discuss it earlier and in more depth than they might otherwise have done.

Finally, individual variations in *information accessing and processing skills* affected the speed with which awareness occurred. This included the prior establishment of schemas - for example, from professional experience, or from contact with sufferers of other illnesses - with which to make sense of information as it emerged. Such schemas could both help those at-risk to explain what HD involves and partners to make sense of the information.

#### Cognitive Processing

Distinctions participants made between decision-making before and after becoming aware of HD revealed *how* it changed with growing awareness. People commented that before becoming aware "there had been no decision to make" that was any different to other people's reproductive decisions, i.e. taking into account desires, financial and travel plans and waiting for the right time (Currie, 1988; Willen, 1994). Accounts of aware decision-making drew on physical metaphors of burden, such as weight and size. These metaphors conveyed the effort required to weigh, balance, or set aside the child-centred and parenting risks associated with HD. As had been surmised, like those facing infertility cited in the introduction, references were made to a sense of loss - the loss of normal expectations for reproductive decisions into *difficult* ones. Difficulty was compounded by complex and multiple analytical concerns, such as those around termination and determining risk status. Decision-making also felt onerous because as awareness increased people felt obliged to account - to themselves and others - for their intentions and actions. This observation provided an early indication that responsibility would become the core concept for the model.

#### Clarifying Links between Awareness, Responsibility and Blame

The process of becoming aware comprised five stages. Who could be held accountable, and how, varied at each stage. In the first *unaware* stage there was no conscious consideration of potential problems in the family. Information may have been available - as can be seen from comments made by other family members - but not divulged. Other explanations were that it was not consciously thought about or that people were not able to put it into words. Consequently, there was no recalled need for explicit consideration of risk in reproductive decision-making. If people had had children at this stage they could feel happy in retrospect that they had 'not really known' and could not therefore themselves be held accountable. Guilt could subsequently be felt about having unwittingly passed on the risk or not having pursued enquiries. However, as the example cited above showed, not knowing could also permit guilt to be displaced as blame on to other people who it was felt had been in possession of information and had not passed it on.

In the second stage, *pre-awareness*, concerns began to be expressed about whether problems existing within the family might have implications for reproduction. This could occur when questions were asked about family illnesses in prenatal clinics. It was not unusual, especially in the older generation, for HD to be attributed to non-genetic causes, such as the menopause or shellshock, which had no implications for other family members. If this had been the case, reassurances from the medical profession enabled reproduction to resume on a 'normal' track. People could subsequently blame those responsible for misdiagnoses or themselves for unquestioningly accepting them.

The third stage, *limited awareness*, arose when HD had been diagnosed in a family member but there was limited factual understanding of the implications this had for reproduction and for other family members. As Daniel explained:

You know, Huntington's disease didn't sort of mean anything. It's just - a word.

As in the previous stage concerns arose in connection with reproduction. How health care professionals reacted to the mention of HD then became the crucial factor in determining how seriously these concerns were taken. Naomi held her doctor accountable for having had a child at-risk in that she felt he had not ensured she and her at-risk husband had understood the magnitude of the risks:

He SHOULD'VE warned us, properly, and MADE us understand, how serious, it all is. And if he had the SLIGHTEST doubt that we didn't understand, he should've SAID it again, and again and again, until that, he was fully satisfied that we, understood FULLY what we were about to enter into.

The fourth stage, *detailed awareness*, encompassed comprehensive understanding of the HD risks. Engaging with the risks associated with HD became a central part of reproductive decision-making. Being aware made it much more difficult to delegate or deny responsibility for decisions that were being made. There was a need to show that one was unable rather than unwilling to act in what might be perceived as a conventional responsible way (Finch and Mason, 1993). Intervening factors (see below) then became crucial in construing responsibility. Many accepted the risks, supporting previous suppositions that although most people in Huntington's families have, or intend to have children, few choose to employ

strategies such as prenatal testing that could ensure that these children would not be at-risk. Modifying strategies comprised timing of reproduction and determining family size. People could even have *more* children than they might otherwise have done, in the hope that at least one had not inherited the mutation and could care for any affected siblings. Avoidance strategies included not having children to seeking gamete donation, clarifying risk by testing and termination and seeking (unsuccessfully) to adopt or foster.

Some participants experienced a fifth stage, *refined awareness*, when risks became certainties as a result of genetic testing or diagnosis. A few chose a non-disclosing form of prenatal testing, prenatal exclusion testing (PNET) which did not clarify the parent's status but enabled them to have children known to be at low risk. This form of testing is complicated to understand. It requires parents to terminate any 'high risk' fetus known to be at the same fifty percent risk of having inherited the HD mutation as the parent. This means that there is a fifty percent chance that it does *not* have the mutation. It works by comparing markers on the gene from three generations, to see if the fetus has inherited the at-risk parent's copy from the affected or non-affected grandparent. The idea becomes to only proceed with pregnancies shown to be at minimal risk in that they have almost certainly not inherited a copy from the affected person. For a detailed account of how couples live with their decisions under these circumstances see Downing (2005). Donating samples to inform testing enabled the older generation to retrospectively define themselves as responsible decision-makers.

Growing awareness also involves acknowledging how reproduction becomes embedded in a 'HD life cycle'. This consists of the following events and stages: being born at-risk, being parented by an affected parent, discovering the risk, developing ways of living with risk, deciding whether to tell prospective partners about HD as a preliminary to reproductive decision-making, undergoing changes in risk status arising from diagnosis and predictive testing, becoming a suffer or carer for parents, one's children or other relatives, and coping with death and loss. People do not necessarily experience all of these events, the order in which they occur can vary, and events may coincide. For example, some people only learn that they were born at-risk when an affected relative died. Learning how HD events are experienced, resolved and map onto the 'normal' life cycle (Erikson, 1963) can elucidate the varied responses that people show to the reproductive risks for HD. As intimated above, how and when people discover their risk seems to affect how HD is taken into account when subsequently making reproductive decisions. Being brought up with the knowledge that HD is in their family, enabled some people to assimilate having at-risk children with a 'normal' life trajectory. Others, appraised of their risk at a later age, found it more difficult to accommodate reproductive risk with normal expectations for reproduction.

#### Redefining as a Response to Awareness

Redefining denotes a strategy of restructuring elements of people's situation, either cognitively or through actions, in order to have a different story to tell. New narratives incorporate changes in their perceptions of risk, fertility, themselves or their partners, and reproduction thereby enabling them to account for having accepted, modified or avoided either or both of the risks in a responsible way.

Deconstructing participants' references to *denial* - a powerful form of redefining - revealed three ways this could manifest. The first corresponded to the accepted clinical psychological definition as unconscious repression of painful information, such as learning of

the risk. The second, 'affected' denial, was associated with the medical onset of HD. The third and most common way was a conscious strategy of 'distancing' oneself from threatening information. In contrast with clinical denial people could acknowledge this was happening and unwanted information remained accessible. Distancing was evidenced by HD only being spoken of in connection with more distant kin, such as cousins, or in relation to physical distancing, or severing contact between affected and non-affected kin.

Resorting to distancing was a function of the degree of fear engendered by the risk, control felt over it, and individual differences in the amount of information that could be tolerated about HD. If distancing was widespread in families, a *closed awareness* context resulted - similar to that identified by Glaser and Strauss (1966) when studying dying patients. As Angela explained, family members knew about HD but colluded in not openly acknowledging or discussing it.

We used to live a life of pretending it isn't happening. .... When we were at home, we used to react as if nothing had happened, and it became like a family trait. If dad choked on some food, you'd sort it out, organize it, and then carry on with the meal as if nothing had happened.

A closed awareness context had considerable and contrasting implications for reproductive decision-making. It provided the freedom for people to take risks without fear of voiced criticism. But, as Frances explained, it also made it hard for those like her marrying into the family to find out what they felt they needed to know in order to make an informed responsible decision.

Every time I've - I tried to talk about Huntington's, to the family, I'm more or less told, "Yeah okay, all right, you know, forget about it, shut up." They don't - they just - just don't want to know.

It also became difficult to implement choices, such as prenatal or linkage testing, which depended on other family members acknowledging the need for their assistance. As Tara said:

We had to plead with him to give his blood, for the comparisons, for mine and my brother's test, because he didn't want to know, and he didn't want us to know.

As well as denying risks, narratives could redefine them in ways that appeared to lessen them or clarify them through genetic testing. Those at-risk, or their partners, whose values allowed them to accept a "semi-biological" child could redefine themselves as temporarily infertile in order to gain access to new reproductive technologies and bypass their genetic risk. Reproduction could be redefined as a challenge to the disease, or as replacing an affected person who had died. Others redefined partners as potential single parents so as to counteract the uncertainties surrounding their own ability to sustain a parenting role. Parenting could also be redefined, for example, through becoming an uncle or by caring for children in a professional capacity which would enable a person to relinquish responsibility on becoming symptomatic.

The versatility of redefining is illustrated in the following examples relating to unplanned pregnancies. Redefining them as chance events made it possible to abdicate responsibility. Maria's comment was typical of this not uncommon response:

Well, if I did just happen to get pregnant ... the child can't turn round to me and say, "You shouldn't have done this to me," because I'd say, "Well - it just happened.

Choosing not to proceed with an at-risk pregnancy enabled one to be redefined by others as acting responsibly. Sabina explained what else she had gained from her unplanned, and terminated, pregnancy. Making decisions about another life allowed her to redefine *herself* as a responsible caring mother, through having:

the experience of GENUINELY meddling in another person's life and shaping it irrevocably so that it can never be as it would have been without my interventions ... knowing that you have taken upon yourself that kind of assumption of responsibility which you have, as a parent, I haven't missed. The, the knowledge of being wholly responsible for life and death of another individual umn, I had with deciding to have an abortion. .... I talked to Sam - Sam's the guy, I named him / her, and sort of explained what I was doing, and all the rest of it, and explained why it was, this was the best life for that person to have. So, I did all of that.

Redefining thus illustrates how people make creative and adaptive responses to difficult dilemmas, consistent with naturalistic decision making (NDM) accounts of personal decision-making (Payne, Bettman and Johnson, 1993).

## The Role Played by Modifying Factors

Modifying and interacting factors which determined *what* and *how* people redefined aspects of their situation comprised: values held about a range of issues, how the future was conceptualized, how social support was perceived and factors subsumed under the heading of individual differences.

## Values

It became crucial for those seeking to justify accepting either or both of the risks to demonstrate that they had been *unable*, rather than unwilling, to have acted otherwise. This can be seen from the references to limited awareness that prefaced many of the accounts of acceptance of genetic risk. After becoming aware of the risks, acceptance began to be justified in terms of the values that the participants held - for example, it was accepted that those known to be strongly opposed to abortion were unlikely to consider prenatal testing.

The grounded theory analysis identified three clusters of interacting values which people found salient in their decision-making. Table 2 indicates how different aspects of these values, relating to reproduction, quality of life and decision-making, influenced reproductive decision-making.

As had been suggested in the introduction (Dallos, 1997; Nelson and Nelson, 1995) values started to form long before any thought of reproduction. Their impact depended on how peripheral or central they were for people. This could vary at different points of time and with the introduction of each person involved in decision making. Those who saw reproduction as central to their existence found it difficult to think of a life without children and easier to downplay other considerations. Flora explained:

I wanted children MORE than I thought the risk would be to her.

Reproductive values, related to	Quality of life values related to	Decision-making values related
feelings about:	perceptions & experiences of:	to the importance placed on:
knowing that you could conceive	a life lived at-risk for HD	Making decisions
or make your partner pregnant		
having a child	Living with definite knowledge	risk taking & avoiding
	of HD status	
having an at-risk child	a life affected by HD	information seeking and
		avoiding
Having a biological child	being parented by an affected	Identity as a responsible
	parent	decision-maker
abortion	the parenting that one might	
	offer to a child	
Envisaged identity as a father or	Alternative opportunities for	
mother	fulfillment of goals seen as	
	related to reproduction	

Table 2. Values Shaping Reproductive Decision-Making in the Face of HD

Conversely, men and women who placed less value on having a child found it convenient to draw on HD as a responsible reason for remaining childless. Matthew's comment was typical of this standpoint:

Plenty of other people are busily producing babies. And if you are unsuitable to provide them, why worry?

More frequently decision-making entailed resolving central but conflicting values, such as wanting to have biological children, not wanting to pass on the risk, valuing an at-risk life, and being opposed to abortion. Thus certain strategies such as prenatal testing or utilizing donated gametes, could be screened out by the values that people held. For a detailed analysis of uptake of various risk limiting and avoiding strategies according to the values that they involved see Downing (2002).

## Vision of the Future

How people viewed the future affected whether or not they were prepared to accept the risks. It also illuminated how some people (who would have been prepared to accept termination of pregnancy) justified rejecting genetic testing. They construed availability of testing as an indication that treatments for HD would follow in the near future. Envisaging the future in terms of *promise* made it easier to redefine the risk as acceptable and reject testing or other risk avoiding strategies. Hearing about on going research into possible treatments fed into this optimism.

Others, more cautious about the future, were more likely to take steps to avoid passing on the risk. Whilst they were able to accept their own existence as a result of the previous generation's limited awareness they doubted that the next generation would necessarily be so tolerant. They saw the future as a time of *reckoning*, when they would become accountable for the decisions that they had made. Their priority was to be seen as responsible decision-makers by their future as well as existing families. As Rupert explained:

I would really HATE to have the discussion with a child of mine which went: "Why did you let me live when you knew I had this disease?"

How the future was conceptualised reflected concepts of agency. Many participants holding a 'chance' perspective of the parenting risk saw their future as predetermined and their role as reactive. Others holding a 'choice' perspective saw themselves as shaping the future. They deployed options such as genetic testing to clarify and control their risk of becoming an affected parent, or had children earlier than they might otherwise have done in the hope that they would be grown up before the parent became symptomatic.<sup>4</sup> This dichotomy accorded with well established psychological concepts relating to internal and external locus of control (Rotter, 1966; Wallston, Wallston and DeVellis, 1978; Wallston, 1997). It clarified contrasting constructions of responsibility in connection with the parenting risk. Those holding chance perspectives took a reactive perspective of responsibility, seeing it as coping with whatever fate presented. Conversely those holding choice perspectives became proactive about responsibility, employing interventions such as predictive testing or gamete donation in order to inform and control outcomes. It became possible to explain why people might not act consistently in regard to the two risks, seeing the parent's risk as predetermined and that for any child as a matter of choice.

## Perceived Social Support

This factor similarly started to shape decision-making long before any thoughts of reproduction. One interesting finding was that those growing up in families affected by HD were more willing to accept the uncertainty associated with their ability to sustain a parenting role *if* they had experienced a supportive sibling relationship. Conversely, only children, and participants whose siblings had not been perceived as supportive, reported feeling more vulnerable and less willing to take on this risk.

Support concurrent with decision-making ranged from endorsement of, or practical assistance with choices made to tolerance of differences. This was more likely to be forthcoming if people could demonstrate that they had acted as responsible decision-makers, or in a manner consistent with their central values. Those whose partners, families and friends were able to show non critical support for whatever path had been chosen acknowledged how helpful this had been. Case studies drawing on the model as a framework illuminated how support changed with repeated decision-making and could come to involve clinicians as well as families (Downing, 2005).

## Individual Factors Contributing to Decision-Making

Individual factors that contribute to decision-making styles include differences in information processing approaches, corresponding to the monitor (information seeking) and blunter (information avoiding) categories suggested by Miller (1987). It has been suggested that these differences are indicative of general patterns of responding to risk (Lopes, 1987), and attitudes to decision-making (Beattie, Baron, Hershey and Spranca, 1994).

Maturational and physical characteristics, such as age, risk status, fertility and gender, also impacted on decision-making. Many participants reported that they had found it easier to distance themselves from the risk when younger. Being able to conceive easily encouraged

<sup>&</sup>lt;sup>4</sup> These terms were coined by Marshall (1995) to account for contrasting attitudes to new reproductive technology.

those undergoing prenatal testing to continue along this path (Downing, 2005). How long a couple could wait for a risk-avoiding strategy to become available, such as genetic testing, was to some extent an interaction of age and gender, in that men could contemplate reproducing at a later age than women.

## Impact of Genetic Testing on Reproductive Decision-Making

A key aim of the study was to explore the impact that the introduction of genetic testing has had on reproductive decision-making. Support was shown for previous consistent findings concerning the reluctance of many facing HD to use predictive testing, seeing it as a *threat* to the hope that uncertainty conveys about their risk status.

The study augments existing work by revealing additional perceptions that people have of genetic testing and how these facilitate or constrain decision-making. Construing genetic testing as *reconciling previously incompatible goals*, such as having biological children and avoiding passing on the risk, facilitated decision-making. Similarly, perceiving testing as an indication of technological and medical progress, implying that treatments for HD would become available in the near future, made decision-making feel less constrained. Participants holding this view could justify *not* employing testing and accepting the risk. Conversely, participants who perceived genetic testing primarily as a *technological imperative*, imposing a responsibility on them to employ it in order to make an informed decision, experienced their decision-making as more constrained. However, because using genetic testing was likely to be interpreted by others as exercising responsibility, support was shown by other family members and clinical staff which mitigated this feeling of constraint.<sup>5</sup> It was not always necessary to employ testing to be judged as acting responsibly. Considering testing was sometimes sufficient to ensure support (Downing, 2005). This observation contributed to the key finding that how decisions were made can be perceived as being as important as what is decided.

The study also makes a significant contribution to what is known about the ways in which deciding to undergo genetic testing changes family dynamics (Roos et al, 1990; Quaid and Wesson, 1995; Cox, 2002, 2003; Sobel and Cowan, 2000). It was shown, for example, that requests for samples to be used in linkage and prenatal testing of other family members enabled some people to view testing as providing an opportunity to *alleviate the guilt* they felt about having unknowingly passed on the risk, and to retrospectively redefine themselves as responsible decision-makers. Conversely, envisaging genetic testing as a *threat*, in that important information about oneself might be revealed to others, made people less likely to cooperate in providing samples and could result in the breakdown of relationships within families.

The impact of genetic testing needs to be seen as a function of which *form* of testing is being employed or considered. Linkage predictive testing, for example, had a limited impact because it was not available for all who considered it. Another constraint encountered in connection with predictive testing, both linkage and mutation, was that this was seen as a decision that could only be made by those at-risk. In contrast, pursuing prenatal testing was seen as a decision involving *both* partners. This observation was an important one as it

<sup>&</sup>lt;sup>5</sup> Although the word 'mitigated' suggests that constraint has a negative effect on decision-making, this may not necessarily be the case. Constraint may help people arrive at a decision by reducing the alternatives.

challenged the appropriateness of the widespread tendency in the literature to refer solely to women when speaking of decision-makers.

Partners of those who did not want to know their status faced making uninformed decisions or having non-disclosing exclusion prenatal testing (PNET). Uptake of PNET has been very low worldwide (Adam et al, 1993; Maat-Kievet et al, 1999; Hayden, 2000, Simpson and Harper, 2001). Fewer than ten prenatal tests were done in Canada between 1990-1995, and the rate in the USA was similarly low. It was therefore a significant achievement for this study to have accessed ten couples who seriously considered employing it, and four who actually used it. Their experiences were very informative about the decision-making of those considering but not using prenatal testing, the problems associated with negotiation of its use, considerations arising with its repeated employment, and the impact of different outcomes.

Some of the ways that using prenatal testing made a difference - such as creating tentative pregnancies (Rothman, 1986) whilst waiting for results, setting up assumptions that people will act on the information received and terminate wanted but high risk pregnancies, or having to involve others in decision-making - were not unique to HD. Neither was the experience of two of the couples who had to cope with the subsequent miscarriage of pregnancies shown to be at low risk (Robinson et al, 1991). The predicament associated with PNET of terminating 'high risk' fetuses which may not be affected, was also not a new one. This dilemma has been previously documented in connection with Duchenne muscular dystrophy and other sex linked conditions. Prior to the introduction of genetic testing some people chose to terminate any male fetus (a proportion of which would not have been affected) in order to ensure that they only had a child that they knew would not be affected (Parsons and Clarke, 1993).

The low uptake of prenatal testing observed in this study appeared to be influenced by weighing these issues in relation to a risk that would almost certainly not manifest until adult life. What the study revealed was that participants were faced with considerations of the quality of a 'tentative life', that is, the life to be lived at-risk. This meant that those who *would* have been prepared to consider termination for early onset disorders could have difficulty in seeing it as appropriate for HD because they saw their, or their partner's, life as worthwhile.

Interesting findings emerged about repeated reproductive decision-making. Those who did not initially use prenatal testing, but then employed it for subsequent pregnancies, were able to accept that their decision-making would result in them having different information about the status of each child. However, couples who started to use a certain type of prenatal testing felt compelled to use it for subsequent pregnancies. This effect was apparent for individuals who had both high and low risk results. What this meant was that couples continued to employ PNET even after a more informative and predictive test, direct mutation prenatal testing, became available. Those who had previously terminated on the basis of a high risk result stated that the distress experienced was preferable to being confronted with the realization that, in the event of the parent having a negative predictive test, termination had not been necessary. This finding illustrated the limitations of classic decision theory models which fail to take account of the input from past decisions. It was consistent however, with studies that have showed what has been called the "sunk cost" phenomenon (Arkes and Blumer, 1985). This refers to a situation that arises favouring continuing with a pattern of behaviour because the investment made - which can be emotional, financial or practical makes it difficult to assess subsequent decisions independently.

Participants saw genetic testing primarily as an extension of *existing* methods - such as sterilization and using donated gametes - offering control over risks. This was compatible with Strathern's (1995) observations of the different ways in which people make sense of new reproductive technology, seeing it as either innovative or linked to established ways of having children. It suggests that there may be underlying consistencies in how people make sense of a range of technological changes.

Conceptualizing testing as an extension of existing methods of control generates additional questions. First, to what extent do the various technological developments offering control over reproduction and risk - which includes improved contraception, early identification of pregnancy and prenatal testing techniques - enable people to actually *feel* more in control of reproductive risks? Perception and tolerance of residual uncertainty associated with testing varied. One of the most interesting findings to emerge from the study was that perceived lack of confidence in *low* risk PNET results formed a significant barrier to its use. Doubts about the reliability of low risk results could also arise later, after repeated high risk results, through actual experience of laboratory error, or from learning that testing for *other* conditions, such as HIV, was fallible. These findings are illustrative of cognitive biases, such as the availability bias (Tversky and Kahneman, 1973), which have been previously well documented in relation to a range of risks (Kahneman, Slovic, and Tversky, 1982). The key point to be made in relation to HD was that although experiences such as these could engender uncertainty about the reliability of testing, they did not necessarily dissuade people from employing it.

A related question arises about how much people *want* to feel in control of reproduction. It was interesting to note that unplanned pregnancies did occur to both those accepting and wanting to avoid passing on the risk. Defining a pregnancy as a chance unplanned event - for which one cannot therefore be held accountable - was shown to facilitate acceptance of the risks. Another indication of ambivalence relating to control was that those who felt strongly about not passing on the risk did not necessarily employ contraception, or, if they did, did not use the most reliable methods.

The few who had learnt their status felt that the increased information clarified their subsequent decision-making. Previous research had suggested that learning there is no longer any risk appears to facilitate implementing previously expressed reproductive intentions. Decruyenaere, Evers-Kiebooms, Boogaerts, Cassiman, et al (1996) et al (1996) showed that within a year of the partner testing negative all 13 couples had had a child, were pregnant or were intending to have a child. In contrast, 13 other couples who learnt that the partner had tested positive had decided not to have a child or more children, were contemplating reproduction only in conjunction with prenatal testing, or were still undecided. However, what people then *did* with this information varied considerably, adding further support to the claim that people do not necessarily prioritize exercising control over the risk in their decision-making. This finding supports that from another study which interviewed couples after a longer time period to show that reproductive behaviour of those who test positive is complex and does not always correspond to pre or post test intentions (Tibben, Frets, van de Kamp, Niermeijer et al, 1993).

## Applying the Model

The model provides a framework in which to re-examine the observed complexity of decision-making. This involves systematically comparing and accounting for:

- variations in responses shown between people, including partners and other family members
- variations in responses to each dimension of risk
- choices of risk avoiding or modifying strategies
- responses to change such as increasing awareness of HD, pre and post testing reproductive decision-making, decision-making with different partners
- making decisions before and after becoming parents
- decision-making at different stages of the reproductive decision-making process including how decision-making becomes shared.

A comprehensive account of these applications is given in Downing (2002). A few examples are presented here to indicate the versatility of the model. The first deconstructs why first and subsequent pregnancies presented different issues, making them feel like different decisions. First pregnancies involved framing decision-making as about *whether* to reproduce and the extent to which identity is shaped by becoming a father or a mother. Families were more supportive of those at-risk having one child for this reason. Making decisions about subsequent children became more complicated - involved reconciling expectations voiced by others outside the family about 'having another one' with diminishing support from other family members, the projected needs of existing children and the at-risk parent, from the perspective of being a parent. Having additional children therefore included acknowledging these needs and presenting oneself as addressing them in a responsible way. Becoming a responsible decision-maker involved deciding which of these needs would be prioritized.

The second example, set out in Table 3, uses the model as a framework to compare different responses to the parenting risk.

	ACCEPT	MODIFY	AVOID
VALUES	Identity as a parent, parenting as "natural"		Value child's needs over need to reproduce
FUTURE	optimistic, focus on present ability		cautious, focus on future limitations
SOCIAL SUPPORT	compensating, forthcoming from partner and previous generation	compensating, forthcoming from partner and previous generation	diminished, children and partners seen as having competing needs
REDEFINING STRATEGIES	risk splitting: between child and parenting risk seeing parenting risk as predetermined self as single parent or dependant child	timing: redefine oneself as a responsible informed decision-maker by having children earlier to lessen possible impact of affected parenting.	risk linking: redefining self as responsible by drawing comparisons with others as selfish redefining at-risk parent as risk free through predictive testing.

# Table 3. Using the Model Of Responsibility to Account for Accepting, Modifying or Avoiding the Parenting Risk

The third example illuminates what is different about risk splitting and linking and provides insights into how couples' decision-making became shared. As Beach (1993; 1996) had suggested, individuals engaged in preliminary restructuring of their situation and options to reflect their own values before negotiating with partners. Negotiation involved checking whether their values were compatible. Difficulties could arise if partners were not in accord - as Ingrid's comments revealed:

You know, I'm not prepared to let those children have what he's had. .... He never agreed with me on that - he said, I should have children at-risk. He didn't see that it was that bad. He said, "I've had a good life. You know, I'm thirty seven. A lot of people don't get to that age, if you're over in Rwanda or something." But maybe, I wonder whether, if, he's in a different position to me, isn't he? Because he's <> if he agrees that you shouldn't have kids, maybe he feels that his life isn't, wouldn't have been worth anything. I mean he'd have been, you know, I don't know if he'd have wiped out, because I don't know the, what how his result would have been.

In the event of disagreement about prenatal testing and terminations the woman's views prevailed. As she explained:

One of you's got to do it, you see, and I, I decided it was MY choice as far as whether to have the tests for the baby, and my choice whether to have a termination and you can't compromise on that. You either do it or you don't do it. So my husband didn't really feature in it, which was TERRIBLE.

For most couples decision-making about predictive testing was seen as the prerogative of the at-risk partner. Focusing on non-disclosing prenatal testing and termination enabled Ingrid to 'split' the risks and redefine both risks: that for the child as negligible and that for her partner as less stressful:

I'm cutting down on my anxiety so it's quite selfish, in a way. ... I want to reduce my anxieties, see? And, at the moment, I'm only worrying about my husband. I don't want to worry about them as well.

In contrast, Martin and his at-risk partner agreed in their views that the risks were firmly 'linked'. He explained, that for them:

The logical approach is Mum wants kids. Mum is at-risk. Mum has the test. If mum's got it, mum cuts the tubes. But only when you've cleared mum, then the whole thing -then, and then you're back to normality. ... UNTIL - until somebody comes up with some - is it cystic fibrosis where they've now got something you can pop up people's noses to do wonderful things against the gene? Fine. If somebody can come up with, with some wonderful thing which takes all your, all your cells out, scrubs the genes over, shakes them into order and bungs them back in again.

## His partner concurred:

If you're not prepared to test yourself, you don't - no, absolutely. ... If you test the child, and then you get Huntington's, that's out of order too - as far as I'm concerned. You DON'T

put a child in a family, knowledgeably, within a Huntington's family. ... Part of the reason you're not having children is not only to pass it on, but not to inflict Huntington's on them. So therefore you should be testing yourself. And if you got a positive ... you don't have them. It's ABSOLUTELY simple.

The model has also been used to clarify clinical encounters with multiple decisionmakers. It is being used as an on-line teaching tool in conjunction with case studies exploring decision-making about prenatal testing (Downing, 2005).

## The Process of Reproductive Decision-Making

The study generated the first detailed account of the *process* of reproductive decisionmaking in the growing awareness of HD. This was shown to involve a multiplicity of decision points, requiring people to draw on a variety of resources to engage in private and shared considerations of a number of areas. These included how genetic risks are perceived, the responsibilities that genetic risks entail, how technological developments are assimilated and, as Richards (1993) and van der Steenstraten et al (1994) had suggested, how identity is constructed. This section provides an overview of what emerged about the four stages comprising the decision-making process: pre-contemplative; pre-implementation; implementation; post-decision evaluation.

## **Pre-contemplative Stage**

The qualitative approach adopted in the study showed the importance of acknowledging an additional, pre-contemplative, stage where no knowledge might be held of HD, nor any considerations entertained about procreation, but that shaped later reactions by forming individual attitudes and values. Support is therefore shown for NDM claims claim that additional insights into the decision-making process will be provided by links made to other domains, such as narrative approaches (Nelson and Nelson (1995) which provide access to how values are formed in families.

This stage could encompass being parented by a parent who was affected but not yet diagnosed. Connections were shown to exist between how affected parenting was experienced, intervening factors such as social support, and the subsequent value placed on avoiding this risk for future generations. Narratives from different family members confirmed the concept of preselection in connection with risk status (Kessler, 1988) and extended it to incorporate particular children being singled out for *abusive* treatment. Younger children in families were more likely to be exposed to affected parenting, but this could be mitigated by sibling support. Only children were the most vulnerable. This meant that affected parenting as well as risk could be *interpreted* quite differently by each family member leading to them placing different values on avoiding this risk for the next generation.

## **Pre-Implementation Stage**

The second, or pre-implementation, stage of decision-making was initiated by becoming aware of the risk and culminated in forming potential responses to it. Intermediary steps involved at this stage were:

- becoming aware of the risks that HD presented;
- assessing the impact of these risks on reproduction;

- defining reproductive decision-making as a difficult decision;
- redefining reproductive goals and evaluating them for compatibility with values held;
- evaluating strategies for their ability to realise these goals and their compatibility with values;
- mental simulation of implementation and outcomes;
- re-assessment.

Decision-making became shared at this stage. Comparison of different family members' accounts revealed how this happened. As the example presented in the previous section shows the findings supported Beach's (1993) claim that individual processing preceded negotiated, or social, decision-making. People considered the issues in relation to the values they held and these values shaped the way they subsequently presented them to others. I had speculated that partners, because they had not shared the same family experiences, might hold different values about passing on the risk, or ability to parent. This was shown not to be the case. Members of a couple did not necessarily hold the same values but acceptance, modification or avoidance responses were made to the risk by *both* those at-risk and their partners. Partners were able to draw on other experiences, such as having a parent die from cancer, to understand the dilemmas that HD presented.

Analysis of how the younger generation's accounts revealed three distinct forms that negotiation could take. These were 'partnership', characterised by collaboration, 'facilitating', where one person made a decision and the other facilitated it, and 'deadlock', where it was difficult to find common ground. These patterns were similar to those identified by Beach (1996) in a corporate context. Successful negotiation created shared values that became a resource to be drawn on in times of stress, for example when facing terminations. This suggested that once shared values were established they became more powerful and took precedence over individual ones.

These patterns of negotiation extended to others who were involved in the decisionmaking process. This included lateral kin such as siblings and cousins, the previous generation from both the affected and non-affected families, work colleagues, friends and health care professionals and sometimes even laboratory staff. Conflicting ideas about appropriate responses could cause temporary breakdowns in relationships between family members. Case studies presented elsewhere documented how the health care professionals and laboratory staff can come to be seen as a substitute family when this happened (Downing, 2005).

Certain areas emerged as not negotiable. One was the decision whether to have a predictive test - which was seen as the prerogative of the at-risk partner. Another was whether to have a termination - which was seen as the woman's decision. However, couples might still *discuss* these decisions and partners decide whether to *facilitate* the choices made.

Information seeking formed another aspect of social decision-making. Constraints could arise according to how HD was managed within a family. If the family coped by not talking this could result in frustration and bitterness if it was not compatible with partners' needs to know. Tensions could also arise if partners had more factual knowledge than the family, especially if it contradicted the myths, incompleteness and misinformation that could exist about HD.<sup>6</sup> For some, this stage included making contact with clinical genetics services.

## Implementation

Passage through these stages could be rapid or slow. The retrospective design of the study made it possible to encompass this variation. For example, one couple waited twelve years for a new strategy - genetic testing - to become available to achieve their goal of a child free from risk. The limitations of studying personal decision-making through hypothetical situations were shown: what people actually did when in a situation could be quite different from what they thought they would do. Further difficulty could arise at this stage from having to confront additional choices, for example, around termination. Previous experience aided implementation of subsequent decisions by providing procedural expertise and feelings of increased confidence. However, it could prove to be disconcerting if new scenarios arose, such as the contrast between reproductive decision-making for first and subsequent children described above.

## **Post-Decision Evaluation**

The final stage of decision-making was post-decision evaluation which could be positive, negative or mixed. Individuals could be reticent about divulging post decision distress in order to protect partners. They could also be reluctant to share their feelings if they sensed that their partner was at a different stage to them in working through post-decision reactions. This was especially noticeable in the accounts given of post termination distress. These observations suggest that decision-making is most likely to be shared in the pre-implementation and implementation stages.

People subsequently justified having children at-risk in diverse ways. Many felt they have been deprived enough - citing lack of nurturing in childhood and repeated bereavements - and questioned why they should have to give up having children as well. Reproduction enabled them to do something normal, reassuring affected relatives that at-risk lives were not constrained, compensating those they love and bringing them a measure of hope. Children continued the generations, replacing other family members who had died. They became symbols of their parent's continuing health, providing a notion of immortality to counteract fears of premature death. One of the commonest fears expressed by those at-risk is that of being abandoned if they develop HD. Having children, provided potential carers - what better carer than someone who has seen it all and seen others coping with it? Having more than one child provided potential carers for the carers should they in turn become affected.

## **Progression Through Stages**

Although identification of these stages suggests a linear progression this was not in fact necessarily the case. For example, mental simulation could incorporate projected post-decision evaluation - in the form of anticipated decision regret - prior to implementation. Previous decisions could also feed into subsequent ones, or function as progress decisions to modify assessments of strategies or goals. Progression was aided by the development of

<sup>&</sup>lt;sup>6</sup> This was especially the case for those who had come across HD in a professional capacity.

competence as outlined by Klein (1999). Family case studies (Downing, 2005) showed this could encompass decision-making by other family members such as siblings.

## Evaluating the Retrospective Approach taken to Studying Decision-Making

There were clearly demonstrable advantages of taking a retrospective approach. As hoped, this provided the flexibility to encompass variations in time taken to make decisions, enabled contact to be made with people at various stages of decision-making, and encapsulated variations in the social contexts experienced by each generation, thus allowing considerable access to what Kemdal and Montgomery (1997: 75) have called the "full picture of all the factors that have been of importance for the decisions". Although the interviews gave a snapshot of one person's views and interpretations at that time, their stories revealed ways in which these had changed. Decisions could also be analyzed in the face of changes which occurred over time, such as moving from being at 25% risk to 50% risk after a parent was diagnosed, the introduction of genetic testing and subsequent altered risk status.

Another way of evaluating the retrospective approach taken was to consider whether the findings might be different, or what could be added, if the study had adopted a prospective design. In fact, some prospective data was generated in that participants were sent a short questionnaire approximately six months after being interviewed, which provided information about how taking part in the study had affected the way HD was discussed in families. Six months was possibly too short an interval to capture relevant developments in risk status, relationships and reproduction. If a longer interval had been scheduled would it have been possible to say something different linking HD status and reproductive decisions?

Some participants remained in contact following their interviews, providing some informal longer term prospective data, including results of predictive tests, which can be drawn on to reflect on this question. What these limited data suggested was that changes other than clarifying risk status influenced participants' responses. In particular, continued contact with participants revealed that decisions made when in one relationship do not necessarily predict those that will be made in the context of a different relationship. Two women who had been adamant that they would not have children had been prepared to do so after forming new relationships. Their behaviour was consistent with the theoretical model, in that it demonstrated how reproductive decision-making was mediated by perceived social support and shared values. These instances reinforced the claim made in this chapter of the need to consider couples, as well as individuals, as the unit of reproductive decision-making.

It is undeniable that retrospective data, such as that gathered in the interviews, is subject to biases of recall and omission. However, it became possible to see these gaps and biases as informative, in that they elucidated *how* people recalled events and structured them as stories in the evaluative stage of decision-making. Corroborations and contradictions between different accounts - from multiple family members - were especially useful in tracking down where these biases arose and in providing indications of what they might signify.

This can be illustrated by the following example, concerning a considerable disparity that occurred in accounts given by a sister and a brother about *when* they learnt of their parent's diagnosis. The sister placed it at a time before either of them had had children, the brother after they both had. Had this time gap actually occurred? Was the initial information not in fact passed on by the woman to her brother - which then tells us something about the time that such information may take to filter through families? Or, was it because her attempt to challenge the existing diagnosis of Parkinson's disease was met with extreme hostility from

other family members, and, a firm refusal to accept that it might be HD until much later when their remaining sibling was diagnosed with HD? Or, did it mean that she recalled the information because it was more salient to her than to her brother, as she was considering having children at that time, and he was not? It was likely that in this case the distress that she could envisage from not having children also helped to structure the vivid memory that she had of having to confront this diagnosis. Bruner (1986;1990), amongst others, has stressed how important emotion and affect are in structuring memories and how events become coloured by the emotional state people are in when they experience and recall them. This example reinforces the point made earlier in this chapter that 'knowing about' HD is not the event of a single moment, but is rather an unfolding awareness which was evidently happening at a different time or pace for these siblings.

Another interpretation of the contradictions in their accounts was that she could accept that she had knowingly had children at-risk whilst he could not. This last point suggests that people may have very good reasons for choosing to remember experiences in ways which may not reflect the actual sequence of events. Evidence existed within their narratives for *all* of these explanations and it was difficult to prioritize any one of them. However, by looking at a number of these instances it became possible to create a detailed picture of the complexity in which events occurred, were recalled and came to have significance for people. Similar conclusions are deduced by Song (1998: 115) in a study of siblings' experiences of working in family businesses. She too had found that "the result of combining and analysing two siblings' interviews was greater than the mere sum of the two interviews." This was not necessarily because of any additional information provided but because "multiple interviews helped reveal the complexities, contradictions, and tensions in people's accounts and in their daily lives."

## CONCLUSION

The specific findings presented above are of particular importance as reproductive decision-making in the face of genetic risk remains an under-researched area. Bekker et al's review cited in the introduction of 547 studies published between 1986 and 1996 categorised as being about informed decision-making yielded only seven that dealt with genetic factors, and none which addressed late onset disorders. Their description of reproductive decision-making as facing "relatively simple clearly-defined options with known risks" does not accord with actual accounts of how decision-making is experienced in the face of genetic risk. The grounded theory analysis revealed that considerable complexities exist in the way genetic risks are perceived which need to be identified and acknowledged in order to understand the multiplicity of responses that people show to them.

From an early stage in the analytic process concepts emerging from the data were able to be integrated with and extend existing NDM theoretical frameworks. Overall similarities in the ways in which NDM models convey the process of personal decision-making enabled their principles to be illustrated by one particular model, image theory (Beach, 1990). Comparisons made between the model of responsibility generated from data and image theory revealed shared features. For example, both models stress that the values that people hold play a crucial part in shaping their decision-making. An important contribution made by the

grounded theory analysis conducted as part of the study was to identify which values participants saw as salient in the context of reproductive decision-making and HD risks. It then became possible to account for varied responses shown to the child-centred and parenting risks by identifying which of these values decision-makers chose to prioritize, in combination with other contributing factors.

Another point of correspondence between image theory and the study model was the recognition of the importance of pre-choice processing, including coming to an understanding of the decision situation. Awareness was shown to be the prime causal condition in this process. Empirical support also was found for the claim forwarded in the introduction that decision-theoretic and health psychology form complementary rather than competing approaches. For example, a related finding that awareness included emotional and cognitive aspects is consistent with theoretical frameworks advanced in health psychology models such as the parallel response model (Leventhal et al, 1980, Leventhal et al, 1997). However, as McAllister (1999) also noted in relation to her model of engagement, the grounded theory analysis revealed that emotional and cognitive aspects interact dynamically rather than as parallel responses. Decision-making became shared in the way that Beach (1996) had suggested, in that individual decision-making preceded shared decision-making. Patterns of negotiation emerging from the data mapped onto those identified by Beach in other contexts. What the study additionally contributed was to reveal how negotiation can engender shared values that become a valuable resource facilitating implementing difficult decisions, such as proceeding with the termination of a wanted pregnancy. These correspondences suggest that NDM models have much to offer as frameworks in which to study decision-making in the face of genetic risk.

## Further Applications of the Model

The introduction alluded to the fact that how risk is negotiated is shaped by being lived in a particular moment in history. Deconstructing how parents viewed reproductive responsibility in relation to genetic inheritance had previously been seen by Callahan (1979) as entering uncharted territory. Opportunities at that time for genetic information were extremely limited and motivations for reproduction had not normally been subjected to detailed scrutiny. However, as Callahan pointed out it was possible to view this question as part of existing work exploring locus of control and intervention in natural processes. Employing this framework yielded contrasting constructions of what could be expected in relation to acting responsibly. If control was perceived as external then parental responsibility could only be reactive, characterised as coping with whatever fate presented. If control was perceived as internal, parental responsibility became proactive, raising expectations that people would employ interventions such as predictive testing in order to inform and control outcomes. These constructions map onto the choice or chance perspectives referred to earlier in this chapter.

Other features of Callahan's analysis accord with the present study's findings. Callahan (1979: 235) distinguished between "exercising responsibility *as* a parent and responsibility in *becoming* a parent", which corresponds to the two aspects of reproductive risk associated with HD. He suggested that individuals could become either more or less accepting of risk when they came into contact with others in their social context, in accordance with a wide range of

findings concerning the "risky shift" phenomenon which states that individuals modify their acceptance of risk in social contexts (Lamm and Myers, 1978). This is consistent with the finding from this study that responsibility involved both individual assessment and social evaluation.

Responsibility was shown to be both a private and a public concern. People had to convince themselves that their decision-making involved acting responsibly to the issues that HD raised for reproduction, and this provided the incentive for cognitive restructuring, or redefining, which could operate in various ways and lead to different outcomes. For example, a few participants redefined themselves as infertile, enabling them to utilize gamete donation in order to have a child and assure themselves that they had acted responsibly because they had bypassed the genetic risk. Redefining was consistent with NDM accounts depicting personal decision-making as making creative and adaptive responses to difficult dilemmas (Payne, Bettman and Johnson, 1993). Creativity involved reconciling discourses of genetic determinism and control within a framework of genetic connection in such a way as to present themselves as acting responsibly

This chapter has presented an account of how decision-making is experienced at a time when it has become possible to clarify uncertainty about risk for HD and utilize strategies to avoid passing this on to the next generation but when no effective treatments or cure yet exists for HD.

The model can be used to study uptake options such as preimplantation genetic diagnosis (PDG) that have subsequently become available. PGD enables people to have children who are known to be gene-negative and avoids ethical issues associated with terminating a pregnancy. Eggs are harvested, fertilized *in vitro* and tests performed on single cells biopsied at the eight-cell embryo (day 3 of development. Only those testing gene-negative are implanted. The main impediments to PGD are its expense, the low efficiency of *in vitro* fertilization (IVF), with only 20% to 30% of couples achieving pregnancy per IVF cycle (Pickering et al.) in the first 100 PDG cycles performed at the Guy's and St. Thomas' Centre in London. Some programs offer non disclosing PDG for couples who do not wish to undergo HD presymptomatic testing. The parental gene status is not revealed during the protocol and only gene negative embryos are implanted.

It is suggested that when effective treatments for HD become a reality concepts of responsibility will change again. Discourses of control could, as Rothman (1998) speculates, generate a more proactive form of decision-making:

'It's genetic' might very quickly not be a throwing-up-of-your-sleeves kind of problem. 'It's genetic' might be coming to mean, so let's fix it, let's engineer it, let's construct it to order. Let us make the determination, and let us predetermine (Rothman, 1998: 210-211).

The factors identified by Slovic (1987) can be used to predict, as shown in Table 4 below, how perceptions of genetic risk might change.

Questions arise about the extent to which it is possible to generalize from the findings from the study. Previous research shows the need for caution when attempting to generalise about responses to risks categorised as belonging to a particular domain. Variations that risks demonstrate on attributes other than risk category contribute to how they are perceived. Hypothetical scenarios constructed within the general category of rail accidents enabled researchers to conclude that responses varied according to specific attributes of trains, cargoes, accidents, locations, and causes of accidents (Kraus and Slovic, 1988). Similar findings were shown in connection with risk perception within other risk domains such as car faults (Slovic, MacGregor and Kraus, 1987), medicines (Slovic, Kraus et al, 1989), and unwanted environmental features (Slovic, 1992).

Table 4. Drawing on Slovic's (1987) Factors to Predict the Impact of Genetic Testing &
Treatments for HD on Perceptions of Genetic Risk

Factor	Genetic risk pre-testing	Genetic risk post-testing	
Dread risk	Uncontrollable	Controllable	
	Dread	Less dread	
	Involuntary	Voluntary	
Unknown risk	Not observable	Observable	
	Impact delayed	Impact predictable	
Number of people	High risk for future generations	Avoidable risk for future generations	

Attributes of genetic risk which seem likely to affect how they are perceived in relation to reproduction include mode of inheritance, perceived severity, possibilities for surveillance, treatment, cure and genetic testing, and age of onset. For example, varied modes of inheritance produce quite different scenarios of reproductive risk decision-making - some of which are briefly sketched out below.

Carriers of *recessive* disorders, such as cystic fibrosis, often remain unaware of the risk until *after* becoming parents. Frequently it is not until after the birth of an affected baby, or the development of symptoms in a previously healthy child, that they perceive their carrier status.<sup>7</sup> Parental anxiety arises from having to cope with the different and extra demands that an affected child creates. Both parents of children with recessive disorders are, and will remain, symptomless carriers. Carriers may also experience increased anxiety on learning of their status (Mennie et al, 1993; Marteau and Anionwu, 1996), which could influence subsequent reproductive decision-making. Reproductive dilemmas arise around whether or not to have *further* children who have a one in four chance of inheriting the disorder (Snowdon and Green, 1994, 1997). As was suggested at the beginning of this chapter, these scenarios differ from those associated with *late onset dominant* disorders, such as HD. A crucial difference is that those who inherit one mutated copy of a dominant gene are likely to *develop* the associated disorder. Thus, unlike carriers of recessive disorders, they face risks in connection with their own health or premature mortality. This can impact on their ability to sustain a mothering or fathering role.

<sup>&</sup>lt;sup>7</sup> This may or may not be their first child. It is only when a child inherits an *affected* copy of the gene from *each* parent that this occurs. Otherwise they are either unaffected or symptomless carriers.

*Gender* considerations arise in relation to another group of disorders - *X-linked* ones - which are passed on by females but mostly affect males.<sup>8</sup> The risk scenario that emerges for these conditions is that having males is perceived as more risky than having females (Parsons and Clarke, 1993; Kay, E. & Kingston H. (2002).

How people frame reproductive decision-making in relation to one late-onset disorder such as HD will not necessarily correspond to other late onset disorders. The likelihood of the parent becoming affected depends on the degree of penetrance shown by the particular mutation.9 Some, such as that associated with HD show almost 100% penetrance whilst others, such as the BRAC1 gene mutation, associated with some inherited breast and ovarian cancers, and that associated with early onset Alzheimer's disease, show a much lower rate. Clarifying risks for these diseases is further complicated in that they also arise sporadically, creating a general population risk for those with no family history. This means that those shown not to carry genes for familial breast, ovarian and bowel cancers can still develop these disorders but those found to carry the genes may not. Other differences exist between these cancers and HD. As mentioned above, those who will develop HD cannot yet benefit from the surveillance measures or treatments which exist for cancer sufferers, and they are likely to develop psychiatric as well as physical symptoms. The first comparative study of the psychological effects of predictive testing of neurodegenerative and cancer syndrome lateonset autosomal dominantly inherited genetic disorders lend support to this in that the groups at-risk for neurodegenerative disorders showed more distress than those at-risk for the cancer syndromes (Dudok de Wit, 1997).<sup>10</sup> Families with neurodegenerative disorders were also more focused upon support for the individual at-risk and partner whilst in families with cancer the focus was upon the post-test treatment options and their outcome (Dudok de Wit et al, 2000).

It may well be that findings in relation to decision-making and HD could have as much, if not more, to contribute to debates about reproductive decision-making in non-genetic situations that have aroused public concern about ability to sustain a parenting role. These include mothers and fathers with learning difficulties (Booth and Booth, 1994), parents who are HIV positive (Bor and Elford, 1998), teenage single mothers (Rains, Davies and McKinnon, 1998) parents who are lesbian and gay (Dunne, 2000), who face deafness (Gregory, 1991), and the experiences of women wishing to become impregnated under unconventional circumstances such as being post-menopausal or wanting to utilize sperm from deceased partners.

## ACKNOWLEDGMENTS

Thanks are due to the participating families for entrusting me with their thoughts and experiences some of which are touched on here; to the Medical Research Council for funding

<sup>&</sup>lt;sup>8</sup> Fragile X primarily affects males but one third of carrier females show some degree of intellectual impairment (Davies, 1989).

<sup>&</sup>lt;sup>9</sup> Penetrance refers to the likelihood of the disorder manifesting.

<sup>&</sup>lt;sup>10</sup> Risk perception was also a function of experience of the particular disorder and the perceived impact it had had on the testee which meant that individual differences in psychological distress between testees at-risk for a particular disorder were larger than the similarities.

the study under their GAHH initiative and to my supervisors, Jo Green and Martin Richards, for their encouragement and guidance. The Huntington's Disease Association and Addenbrookes Regional Clinical Genetics Centre shared their expertise with me and helped recruit participants to the study. I am also grateful to Martin Richards, Nina Hallowell, Helen Statham, Robert Osborne and Helena Willen for their helpful comments on subsequent drafts of this chapter. I am currently funded by The Wellcome Trust (Award Reference: 065207, 2002-2005) under their biomedical ethics programme.

## REFERENCES

- Adam, S., Wiggins, S., Whyte, P., Bloch, M., Shokeir, M. H. K., Soltan, H., et al.. (1993) Five year study of prenatal testing for Huntington's disease: demand, attitudes, and psychological assessment. *Journal of Medical Genetics*, 30, 549-56.
- Arkes, H. R. & Blumer, C. (1985) The psychology of sunk costs. Organizational Behaviour and Human Decision Processes, 35, 124-40.
- Barette, J. & Marsden, C. D. (1979) Attitudes of families to some aspects of Huntington's chorea. *Psychological Medicine*, 9, 327-36.
- Bartlett, J. (1994) Will you be mother? Women who choose to say no. London: Virago.
- Beach, L. R. (1990) *Image theory: Decision making in personal and organisational contexts.* Chichester: Wiley.
- Beach, L. R. (1993a) Making the right decision: organizational culture, vision, and planning. Englewood Cliffs, NJ: Prentice Hall.
- Beach, L. R. (1993b) Image theory: Personal and organizational decisions. In G. A. Klein, J. Orasanu, R. Calderwood, C. E. Zsambok (eds.), *Decision making in action: models and methods*, pp.148-57. Norwood, NJ: Ablex.
- Beach, L. R. (1996) *Decision making in the workplace: a unified perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Beach, L. R., Townes, B. D., Campbell, F. L. & Keating, G. W. (1976) Developing and testing a decision aid for birth planning decisions. *Organizational Behaviour and Human Performance*, 15, 99-116.
- Beach, L. R. & Lipshitz, R. (1993) Why classical decision theory is an inappropriate standard for evaluating and aiding most human decision making. In G. A. Klein, J. Orasanu, R. Calderwood & C.E. Zsambok (eds.), *Decision making in action: models and methods*, pp.21-35. Norwood, NJ: Ablex Publishing Corporation.
- Beattie, J., Baron, J., Hershey, J. C., & Spranca, M. D. (1994) Psychological determinants of decision attitude. *Journal of Behavioural Decision Making*, 7, 129-44.
- Beck, U. & Beck-Gernsheim, E. (1995) The normal chaos of love. Cambridge: Polity Press.
- Becker, M. H. (1974) The Health Belief Model and sick role behaviour. *Health Education Monographs*, 2, 409-19.
- Bekker, H., Thornton, J., Airey, M., Connelly, J., Hewison, J., Lilleyman, et al (1999) Informed decision making: an annotated bibliography and systematic review. UK: Health Technology Assessment, No. 3, National Health Service, Research & Development.
- Bloch, M. & Hayden, M. R. (1990) Opinion: predictive testing for Huntington disease in childhood: challenges and implications. *American Journal of Human Genetics*, 46, 1-4.

- Booth, T. & Booth, W. (1994) *Parenting under pressure: mothers and fathers with learning difficulties.* Buckingham: Open University Press.
- Bor, R. & Elford, J. (1998) *The family and HIV today: recent research and practice*. London: Cassell.
- Bostrom, A., Fischhoff, B. & Granger Morgan, M. (1992) Characterizing mental models of hazardous processes: a methodology and application to radon. *Journal of Social Issues*, 48, 85-100.
- Boulton, M. & Fitzpatrick, R. (1997) Evaluating qualitative research. Evidence Based Health Policy and Management, 1, 83-5.
- Broadstock, M. & Michie, S. (2000) Processes of decision making: theoretical and methodological issues. *Psychology and Health*, 15, 191-204.
- Bruner, J. (1986) Actual minds, possible worlds. Cambridge, MA: Harvard University Press.
- Bruner, J. (1990) Acts of meaning. Cambridge, MA: Harvard University Press.
- Busfield, J. & Paddon, M. (1977) *Thinking about children: sociology and fertility in post-war England.* Cambridge: Cambridge University Press.
- Callahan, S. (1979) An ethical analysis of responsible parenthood. In A. M. Capron, M. Lappe, & R. F. Murray (eds.), *Genetic counseling: facts, values and norms. Birth Defects: Original Articles Series*, XV, pp.217-38. New York: Alan R. Liss Inc.
- Campbell, E. (1985) *The childless marriage: an exploratory study of couples who do not want children.* London: Tavistock Publications.
- Cannon-Bowers, J. A., Salas, E. & Pruitt, J. S. (1996) Establishing the boundaries of a paradigm for decision-making research. *Human Factors*, 38, 193-205.
- Carter, C. O. & Evans, K. (1979) Counselling and Huntington's chorea. Lancet, ii, 470-2.
- Conrad, P. & Gabe, J. (1999) (eds) *Sociological Perspectives on the New Genetics*. Oxford & Malden, Ma.: Blackwell Publishers Ltd.
- Cox, S.M. (Ed.) (2002). *Personal Perspectives on Genetic Testing for Huntington Disease*. Cambridge, Ontario: Huntington Society of Canada.
- Cox, S. M. (2003). Stories in Decisions: How At Risk Individuals Decide to Request Predictive Testing for Huntington Disease. *Qualitative Sociology*, 26, 257-280.
- Creswell, J. W. (1998) *Qualitative inquiry and research design: choosing among five traditions*. Thousand Oaks: Sage Publications Inc.
- Currie, D. (1988) Re-thinking what we do and how we do it: a study of reproductive decision making. *Review of Canadian Sociology and Anthropology*, 25, 231-53.
- Dallos, R. (1997) Interacting stories: narratives, family beliefs and therapy. London: Karnac.
- Davenport, C. B. & Muncey, E. B. (1916) Huntington's Chorea in relation to heredity and eugenics. *Eugenics Record Office*, Bulletin No.7, 195-220.
- Decruyenaere, M., Evers-Kiebooms, G., Boogaerts, A., Cassiman, J. J., Cloostermans, T., Demyttenaere, K., et al (1996) Prediction of psychological functioning one year after the predictive test for Huntington's disease and impact of the test result on reproductive decision making. *Journal of Medical Genetics*, 33, 737-43.
- Douglas, M. (1985) *Risk acceptability according to the social sciences*. New York: Russell Sage Foundation.
- Downing, C. (2002). Reproductive decision-making in families at risk for Huntington's disease: perceptions of responsibility. An unpublished dissertation submitted to the University of Cambridge for the Degree of Doctor of Philosophy

Downing, C. (2005) Negotiating responsibility: case studies of reproductive decision-making and prenatal genetic testing in families facing Huntington disease. Journal of Genetic Counseling, Vol. 14, 3, pp. 219-234.

Dowrick, S. & Grundberg, S. (1980) Why children? London: The Women's Press.

- Drake, H., Reid, M. & Marteau, T. (1996) Attitudes towards termination for fetal abnormality: comparisons in three European countries. *Clinical Genetics*, 49, 134-40.
- Dudok de Wit, A. C. (1997) To know or not to know: the psychological implications of presymptomatic DNA testing for autosomal dominant inheritable late onset genetic disorders. PhD thesis, Rotterdam, Erasmus University of Rotterdam.
- Dunne, G. A. (2000) Opting into motherhood: lesbians blurring the boundaries and transforming the meaning of parenthood and kinship. *The Journal of Gender and Society*, 14, 11-35.

Erikson, E. H. (1963) Childhood and society. (2nd ed.). New York: Norton.

- Evers-Kiebooms, G., Swerts, A. & van den Berghe, H. (1990) Partners of Huntington patients: implications of the disease and opinions about predictive testing and prenatal diagnosis. *Genetic Counseling*, 39, 151-9.
- Evers-Kiebooms, G., Denayer, L., Decruyenaere, M. &Van den Berghe, H. (1993) Community attitudes towards prenatal testing for congenital handicap. *Journal of Reproductive and Infant Psychology*, 11, 21-30.

Festinger, L. (1957) A theory of cognitive dissonance, Evanston, IL: Row Peterson.

- Finch, J. and Mason, J. (1993) Negotiating family responsibilities. London: Tavistock Routledge.
- Fletcher, J. F. (1980) Knowledge, risk, and the right to reproduce: a limiting principle. In A. Milunsky & G.J. Annas (eds.), *Genetics and the law II*, pp.131-5. New York: Plenum Press.
- Fox, G.L. (1982) *The childbearing decision: fertility attitudes and behavior*. Beverley Hills: Sage Publications Inc.
- Gifford, S. M. (1986) The meaning of lumps: a case study of the ambiguities of risk. In C. R. Janes, R. Stall & S. M. Gifford (eds.), *Anthropology and Epidemiology: interdisciplinary approaches to the study of health and disease*, pp.213-46. Boston: Reidel.
- Gladwin, C. H. (1989) *Ethnographic decision tree modeling*. Newbury Park: Sage Publications.
- Glaser, B. G. & Strauss, A. L. (1966) Awareness of dying. London: Weidenfeld and Nicolson.
- Green, J. (1990) Calming or harming? A critical review of psychological effects of fetal diagnosis on pregnant women. London: The Galton Institute.
- Green, J. M., Statham, H. & Snowdon, C. (1992) Screening for fetal abnormalities: attitudes and experiences. In T. Chard & M. P. M. Richards (eds.), *Obstetrics in the 1990's: current controversies*, pp. 65-89. London: McKeith Press.
- Green, J. M., Statham, H. & Snowdon, C. (1994) *Pregnancy: A testing time*. Report of the Cambridge Prenatal Screening Study. Centre for Family Research, University of Cambridge.
- Gregory, S (1991) Challenging motherhood: mothers and their deaf children. In A. Phoenix,A. Woollett, & E. Lloyd (eds.), *Motherhood: meanings, practices and ideologies*, pp.123-42. London: Sage Publications Ltd.
- Hallowell, N. (1999) Doing the right thing: genetic risk and responsibility. *Sociology of Health and Illness*, 21, 597-621.

- Hamburg, K., Johansson, E., Lindgren, G. & Westman, G. (1994) Scientific rigour in qualitative research - examples from a study of women's health in family practice. *Family Practice*, 11, 176-81.
- Hans, M. B & Koeppen, A. H. (1980) Huntington's chorea: its impact on the spouse. *The Journal of Mental and Nervous Disease*, 168, 209-14.
- Harper, P. (1996a) The epidemiology of Huntington's disease. In P. S. Harper (ed.), *Huntington's disease.* (2nd edn.) pp.201-39. London: W.B. Saunders Company Ltd.
- Harper, P. (1996b) The natural history of Huntington's disease. In P. S. Harper (ed.), *Huntington's disease*. 2nd edn. pp.123-136. London: W.B. Saunders Company Ltd.
- Harper, P. S., Walker, D. A., Tyler, A., Newcombe, R. O. & Davies, K. (1979) Huntington's chorea: the basis for long term prevention. *Lancet*, 8138, 346-9.
- Harper. P.S., Lim, C. & Craufurd, D. (2000) Ten years of presymptomatic testing for Huntington's disease: the experience of the UK Huntington's Disease Prediction Consortium. *Journal of Medical Genetics*, 37, 567-71.
- Hayden, M. (2000) Predictive testing for Huntington's disease: the calm after the storm. *Lancet*, 356, 1944-5.
- Hayes, C. V. (1992) Genetic testing for Huntington's disease a family issue. The New England Journal of Medicine, 327, 1449-51.
- Heider, F. (1946) Attitudes and cognitive organization. Journal of Psychology, 21, 107-12.
- Heider, F. (1958) The psychology of interpersonal relations. New York: John Wiley.
- Hill, S. A. (1994) Motherhood and the obfuscation of medical knowledge: the case of sickle cell disease. *Gender and Society*, 8, 29-47.
- Huys, J., Evers-Kiebooms, G. & d'Ydewalle, G. (1992) Decision making in the context of genetic risk: the use of scenarios. *Birth Defects:Original Article Series*, 28, 17-20.
- Johnson, B. T. & Eagly, A. H. (1989) Effects of involvement on persuasion: a meta-analysis. *Psychological Bulletin*, 106, 290-314.
- Jungermann, H. (1997) Advice giving and taking. Paper presented at the 16th bi-annual Conference on Subjective Probability Utility and Decision Making (SPUDM), Leeds.
- Kahneman, D. (1991) Judgment and decision making: a personal view. *Psychological Science*, 2, 142-45.
- Kahneman, D. & Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica*, 49, 263-91.
- Kahneman, D., Slovic, P., & Tversky, A. (1982) Judgement under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press.
- Kahneman, D. & Tversky, A. (1984) Choices, values and frames. American Psychologist, 39, 341-50.
- Karlsson, G. (1987) A phenomenological psychological method: theoretical foundations and empirical applications in the field of decision making and choice. PhD thesis, Stockholm, University of Stockholm.
- Kates, R. W. & Kasperson, J. X. (1983) Comparative risk analysis of technological hazards. *Proceedings of the National Academy of Science, U.S.A.*, 80, 7027-38.
- Kay, E. & Kingston H. (2002) Feelings associated with being a carrier and characteristics of reproductive decision making in women known to be carriers of X-linked conditions. *Journal of Health Psychology*, 7 169-81.
- Kelly, G. A. (1955) The psychology of personal constructs. New York: Norton.

- Kemdal, A. B. & Montgomery, H. (1997) Perspectives and emotions in personal decision making. In R. Ranyard, W. R. Crozier and O. Svenson (eds.), *Decision making: cognitive models and explanations*, pp.72-89. London: Routledge.
- Kenen, R. (1994) The Human Genome Project: creator of the potentially sick, potentially vulnerable and potentially stigmatized? In L. Robinson (ed.), *Life and death under high technology medicine*, pp.49-64. Manchester: Manchester University Press.
- Kenen, R. H. (1996) The at-risk health status and technology: a diagnostic invitation and the 'gift' of knowing. *Social Science and Medicine*, 42, 1545-53.
- Kessler, S. (1988) Invited essay on the psychological aspects of genetic counseling: V. Preselection: a family coping strategy in Huntington disease. *American Journal of Medical Genetics*, 31, 617-21.
- Kessler, S. (1989) Psychological aspects of genetic counseling. VI. A critical review of the literature dealing with education and reproduction. *American Journal of Medical Genetics*, 34, 340-53.
- Kiernan, K., E. (1989) Who remains childless? Journal of Biosocial Science, 21, 387-98.
- Kitcher, P. (1996) *The lives to come: the genetic revolution and human possibilities*. London: Penguin Books.
- Klein, G. (1997) Naturalistic decision making: where are we going? In C. E. Zsambok & G. Klein (eds.), *Naturalistic decision making*, pp. 383-97. Mahwah, NJ: Lawrence Erlbaum Associates.
- Klein, G. (1999) *Sources of power: how people make decisions*. Cambridge, MA: The M.I.T. Press.
- Klein, G. A., Orasanu, J., Calderwood, R. & Zsambok, C. E. (eds.) (1993) *Decision making in action: models and methods*. Norwood, NJ: Ablex Publishing Corporation.
- Korer, J. and Fitzsimmons, J. S. (1985) The effect of Huntington's chorea on family life. *British Journal of Social Work*, 15, 581-97.
- Kraus, N. N. and Slovic, P. (1988) Taxonomic analysis of perceived risk: modelling individual and group perceptions within homogenous hazard domains. *Risk Analysis*, 8, 435-55.
- Krimsky, S. and Golding, D. (eds.) (1992) Social theories of risk. Westport, CT: Praeger.
- Lamm, H. and Myers, D. G. (1978) Group induced polarization of attitudes and behavior. In L. Berkovitz (ed.), *Advances in experimental social psychology*, 11, pp.145-95. New York: Academic Press.
- Lee, R. M. (1993) Doing research on sensitive topics. London: Sage Publications Ltd.
- Leventhal, H., Meyer, D., and Nerenz, D. (1980) The common-sense representation of illness danger. *Medical Psychology*, 2, 7-30.
- Leventhal, H., Nerenz, D. R. & Steele, D. J. (1994) Illness representations and coping with health threats. In A. Baum, S. E. Taylor & J. E. Singer (eds.), *Handbook of psychology* and health, Vol. IV: Social aspects of health. pp. 210-52. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Leventhal, H., Benyamini, Y., Brownlee, S., et al (1997) Illness representations. In K.J. Petrie & J.A. Weinman, *Perceptions of health and illness*, pp. 19-45. Amsterdam: Harwood Academic Publishers.
- Lippman, A. (1992) Mother matters: a fresh look at prenatal genetic testing. *Issues in Reproductive and Genetic Engineering*, 5, 141-54.

- Lipshitz, R. (1993) Converging themes in the study of decision making in realistic settings. In G. A. Klein, J. Orasanu, R. Calderwood & C. Zsambok (eds.), *Decision making in action: models and methods*. pp.103-37. Norwood, NJ: Ablex Publishing Corporation..
- Lopes, L. L. (1987) Between hope and fear: the psychology of risk. Advances in Experimental Social Psychology, 20, 255-95.
- Maat-Kievit, A., Vegter-van der Vlis, M., Zoeteweij, M., Losekoot, M., van Haeringen, A., Kanhai, H. et al. (1999) Experience in prenatal testing for Huntington's disease in the The Netherlands: Procedures, results and guidelines (1987-1997). *Prenatal Diagnosis*, 19, 450-57.
- Mansfield, P. & Collard, J. (1988) *The beginning of the rest of your life: a portrait of newly-wed marriage*. London: The Macmillan Press Ltd.
- Marshall, S. E. (1995) Choosing the family. In C. Ulanowsky (ed.), *The family in the age of biotechnology*, pp.105-117. Aldershot: Avebury.
- Marteau, T. M., Drake, H. & Bobrow, M. (1994) Counselling after diagnosis of fetal abnormality: the differing approaches of obstetricians, clinical geneticists and genetic nurses, *Journal of Medical Genetics*, 31, 865-67.
- Marteau, T. & Anionwu, E. (1996) Evaluating carrier testing: objectives and outcomes. In T.Marteau and M. Richards, (eds.), *The troubled helix: social and psychological implications of the new human genetics*, pp.123-29. Cambridge: Cambridge University Press.
- Mason, M. C. (1993) Male infertility: men talking. London: Routledge.
- Mastromauro, C. A., Myers, R. H. & Berkman, B. (1987) Attitudes towards presymptomatic testing in Huntington's disease. *American Journal of Medical Genetics*, 26, 271-82.
- Mays, N. & Pope, C. (1995) Rigour and qualitative research. *British Medical Journal*, 311, 109-12.
- McAllister, F. and Clarke, L. (1996) *Choosing childlessness*. London: Family Policy Studies Centre.
- McAllister, M. (1999) Predictive testing for Hereditary Non-Polyposis Colorectal Cancer (HNPCC): a theory of engagement. Unpublished PhD thesis. Centre for Family Research, University of Cambridge, Cambridge.
- Mennie, M. E., Compton, M. E., Gilfillan, A., Liston, W. A., Pullen, I., Whyte, et al (1993) Prenatal screening for cystic fibrosis: psychological effects on carriers and their partners. *Journal of Medical Genetics*, 30, 543-48.
- Miller, S. M. (1987) Monitoring and blunting: validation of a questionnaire to assess styles of information seeking under threat. *Journal of Personality and Social Psychology*, 52, 345-53.
- Montgomery, H. (1993) The search for a dominance structure in decision making: examining the evidence. In G. A. Klein, J. Orasanu, R. Calderwood, C. E. Zsambok (eds.), *Decision making in action: models and methods*, pp.182-87. Norwood, NJ: Ablex Publishing Corporation.
- Montgomery, H. (1999) NDM: the challenge to JDM. Book review of *Naturalistic Decision Making* (1997), C. E. Zsambok & G. Klein (eds.). Mahwah, NJ: Lawrence Erlbaum Associates. In *Journal of Behavioural Decision Making*, 12, 340-42.
- Montgomery, H. & Svenson, O. (eds.) (1989) Process and structure in human decision making. Chichester, NY: John Wiley & Sons Ltd.

- Morell, C. M. (1994) Unwomanly conduct: the challenges of intentional childlessness. New York: Routledge.
- Nelson, H. L. and Nelson, J. L. (1995) *The patient in the family: an ethics of medicine and family*. London: Routledge.
- Orasanu, J. & Connolly, T. (1993) The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood, C. E. Zsambok (eds.), *Decision making in action: models and methods*, pp. 3-20. Norwood, NJ: Ablex Publishing Corporation.
- Palmlund, I. (1992) Social drama and risk evaluation. In S. Krimsky and D. Golding (eds.), Social theories of risk, pp.197-212. Westport, CT: Praeger.
- Parsons, E. & Clarke, A. J. (1993) Genetic risk: women's understanding of carrier risks in Duchenne muscular dystrophy. *Journal of Medical Genetics*, 30, 562-66.
- Payne, J. (1978) Talking about children: an examination of accounts about reproduction and family life. *Journal of Biosocial Science*, 10, 367-74.
- Payne, J. W., Bettman, J. R. & Johnson, E. J. (1993) The adaptive decision maker. Cambridge: Cambridge University Press.
- Pfeffer, N. & Woollett, A. (1983) The experience of infertility. London: Virago.
- Proctor, H. (1981) Family construct psychology. In S. Walrond-Skinner (ed.), *Family therapy and approaches*. London: Routledge & Kegan Paul.
- Quaid, K. & Wesson, M. K, (1995) Exploration of the effects of predictive testing for Huntington disease on intimate relationships. *American Journal of Medical Genetics*, 45, 46-51
- Rains, P., Davies, L. & McKinnon, M. (1998) Taking responsibility: an insider view of teen motherhood. *Families in Society: The Journal of Contemporary Human Services*, 79, 308-19.
- Read, J. (1995) Counselling for fertility problems. London: Sage Publications Ltd.
- Reid, M. & Armstrong, D. (1998) Guidelines on evaluating qualitative research proposals in health services research. Unpublished guidelines prepared for the Cancer Research Council, Educational and Psychosocial Research Committee, London.
- Richards, L. (1978) *Having families: marriage, parenthood and social pressure in Australia.* Harmondsworth: Penguin Books Ltd.
- Richards, M.P.M. (1993) The new genetics: some issues for social scientists. Sociology of *Health and Illness*, 15, 567-86.
- Richards, M. P. M. & Green, J. M. (1993) Attitudes toward prenatal screening for fetal abnormality and detection of carriers of genetic disease: a discussion paper. *Journal of Reproductive and Infant Psychology*, 11, 49-56.
- Richards, T. J. & Richards, L. (1994) Using computers in qualitative research. In N.K. Denzin and Y.S. Lincoln (eds) *Handbook of qualitative research*, pp. 445-62. Thousand Oaks: Sage Publications.
- Robertson, A. (2000). Embodying risk, embodying political responsibility: women's accounts of risks for breast cancer. *Health, Risk and Society*, 2, 219-35.
- Robinson, G. E., Carr, M. L., Olmsted, M. P., & Wright, C. (1991) Psychological reactions to pregnancy loss after prenatal diagnostic testing: preliminary results. *Journal of Psychosomatic Obstetrics and Gynecology* 12, 181-92.
- Rogers, W. (1994) Changing health-related attitudes and behaviour: the role of preventative health psychology. In J. J. Harvey, E. Maddox, R. F. McGlynn, R. F. & C. D. Stoltenberg

(eds.), *Social perception in clinical and counseling psychology*, Vol. 2, pp, 91-112. Lubbock, TX: Texas Technical University Press.

- Roos, R. A., Vegter van der Vlis, M., Tibben, A., Skraastad, M. I. & Pearson, P.L. (1990) [Procedure and initial results of presymptomatic DNA studies in Huntington's Chorea]. *Ned. Tijdschr. Geneeskd.*, 134, 704-7.
- Rothman, B. K. (1986) The tentative pregnancy: prenatal diagnosis and the future of *motherhood*. New York: Viking.
- Rothman, B. K. (1998) Genetic maps and human imaginations. The limits of science in understanding who we are. New York: W. W. Norton & Company.
- Rotter, J. B. (1966) Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 609, (whole).
- Sandelowski, M. & Jones, L. C. (1996) 'Healing fictions': stories of choosing in the aftermath of the detection of fetal abnormalities. *Social Science and Medicine*, 42, 353-61.
- Seale, C. (1999) The Quality of Qualitative Research. London: Sage Publications Ltd.
- Shakespeare, J. (1992) Communication in Huntington's disease families. Unpublished paper presented at the Third European Meeting on Psychosocial Aspects of Genetics, Nottingham, University of Nottingham.
- Shiloh, S. (1996) Decision-making in the context of genetic risk. In T. Marteau & M. Richards (eds.), *The troubled helix: social and psychological implications of the new human genetics*, pp.82-103. Cambridge: Cambridge University Press.
- Simon, H. A. (1955) Behavioural Model of Rational Choice. Quarterly Journal of Economics, 69, 99-108.
- Simpson, S. A. & Harding, A. E. (1993) Predictive testing for Huntington's disease: after the gene. *Journal of Medical Genetics*, 30, 1036-38.
- Simpson, S.A. and Harper, P.S. (2001) Prenatal testing for Huntington's disease: experience within the UK 1994-1998. Letter to the Editor. *Journal of Medical Genetics*, 38, 333-35.
- Sloan, T. S. (1986) Deciding: Self-deception in life choices. New York: Methuen.
- Slovic, P. (1987) Perception of risk. Science, 236, 280-85.
- Slovic, P. (1992) Perception of risk: reflections on the psychometric paradigm. In S. Krimsky & D. Golding (eds.), *Social theories of risk*, pp.117-52. Westport, CT: Praeger.
- Slovic, P., Lichstenstein, S. & Fischhoff, B. (1984) Modeling the societal impact of fatal accidents. *Management Science*, 30, 464-74.
- Slovic, P., MacGregor, D. & Kraus, N. N. (1987) Perception of risk from automobile safety defects. Accident Analysis and Prevention, 19, 359-73.
- Slovic, P., Kraus, N. N., Lappe, H., Letzel, H. & Malmfors, T. (1989) Risk perception of prescription drugs: report on a survey in Sweden. *Pharmaceutical Medicine*, 4, 43-65.
- Smith, J.A., Stephenson, M. & Quarrel, O. (1999). Factors influencing the decision whether to take the genetic test for Huntington's disease: an interpretative phenomenological analysis. Unpublished paper presented at the Qualitative Methods and Genetics Group, Centre for Family Research, University of Cambridge.
- Sobel, S. K. & Cowan, D. B. (2000) Impact of genetic testing for Huntington's disease on the family system. *American Journal of Medical Genetics*, 90, 49-59.
- Song, M. (1998) Hearing competing voices: sibling research. In J. Ribbens and R. Edwards (eds.), *Feminist dilemmas in qualitative research*, pp.103-18. London: Sage Publications Ltd.

- Strathern, M. (1995) New families for old? In C. E. Ulanowsky (ed.), *The family in the age of biotechnology*, pp.27-45. Aldershot: Avebury.
- Strauss, A. and Corbin, J. (1998) *Basics of qualitative research: techniques and procedures* for developing grounded theory .(2nd edn.) Thousand Oaks: Sage Publications Inc.
- Svenson, O. (1991) Differentiation and consolidation theory of human decision making. Paper presented at the 13th SPUDM Conference, Fribourg, Switzerland.
- Svenson, O. (1996) Decision making and the search for fundamental psychological regularities: what can be learnt from a process perspective? Organizational Behaviour and Human Decision Processes, 65, 252-67.
- Svenson, O. and Maule, A. J. (eds.) (1993) Time pressure and stress in human judgement and decision making. New York: Plenum Press.
- Taylor, S.D. (2004). Predictive genetic test decisions for Huntington's Disease: context, appraisal and new moral imperatives. *Social Science and Medicine*, 58, 137-149.
- Teigen and Brun (1997) Anticipating the future. In R. Ranyard, W. R. Crozier & O. Svenson (eds.), *Decision making: cognitive models and explanations*, pp.112-127. London: Routledge.
- Thom, D. and Jennings, M. (1996) Human pedigree and the 'best stock': from eugenics to genetics. In T. Marteau & M. Richards (eds.), *The troubled helix: social and psychological implications of the new human genetics*, pp.211-234. Cambridge: Cambridge University Press.
- Tibben, A., Frets, P. G., van de Kamp, J., J. P., Niermeijer, M. F., Vegter-van de Vlis, M., Roos, R. A. C., et al. (1993b) On attitudes and appreciation, 6 months after predictive DNA testing for Huntington disease in the Dutch program. *American Journal of Medical Genetics*, 48, 103-11.
- Tolmie, J. L., Davidson, H. R., May, H. M., McIntosh, K., Paterson, J. S. & Smith, B. (1995) The prenatal exclusion test for Huntington's disease: experience in the west of Scotland. *Journal of Medical Genetics*, 32, 97-101.
- Tversky, A., & Kahneman, D. (1973) Availability: a heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-32.
- Tversky, A. and Shafir, E. (1992) Choice under conflict: the dynamics of deferred decision. *Psychological Science*, 3, 358-61.
- Tymstra, T. J. (1989) The imperative character of medical technology and the meaning of "anticipated decision regret". *International Journal of Technology Assessment in Health Care*, 5, 207-13.
- Van der Steenstraten, I.M., Tibben, A., Roos, R.A.C., van de Kamp, J.J.P. and Niermeijer, M.R. (1994) Predictive testing for Huntington disease: nonparticipants compared with participants in the Dutch program. *American Journal of Human Genetics*, 55, 618-25.
- Verplanken, B., & Svenson, O. (1997) Personal involvement in human decision making: conceptualisations and effects on decision processes. In R. Ranyard, W. R. Crozier & O. Svenson (eds.), *Decision making:Cognitive models and explanations*, pp.40-57. London: Routledge.
- Wallston, K. A. (1992) Hocus-pocus, the focus isn't strictly on locus: Rotter's social learning theory modified for health. *Cognitive Therapy and Research*, 16, 183-99.
- Wallston, K. A. (1997) Perceived control and health behaviour. In A. Baum, S. Newman, J. Weinman, R. West & C. McManus (eds.), *Cambridge Handbook of psychology, health and medicine*, pp.151-54. Cambridge: Cambridge University Press.

- Wallston, K. A., Wallston, B. S. & DeVellis, R. (1978) Development of the Multidimensional Health Locus of Control (MHLC) scale. *Health Education Monographs*, 6, 161-70.
- Wartofsky, M. W. (1982) Medical knowledge as a social product: rights, risks, and responsibilities. In W. B. Bondeson, H. T. Englehardt, S. F. Spicker, & J. M.White (eds.), *New knowledge in the biomedical sciences: some moral implications of its acquisition, possession and use,* pp.113-30. Dordrecht: D. Reidel Publishing Company.
- Wertz, D. C., Sorenson, J. R., and Heeren, T. C. (1984) Genetic counseling and reproductive uncertainty. *American Journal of Medical Genetics*, 18, 79-88.
- Wexler, N. S. (1979) Russian roulette: the experience of being at risk for Huntington's disease: harbinger of the new genetics. In S. Kessler (ed.), *Genetic counseling: psychological dimensions*, pp. 199-220. New York: Academic Press.
- Wexler, N. (1980) "Will the circle be unbroken?" Sterilizing the genetically impaired. In A. Milunsky & G. J. Annes (eds.), *Genetics and the law II*, pp.313-329. New York: Plenum Press.
- Wexler, N. S. (1992) Clairvoyance and caution: repercussions from the Human Genome Project. In D. J. Kevles & L. Hood (eds.), *The code of codes: scientific and social issues in the Human Genome Project*, pp. 211-43. Cambridge, MA: Harvard University Press.
- Willen, H. (1994) The childbearing decision: motivational aspects, process characteristics and consequences of the decision. Ph.D.thesis, Department of Psychology, Goteborg University. *Goteborg Psychological Reports*, 24, No.1.
- Willen, H. & Montgomery, H. (2000) Creating a realisable option on the ruins of a wrecked marriage. In C. M. Allwood & M. Selart (eds.), *Decision-making: social and creative dimensions*. Dordrecht: Kluwer Academic Publishers.
- Yow, V. R. (1994) *Recording oral history: a practical guide for social scientists.* Thousand Oaks: Sage Publications Inc.
- Zsambok, C. E. and Klein, G. (eds.) (1997) *Naturalistic decision making*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chapter 16

## DECISION MAKING PROCESS AND HEALTH SEEKING PATTERNS OF YOUNG WOMEN WITH UNPLANNED PREGNANCIES, BANGKOK, THAILAND\*

Wanapa Naravage<sup>1</sup> and Rungpetch C. Sakulbumrungsil<sup>2</sup>

<sup>1</sup>Program Manager Contraceptive Technologies, Concept Foundation, Thailand <sup>2</sup>Faculty of Pharmaceutical Sciences, Chulalongkorn University, Thailand

## ABSTRACT

*Objectives*: This study examined the decision making process and health seeking patterns of low-income young women with unplanned pregnancies who opted for abortion, putting baby up for adoption, or keeping the baby.

*Methods*: Research methodology in this study focused on formative research utilizing qualitative data. The study was conducted covering all five shelters and low-income communities in Bangkok, Thailand. The data collection was through Focus Group Discussions, and in-depth interview. Forty-five cases were purposively recruited into the study. The data were analyzed using content analysis.

*Results*: From the qualitative research, the results indicated that the majority of the young women tended to select abortion as their first choice, while a few cases continued their pregnancies to term without any attempt to terminate the pregnancy. These women tended to delay seeking assistance once they found out their pregnancy situation. Majority of the low-income young women were more likely to resort self-medication as their first alternative, which was sometimes hazardous to their health. Decision on self-medication or abortions with unskilled personnel sometimes resulted in serious and life-threatening complications. Many pregnancy-termination situations resulted from self-medication. While most of these women made the important decision themselves, they still sought advice, guidance, and support from their partners, peers and parents. When the first attempt failed, they would seek a second or third attempt until they felt it was not possible to achieve what they had planned.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

Most women with unplanned pregnancies knew that they had options regarding termination of their pregnancies but their main concern was confidentiality. Consequently, the women who wanted to terminate their pregnancies adopted three patterns of action, i.e., 1) visiting drugstores / grocery stores, 2) visiting private clinics or hospitals, and 3) using physical pressure or vigorous actions. Most of the women realized that private clinics provided effective methods for terminating pregnancy, but due to its high cost, they would first resort self-medication or self-management. If they were not successful, they would then visit a private clinic and asked others to support the cost.

Understanding the women's decisions making process and their health seeking patterns utilization allows us to understand their decision and course of actions. The program managers, implementers, providers, partners, parents/relatives, and friends should do as much as possible to support the decision of the women in order to provide better information and services to reduce the impact, both physical and mental, of their selected choice.

## **1. BACKGROUND AND RATIONALE**

#### 1.1. Situation of Unplanned Pregnancy Among Young Women

Worldwide, there is an estimated 15 million adolescents age 15 to 19 who give birth, approximately for up to one-fifth of all births each year. Furthermore, each year 1 million to 4.4 million adolescents in developing countries undergo abortion, and most of these procedures are performed under unsafe conditions (RHO, 2002) due to unplanned pregnancies. These pregnancies lead to their health problems in two ways: first, many unplanned pregnancies can threaten the young women's health or well being; they may face a medical, psychological, social problem and lack of resources to support themselves during pregnancy and raise a child. Second, if young women do not have access to or cannot afford safe abortion services, many unplanned pregnancies are terminated using unsafe procedures that can lead to the women's death or disability.

#### Legal Situation of Abortion in Thailand

Induced abortion is a crime under the Articles 301-305 of the 1957 Penal Code of Thailand. Both the women and the person terminating the pregnancy are subject to legal penalty. The women can be sentenced to three years in prison and a fine of 6,000 baht (US\$ = 171; 1US\$ = 35 baht). Heavier prison sentences and fines are prescribed for the person who conducts the pregnancy termination. However, attempted but unsuccessful termination of pregnancy is not punishable under some circumstance, which allows women who have had the operations to be treated in hospital. The other major exception to the law is that medical practitioners are permitted to terminate pregnancies in case of rape or if continuation of pregnancy will endanger the health of the women. During the last few decades' abortion law was debated worldwide, as well as in Thailand. At present, there are some movements of many non-government organizations (NGOs) especially those working in women issues. Some groups of people from governmental organization are calling in for reform abortion law. Until now, nothing in the law has been changed (Gray et al., 1999).

However, changing law takes a long period of time. Prevention and care for long and short-term plan can help solve the problem of the unplanned pregnancies.

## Sexual Health and Risk of Unplanned Pregnancy

Along with increase exposure to unplanned pregnancy, Sexually Transmitted Infection, and Human Immunodeficiency Virus (STI/HIV), young women who engage in sexual activity outside of marriage may face stigmas, family conflicts, problems with school and the potential need for unsafe abortion. Married adolescents and youth who become pregnant may not encounter the same social risk as the unmarried women, but they may face the same complications from STI/HIV and the health risk of early pregnancy, which will cause both physical and psychosocial problems.

Among Thai youth who are sexually active, the median age at first sexual encounter is between 16 to 18 years. The median age for girls was higher compared to that of boys. More than half of those who have had sex had unprotected first sexual intercourse. Moreover, it was found that a relatively small proportion of males had commercial sex workers (CSWs) as their first partner. This behaviour has been changed as compared with the period before the HIV/AIDS epidemic when men tended to go to CSWs more (UNESCO/Thailand, 2001). Now their partner has changed to their girl friend(s). There is clear evidence from many studies showing that sex is initiated at a younger age and sex partner are either friends or lovers of similar age (Boonmongkon et al., 2000). Thus, this situation puts their girlfriend(s) at a high risk of unplanned pregnancies, and STI/HIV.

Among female adolescents, they revealed that unplanned pregnancy was a major problem (Boonmongkon et. al., 2000; Porapakkham et al., 1985 and Deemar, 1980 as cited in Soonthornthada, 1996). Findings from a study of sexual experience of school adolescents in Bangkok showed that 35 percent of sexually active male adolescents stated that their girlfriends become pregnant and 4 out of 5 pregnancies ended with abortion. Moreover, 30 percent of sexually active girls stated that they had had abortion (Porapakkham et al., 1985 and Deemar, 1980, as cited in Soonthornthada, 1996). A study by Soonthornthada (1996) found that out of school adolescents (factory workers) were more likely to accept abortion when compare with school adolescents. Thus from reviewing many studies, it is found that female out of school adolescents.

Furthermore, studies estimated that one out of three pregnancies are unplanned (Chayowan and Nodel, 1992) and 200,000-300,000 women at the reproductive age terminated their pregnancies each year (Koetsawang, 1993). Adolescents are more likely than adults to hide a pregnancy, seeking late term abortions, and having a procedure performed by untrained providers under unsafe conditions, often leading to permanent disability or death (Sertthapongkul et al., 1993; Koetsawang, 1993).

#### Psychological Consequences

Major and Gramzow (1999) found that apart from physical consequences, women who feel stigmatized about their pregnancy are more likely to feel a need to keep it hidden from family and friend. Secrecy was related positively to suppressing thoughts of abortion and negatively to disclosing related emotion to others. More importantly, suppression was associated with experiencing intrusive thoughts and distress. Both suppression and intrusive thoughts, in turn, were positively related to psychological distress over time.

## Impact on Health Services

Spontaneous abortion or uncomplicated case is rarely a fatal and seldom presents complications. It may require up to 3 days of hospitalization, complicated cases may need a stay of up to 5 times longer. The treatment of abortion complications in hospital uses a disproportionate share of resources, including hospital beds, blood supply, medication, as well as access to operating theatres, anesthesia and medical specialists. Thus, consequences of unsafe abortion place great clinical, material and financial demands on the scarce hospital resources of many developing countries (WHO, 1997). In Thailand, it is calculated that cost per case of severe complication per abortion is 21,024 baht (Warakamin and Boonthai, 2001) where as the first trimester abortion is about 3,000-5,000 baht.

## 1.2. Causes of Unplanned Pregnancy

Unplanned pregnancies are caused by several factors including the nature of transition from childhood to adulthood; lack of knowledge, moral, values, and education at home and in school; relationship with parents; gender inequalities situation; inadequate reproductive health resources for providing information, education, and communication to the young people. Moreover, all these factors intern effect on health seeking patterns and access to health care of the young women. Details are as follows:

## Transitional from Childhood to Adulthood Especially Physical and Emotional Development

Adolescence is a transitional phase from childhood to adulthood. While becoming physiologically mature during this transitional period, they become less dependent on their parents and more involved with peers. They begin to form identities as individuals and develop further capacity for interpersonal relationships with others. Moreover, it is a critical period that lays the foundation for reproductive health of the individual's lifetime. It is also a period when *"sexuality"* emerges in the form of physical body changes. Feeling, psychological changes including emotions and consciousness about one's sexuality and these of the opposite sex also occur at this time. It is also a phase of life where one searches for self-identity, is vulnerability to sexual risk behavioral including unplanned pregnancy (AIDC, 1999).

## Lack of Knowledge, Moral and Values, Education at Home and in School

Many young people lack accurate knowledge of reproductive health anatomy, physiology and the pregnancy process including the consequences of unprotected sex. Also, they might lack knowledge of pregnancy prevention and access to family planning methods due to several factors, such as, social taboo on unmarried young women seeking for contraceptive methods.

A large-scale study by Maungman et al. (1983) in Bangkok found that schooled females had different knowledge and attitude about sex than out of school female youth. School adolescent had higher level of basic knowledge about reproductive biology, conception, and awareness of sexuality than out of school adolescents. A subsequent study in 1996, conducted among in and out of school youth revealed that among the sexually active out of school

youth, less than half used contraceptives while having sex. Among the non-users, some of them reported that they did not use any method of birth control due to lack of adequate knowledge while some of them feared side effects (Soonthornthada, 1996).

Another source of information revealed that parents and teachers rarely talk and openly discuss sexual knowledge, values, and moral with their children. This may lead youth to lack of concern for the outcome of irresponsible sexual behavior and to seek information from many unreliable and harmful sources such as pornography, magazine, Internet sites and etc. to educate themselves. Although, some issues on sexual health have begun to be taught in some school effective sex education requires trained personnel which is lacking at the present time.

#### Gender Inequality

In Thailand, young women are more restricted than men in their personal movements because of concerns over their virginity and chastity. After puberty, daughters are not allowed to go out alone. Daughters are kept under strict supervision, while brothers take advantage of their social and sexual freedom. Young men gain status if their peers believe they have seduced many women, whereas young women lose their reputation if people hear that she slept with a man (Boonmongkon et al., 2000; Gray and Pungpuing, 1999; Sainsbury, 1997; Ford and Kittisuksathit, 1996).

This inequality situation puts young women at a disadvantage in controlling sexual relations and contraceptive use. Social experiences and pressures define what is or is not acceptable for a young woman to do. These make it difficult for them to protect themselves from unplanned pregnancy. The inequality among man and woman include social taboos and power over women often prevents her from using contraceptives. Opposition from their partners/lovers is one of the most common reasons women give for not using barrier contraception.

#### Stigmatizing Young Women in Utilizing Reproductive Health Resources

In Thailand, there are limited reproductive health service resources for young people. In the last decade, the existing government health services were for adults, which their operations based on the assumption that the adolescent group did not require reproductive and sexual health services. The use of health care services is complicated and stigmatizes adolescents who come to use the services for their sexual health problems (Limsampan, 1997 and Pracharat, 1990 as cited in Boonmongkok et al., 2000).

#### Health Seeking and Access to Reproductive Health Care

The main obstacle for women in Thailand in seeking abortion once they are faced with unplanned pregnancies is that it is illegal. There are no public health facilities that provide abortion if the pregnancy does not result from specific conditions as stated in the law. There are some private professionals who provide such services, but the cost is high. Prohibitively most young women cannot afford it, thus, they will seek alternative services from other sources, such as, traditional healers, and drugstore personnel. Drugstore personnel are more popular than traditional healers among young people because they do not need to answer questions nor reveal their identity but the most important is easy access to products that young women believe can terminate their unplanned pregnancy (PATH, 2001).

Drugstores are an important source of health services for Thai communities. When faced with mild illness, most of people tend to go to a drugstore as their first choice. In the Northern part of Thailand, a study revealed that 66 percent visited drugstores (Chuamanocharn et al., 2000), where as 80 percent in Bangkok (Punyawuthikrai et al., 2001) and 97.7 percent in the Northeastern part of Thailand respectively (Kanchanaraj et al., 2001). Young people also share the same pattern of health seeking that is they will seek services from drugstores as their primary resource especially when they have to seek for sexual health products; abortifacient for terminating pregnancy; or when faced with any sexual health problems.

Another study in 1998 by PATH/Thailand in Had Yai, Songkla province, revealed that vocational school students utilizing drugstores as the common place to seek health services; particularly in regard to sexual and reproductive health issues. The research showed that youth prefer to seek reproductive health advice from drugstores rather than from the government health care service, because they cannot be identified by medical record showing name, marital status, etc. Youth also wish to avoid confronting sometime-negative attitudes of healthcare providers towards unmarried adolescents, particularly females, seeking family planning services, diagnosis and treatment of Sexually Transmitted Diseases (STDs), pregnancy tests, etc.

Peers also influence young people to seek services. When young women missed their menstruation, they usually consult their female friends who would recommend that they get a pregnancy test from the drugstore or private clinic in town. No young women utilize the government health services for a pregnancy test because they were afraid that their secrecy about their sexual relations would be revealed to the public (Boonmongkon et al., 2000).

#### 1.3. Ways to Solve Problem

Unplanned pregnancy is one of the most difficult life experiences among young women. Women are often confused, and then seek for help and support. When the problem occurs, the woman has three choices including parenting the baby, make a plan for adoption, or terminate the pregnancy. However, to choose one of the options, it is hard for the young women to make a decision.

Thus, to implement or initiate any strategies to help solve the unplanned pregnancy problem, it is important to understand thoroughly how young women making decision, and seeking for help - or health services; when, where, how and why they decide in that ways.

Based on the existing evidences on this issue, so far, there is not a single study that addressed, in a holistic manner, health-seeking patterns among young women on unplanned pregnancy. Moreover, research on abortion failed to take into account the role of unplanned pregnancy, which is an important determinant of abortion. The aim of this study is to formulate a theoretical model of health seeking patterns of service utilization among young women with unplanned pregnancy. Findings are expected to support policy and decision makers in designing interventions and services to help young women to make the transition from adolescent to adulthood without physical and psychosocial trauma from unintended pregnancy. Moreover, this explanatory model will help implementers and providers understand the decision making process, why and how women made decision to solve their problems which in turn can provide more appropriate interventions and services that serve the needs for young women.

## 1.4. Conceptual Framework of the Study

There are many theoretical models on health seeking behavior that attempt to explain utilization of health services. Furthermore, there are many determinants that affect the decision- making options of women with unplanned pregnancies, and their health seeking behavior. However, there has been no definitive conclusion about which factors or variables influence their decision-making options, or their seeking of health care utilization. This is because health behavior involves a number of social and behavioral factors including attitude, perception, beliefs and interpersonal relationships with peers, partners, and family members, as well as the influences of culture, norms, law, and health delivery systems and policies.

Thus, the study conceptual framework of this study is derived from a combination of factors recognized to be important determinants for the decision-making process of options for women with unplanned pregnancies from Ratchukoon (1998) "A Model on Decision-making Process for Terminating Pregnancy".

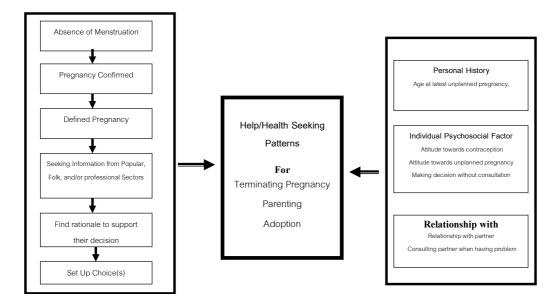


Figure 1.1. Conceptual Frame Work of the Study.

In addition, the "Symbolic Interaction Theory" plays special attention to the symbols individuals use to interpret and define themselves, the actions of other people, and all other things and events. This theory helps us understand how young women interact with themselves when facing the problem of unplanned pregnancy. Moreover, Kleinman's theoretical framework (1980) called "The Health Care System". According to Kleinman, the health care system is a type of cultural system, which integrates all health related components of a society. These components can include the patterns of belief about what causes health problems; the norms that govern choice and the evaluation of treatment and care; power of relationship; interaction settings, and institutions. Kleinman also proposed structural components of the health care system; he divides his model into three overlapping parts: the popular, the professional, and the folk sectors. Moreover, the study's conceptual framework also takes into account some selected sexual and reproductive health variables of the young

people. In addition to services characteristics, indicators that are used for measuring the youth friendly services are selected from Nelson's study (2000).

In conclusion, based on the theoretical models, the study of Naravage et al (2005) on factors affecting decision making of low income young women with unplanned pregnancies, this study classifies determinants of health service utilization into three sets of factors. These sets of factors are including: *1) personal history*. The variables include age at latest unplanned pregnancy; *2) individual psychosocial factors*. The variables used are attitude towards contraception, attitude towards unplanned pregnancy, and making decision without consultation; and *3) relationship factors*. The relationship variables used are relationship with partner, and consulting partner while having problem.

## 1.5. Purpose of the Study

This study is intents to understand when, where, why and how women with unplanned pregnancies seeking for health services, and patterns of health seeking behavior, their access to health care, and their reasons for using or not using available health services. These results provide valuable insights regarding how best to improve the utilization of existing interventions by influencing health behavior and care seeking choices. The study employs qualitative methods to answer the research questions. The qualitative approaches are employed to provide in-depth information to explain a woman's decision-making process to utilize the service facilities to ultimately formulate a theoretical model of help- or health-seeking patterns of women with unplanned pregnancies.

## Research Objectives

- 1. To explore decision-making processes of the young women with unplanned pregnancy who choose for abortion, parenting, or adoption.
- 2. To explore health-seeking patterns of the unplanned pregnancy women who choose for abortion, for birth, or adoption.
- 3. To explore the characteristics of service facilities that influence utilization of unplanned pregnancy services among young people.

## 1.6. Key Themes and Main Issues for the Study

The main part of this study focused on the collection and analysis of qualitative information to understand when, where, how and why women with unplanned pregnancies making decision and seeking for health and social services, their patterns of seeking services, and the gaps between their needs and the services available. However, key themes and main issues are proposed as follows:

## 1.6.1. Definition of Pregnancy and their Interaction

- Definition of unplanned/unwanted pregnancy by women.
- What will happen if women keep the baby to full term?

- Interaction and feelings among women and their significant others, when facing unplanned/unwanted pregnancies.
- Feeling towards the fetus, and pregnancy.

# 1.6.2. Process of Decision-Making, Seeking for Services and their Interaction

- When and why women define the pregnancy as unplanned/ unwanted.
- Sources of information and person that women go for consulting.
- Process of decision making of women who opt for abortion, adoption, or parenting and feeling towards the choice.
- Rationale behind the selected choices
- Choice of services and care seeking patterns of women who select each option.
- Rationale behind the choice of services
- Interaction and feeling towards themselves and service providers.

# 1.6.3. Social Culture and Stigmatization on Sex, Sexuality, and Culture

• Reaction of others towards sex, sexuality and pregnancies of young women with unplanned pregnancies and ways the women response to the reactions.

# 1.7. Expected Outcomes

The Specific Outcomes of the Study are Expected as Follows:

- To understand woman's decision making process and her utilization of services and patterns in order to provide information for policy development, program managers, service providers, teachers, family, and adolescences.
- To provide more in-depth understanding of the relevant factors which facilitate or obstruct the use of certain types of health services by women with unplanned pregnancy. The results also help understanding the low utilization of the government health care facilities.
- To provide valuable insights regarding how best to improve the utilization of existing interventions by influencing help/health seeking choices care, which may lead to danger for women's health.
- Results use as case study to advocate policy development, decision makers, women's NGOs for revision of abortion law, unfair regulation in school and labor force, and decrease in restrictions criteria for doctors to perform induced abortion.

# **2. Research Methodology**

# 2.1. Research Design

Research methodology in this study focused on formative research utilizing qualitative data. The study was conducted covering all five shelters and low-income communities in Bangkok, Thailand. The data collection was through Focus Group Discussions, and in-depth

interview. Forty-five cases were purposively recruited into the study. The interview guides important for the formative research because it relies entirely on the spontaneous generating of questions as they come out naturally from the free –flowing discussion between researcher and respondents. So, a researcher concentrates on guiding the discussion around the themes. The data were analyzed using content analysis

#### 2.2. Sites Selected and Sampling Technique

The study conducted in Bangkok which most of the shelters are located. Listed of all the five active government and non-government shelters, and a foster home both located in Bangkok and outskirt in the women networks' bulletin were purposively selected. A foster home was selected for recruiting unplanned pregnancy women during they visited the home for putting the baby for adoption, or for temporary cares. Moreover, some low-income communities were purposively drawn based on voluntary participating.

# 2.3. Study Samples

The women with unplanned pregnancies were recruited from women who utilized the five government and private (NGO) shelters, a foster home, and women who lived in lowincome communities. Verbal reports by the social workers at the shelters disclosed that women with unplanned pregnancies were aged 13 years and more, and 13 was the youngest age of those temporarily staying in the shelters. Thus, this study selected women aged 13-24 who volunteered to participate. Also, young women aged between 13-24, who had experienced unplanned pregnancies at least once, or women currently experiencing unplanned pregnancies, were selected, including young women in and out of school, and married and unmarried.

#### Inclusion Criteria for the Women's Group

To meet the study objectives, the selection criteria for the young women with unplanned pregnancies were as follows:

Experience of unplanned pregnancy.

Willing to participate in the study.

Total family incomes less than 10,000 bath per month.

#### Number of Samples

There were no fixed figures for the number of young women with unplanned pregnancies. The numbers of the study sample were elicited until information related to the objectives became saturated.

#### Data Collection

It took six months, after getting permission from the directors of the shelters to collect the data.

#### 2.4. Study Instruments

Data collection for the formative research included group discussions, and in-depth interviews. Details of each method are as follows:

## Focus Group Discussion (FGDs)

FGDs guidelines were constructed to understand the group attitudes and perceptions towards series of themes which they expressed in the public, commencing with the more neutral subjects of group's expectations towards the family and pregnancy; feelings and concerns about unexpected pregnancy; the meaning of unplanned pregnancy; choices for women concerning unplanned pregnancy; social reactions towards women who terminated pregnancy, parenting, or adoption; and expectations of services for young women.

Data collection began with introductory group discussions in order to get to know each other and build understanding and a sense of trust between the researcher (a moderator in the FGDs) and the study group. Moreover, they helped to promote, discuss, and foster a friendly atmosphere between the researcher and the sample population prior to in-depth one-on-one discussions. The results of the FGDs were analyzed to identify key issues that needed further investigation in the in-depth interviews. More important, the results of the FGDs provided an independent cross-check on the validity of the in-depth interviews among the samples. After the group discussion, the researcher met with the women who were willing to participate in the in-depth interviews to make next appointments for the interviews. The tapes were sent to research assistants for transcription. The tapes were transcribed in Thai and entered into computer files to facilitate further analysis.

#### In-Depth Interview

In-depth interviews with the women were developed to assess the women's experiences of unplanned pregnancy, interaction with their significant persons, the community and providers, the patterns of help- or health-seeking behaviors, and the reasons behind the choice(s) regarding the unplanned pregnancy. In-depth interviews were employed for both the young women with unplanned pregnancy at the shelters and the women who lived in the lowincome communities in Bangkok. Details of themes in the guideline are as follows:

## Key Themes for Data Collection

The study focused collecting and analyzing qualitative information, to understand when, where, how and why women with unplanned pregnancies sought health and social services, their service-seeking patterns, and the gaps between their needs and the services available. Key themes and issues included the definition of unplanned pregnancy and their interaction with themselves once they faced the problem; interactions and feelings among the women towards their significant persons; sources of information and the persons the women consulted; the processes and patterns of decision making of the women who opted for abortion, adoption, or parenting, and their feelings about their choice; interactions and feelings between themselves and service providers; social culture and stigmatization of sex, sexuality, and their pregnancy; and the reactions of others towards their sex, sexuality and pregnancy, and the ways the women responded to these reactions. The information was a foundation for constructing a grounded theory on the decision-making process, and healthseeking patterns of young women with unplanned pregnancies.

#### Process of Instrument Development and Pre-Testing

After the research guidelines had been constructed based on information obtained through the reviewed literature, some ideas and suggestions were obtained through consultation with the researchers, and experts in adolescent health. The constructed guidelines and interview forms were pre-tested using face-to-face interviewing with women who complied with the same inclusion criteria of the study sample. The pre-testing and the sample group were different groups so that it was confident that the instruments could be used with the study group. Pre-testing allowed the researchers to crosscheck appropriate translation of the questions, the appropriateness of the order, the thoroughness of the response, and assessment of the reliability and validity of the interviewees' responses (for the attitudinal section). Moreover, pre-testing was conducted to permit the researcher to know how much time was required to conduct the interview, and the most appropriate place to conduct the interview. Pre-testing was part of the interviewer training process, to gain more skills before the actual data collection. The women who were exposed to pre-testing were excluded from the study.

The in-depth interview and FGD guidelines were pre-tested. Participants were elicited for the pre-testing until the researchers felt confident about the quality of the instruments, questions, responses, and the interviewers' skill levels. Participants who completed the pretesting had similar characteristics to the intended study participants.

## 2.5. Definition of Terms

a) Young women or young people refer to women aged 13-24 years. However, there are different words, definitions, and age ranges used to describe the transition from childhood to adulthood among this age group. Thus in this study, the terms "Adolescents", "Youth", "Young Adults" and "Young People" are interchangeable.

**b)** Health-seeking patterns refers to young, low-income women with unplanned pregnancies utilizing the services of both the formal and/or informal sectors (popular, professional, and folk sectors), to seek care, support, or treatment once they have decided to opt for abortion, birth and adoption, or birth and keeping the baby. Health-seeking patterns are influenced by several factors, including socio-demographic characteristics, individual psychosocial factors, significant others, and environmental factors of the service facilities.

c) **Patterns** refers to the regularity (similarity, commonality) in what women with unplanned pregnancies disclose about their processes of seeking help and/or utilizing services based on selecting choice(s).

d) **Popular sector** is a matrix with several levels, including the individual, the family, social network, and community beliefs and activities. It is the lay, non-professional, non-specialist, popular culture arena in which illness is first defined and healthcare activities are initiated based on beliefs, attitudes, relationships, interactional settings, and institutions (Kleinman, 1980).

e) **Professional sector** is composed of professional physical and mental healing personnel. They have completed training in modern scientific medicine, and comprise trained drugstore personnel, nurses, doctors, social workers, and paramedics (Kleinman, 1980).

**f)** Folk sector is composed of non-professional healing personnel; traditional birth attendants (TBA), and traditional healers (TH) are in this sector. Folk medicine is a mixture of many different components; some are closely related to the professional sector, but most are related to the popular sector (Kleinman, 1980).

**g)** Low-income is defined as a total family income of less than 10,000 Baht per month (Pisalbutr, 1997).

**h)** Choices is defined as options available for young women to solve unplanned pregnancies, including abortion, parenting, or adoption.

i) Decision making process is adapted from Rutchukul (1998) which defined as a process that young women with unplanned pregnancies use to solve the problem by 1) seeking information; 2) arranging their choices; 3) finding rationales to support their choice; and 4) performing as planned. Women who made their decisions might seek services from the formal and informal sectors for abortion, birth and adoption, or birth and keeping the baby. Decisions were influenced by various factors, including socio-demographic characteristics, individual psychosocial factors, relationships with significant others, and environmental factors.

#### 2.6. Data Collection and Management

Data collection took 6 months, from the first week of October 2002 to the end of March 2003, at the five shelters, and in the communities through peers in Bangkok. The researcher interviewed all the study women individually because the questions and discussion concerned premarital sex, pregnancy, and pregnancy termination, and were very sensitive and illegal in Thai culture, and also because the qualitative research had not targeted a sample size number. For this reason, it was important that it was the researcher who made the decision on when the data were saturated and when to stop further recruitment. In addition, most of the women did not disclose their experiences to anyone they were not familiar with. Moreover, by collecting data alone the researcher could ask for and probe in-depth information about each woman's experiences. At the beginning of data collection, the researcher visited each shelter to join their routine activities and assist the social worker teach how to prepare to be a healthy mother. At the session, the social worker introduced the researcher and let her share experiences of pregnancy with the women in the shelters. Thus, a sense of trust was gained from the women, which made it easier to get collaboration from the women. The steps of data collection and management were as follows:

# 1) Preparation

Prior to data collection, we as the researchers coordinated with peers, social workers, and the directors of all the shelters to make appointments to ask for permission to conduct the study. After receiving permission from the directors and commitment from the peers, the study sites and logistics were prepared, and the data collection tools were developed and pretested.

# 2) Training Research Assistants

The roles of the research assistants were to be note-takers while conducting the FGDs, transcribing tapes, and recruiting and screening the study participants for the researcher. Having the research assistants as the note-taker is important in case the samples do not give the verbal expression in response to the specific issues; then the note-taker can make a note on the fieldnotes. The inclusion criteria for the research assistants were that all of them had the opportunity to contact the potential study group, and had positive attitudes towards premarital sex, sexuality, and abortion. The training content included the sexuality in the Thai context, socio-political values and norms of pregnancy among young unmarried women, unplanned pregnancy and choices. Moreover, the research assistants were briefed about the how to take note in the qualitative ways. The pre-test guidelines were used as a training tool to reinforce the research assistants' skills.

#### 3) Steps in Data Collection

The young women with unplanned pregnancies who complied with the inclusion criteria were elicited for the study. The researcher explained the purpose of the study and read the consent form to the participants in order to get permission to gather information according to the guidelines and the interview forms.

The main concerns at this stage were confidentiality and breaking the ice so that the respondents trusted and felt confident about disclosing their problems freely. The process maintained respondent confidentiality at all times. Information collected through interviews and observation was made permanent by removing the names and other identifying information from the data, to ensure that the names of the participants did not appear on any paper, in any circumstance. A signed consented form was utilized with all of the cases. All data were kept in a secure place to protect them from the authorities.

Tape recordings were used with all cases, if permitted. If any case did not permit tape recording, the researcher used short notes, key words and expanded the notes after returning from the field. The recorded tapes were transcribed and prepared for data analysis. Apart from the notes, comments about the appearance and reactions of the respondents were observed and noted.

# 4) Performing Quality Control Checks

All of the data were checked in the field to ensure that all the information was properly collected and recorded. Before and during data processing, the data and tapes were checked again for completeness and internal consistency before being sent to the research assistants for transcription. After the tapes had been transcribed, a hardcopy of the transcription results were rechecked again for data consistency. Moreover, the results were cross-checked with other sources of information, including observation, FGDs, and in-depth interviews.

#### 5) Data Processing: Categorizing and Coding

For the qualitative data, preliminary analysis identified the key themes and issues for further analysis. The preliminary data analysis was processed by hand on a master tally sheet, then content analysis was employed for data analysis.

## 2.7. Data Analysis

During data collection, the researcher analyzed information case-by-case and built a set of basic key issues for further study and investigations in subsequent interviews, which formed the basis for the grounded theory. The accumulation of information with each case was gradually adjusted and the theory clarified until it had reached a stage of theoretical saturation. At this point, the researcher stopped recruiting new participants. Primary data analysis took place immediately after collecting the data for each case, so that the researcher could complete the information while it was still fresh in memory.

Content analysis was used for analyzing the total data. The choice of options by the young women with unplanned pregnancies (abortion, parenting, or adoption) were constantly compared to identify commonality or difference for each choice that the women made to solve the unplanned pregnancy, factors that related to the options considered by the young women, their decision-making processes, their health-seeking patterns, and factors affecting their health-seeking behaviors. Using this method, the researcher could make inductive conclusions from the findings into a larger theoretical picture, or theory, of unplanned pregnancy.

#### 2.8. Ethics and Confidentiality

This study had been reviewed and approved by the ethical committee under the Medical Sciences Faculty, Chulalongkorn University, Bangkok, Thailand. During and after data collection, the maintenance of privacy and confidentiality was very strict. These issues are important because it is a sensitive topic that deals with sexuality, abortion, and political issues. The interviews were conducted where questions and responses could not be overheard. Thus, all of the sample population was informed about confidentiality. In addition, all information was kept confidential according to human subject protection guidelines. Anonymity was employed and maintained during the study. No full names or other information that could identify participants were recorded during any portion of the study. Field notes, tape transcripts, and any other field data collection forms used during the research were collected and stored in a secure location where unauthorized persons could not access them. More importantly, all of the participants in the study were voluntary. The voluntary nature of the study was stressed at the time of recruitment and again at the start of the indepth interviews. Also, at any time during an interview, a participant was free to leave or terminate the session.

According to the ethics committee standards in Thailand, for women aged under 15 years, the representative of the shelters is the one who is authorized to sign the consent form on behalf of these girls, while women aged 15 years or older were authorized to sign the consent form themselves.

# **3.** FINDINGS

Results of the study present the analysis of focus group discussions and in-depth interviews undertaken among young women. Thus, the main discussion focuses on individuals, partner, family, environment, and socio-demographic characteristics that affect the options considered by young women and their health-seeking patterns--abortion, parenting, or adoption.

#### 3.1. Profile of Samples

#### 3.1.1. Focus Group Discussions (FGDs)

Five Focus Group Discussions (FGDs) were conducted among the women with unplanned pregnancies, between October to December, 2002. The participants were recruited from women who lived in the five selected shelters in Bangkok. Since the discussion topics were sensitive, dealing with feelings towards unplanned pregnancy and sexual health, and the participants were in this situation, discussion initially was difficult and it took time to "open up".

The focus group discussions were held in a room where strict privacy and a lack of interruptions could be assured. The general atmosphere in the group discussion was informal, in order to get to know each other and gain more trust among the participants. Snacks and soft drinks were served at the beginning of the group discussion. The researcher conducted the group discussion as a moderator with an assistant who was a note-taker, and was also a social worker. The number of participants in each group generally ranged from 6-10, although there was one group comprising 4 participants, which was too few due to a small number of women who fall under the inclusion criteria were limited. There were a total of 37 participants aged 14-24 years.

#### 3.1.2. In-depth Interviews

In-depth interviews were conducted among young women with unplanned pregnancies who stayed at the shelters, and women in low-income communities in Bangkok. A total of 45 cases participated in the study during the period October 2002-end March 2003 (Table 3. 1). The women who decided to raise the baby were the majority in this group, and the remainders were adoption and abortion, which were 28, 11, and 6, respectively. It was found that the proportion of the middle adolescent and late adolescence was equal (22:22), whereas one case was 14 year old with mean age of 19.7 years of the total sample. More than half (28 participants) was recruited from shelters, whereas 5 cases used the snowball technique, and the remainder was recruited through community health centers' record. Regarding the participants' status, 28 were single, while 17 participants were married. Thirteen participants had studied beyond grade six. For living status, 14 of the participants lived in a dormitory, and 16 lived with their parents or caretakers. The remainder or 15 lived with their partners.

General Characterstics	Number	Percentage
Adolescence Stages (Mean age =19.7 years)		<u> </u>
Early adolescence (13-14 years)	1	2.2
Middle adolescence (15-19 years)	22	48.9
Late adolescence (20-24 years)	22	48.9
Total	45	100
Women Recruited from		
Shelters	28	62.2
Communities/Snowball technique	17	37.8
Total	45	100
Marital Status		
Single	28	62.2
Married	17	37.8
Total	45	100
Educational Level		
Primary education (grade 1-6)	13	28.9
Beyond grade 6	32	71.1
Total	45	100
Living Status		
Alone (Dormitory)	14	31.1
With parents/care taker	16	35.6
With Partner	15	33.3
Total	45	100

# Table 3.1. General Characteristics of Study Women with Unplanned Pregnancy

# 3.2. Experiences of Young Women with Unplanned Pregnancies

In Thai culture, men and women are not treated equally. The biases start from when the baby is born. Gender biases are clearly shown regarding sexuality. For men, pre-marital sex is socially accepted, but it is not for women. Only sex among married women is socially accepted. Moreover, as mothers, women are expected to raise, care, and feed the baby. They must not only care for the baby, but it is socially expected that women should care for all family members. In addition, if women want to terminate their pregnancy they will incur the blame of society. Furthermore, it is illegal to do so in Thailand. This situation puts young women in a crisis situation once they are faced with an unplanned pregnancy.

If women have sex before marriage, or while they are students, their parents, family members, teachers, friends, and other people in their community will blame them. Moreover, if a woman gets pregnant without a responsible man, society judges her as being promiscuous. Thus, most of the young women with unplanned pregnancies are afraid and want to hide themselves once they face trouble. They do not dare to confront anyone they know. The most important factor is that these women want to avoid seeing the reactions of these people, especially their parents. Thus, this section presents the interactions and meanings of the young women while they are having trouble with their significant persons, including their partners, parents, peers, themselves, and their providers, and the reasons for unplanned pregnancy. These results help us understand the young women's experiences, their thinking and to find opportunities to assist those who have unplanned pregnancies endure and cope with the critical situation with physical and mental well-being.

# 3.2.1. Terms and Meaning of Unplanned Pregnancy

The meanings and terms for unwanted pregnancy respond to individual situations during pregnancy and after delivery. Most important is the relationship of the women and their sexual partners. If the relationship is good, the terms are more positive, while if the relationship is bad or there is no relationship, the terms are negative. The following are the terms raised by participants during the FGDs:

The participants mentioned "Thong mai prom" and "Thong mai thang jai" most frequently. These two terms reflected similar situations for the women faced with an unplanned pregnancy. Some of the participants were still loyal to their partners even though they had left them. However, the women felt unprepared for raising the baby because some of them were studying, unemployed, or their parents did not accept the pregnancy. The majority of them raised the baby by themselves, whereas some of them put the baby up for adoption.

"The reason it was "mai prom" was because I intended to have a baby, but I had this problem (partner leaving with another woman). Thus, I felt lost" (Lee, married, 21 year-old factory worker, put the baby up for adoption).

Local Terms	English	Meanings
"Thong mai thong	Unwanted	The woman did not want to have a baby at all.
karn"	pregnancy	She attempted to terminate the pregnancy using
		various methods, with negative attitudes
		towards the man.
"Thon mai thang jai"	Unintended	The baby is wanted but the woman is not ready
	pregnancy	to have the baby at that time, because of study,
		work, or unemployment. It reflects a positive
		relationship with the sexual partner.
"Thong mai prom"	Unplanned	In some situations, the baby is wanted, while in
	pregnancy	others, not. The woman's relationship with the
		partner is positive.
Thong mai kadkit	Unexpected	The woman did not want to have the baby at
	pregnancy	that time because of a lack of mental and
		physical preparedness.

"Thong mai thong karn" was mentioned among the women who had negative relationships with the men, because some of them were raped by both known and unknown men. Some of the women had very negative impressions of the man, or felt the enormous burden of having a new baby. Women in these situations tended to put the baby up for adoption, whereas some of them raised the baby by themselves because they became attached to the baby during pregnancy and/or after delivery and after having raised the baby for a while. During these periods, attachments bonded subtly.

"The reason for "Thong mai thong karn" is because I was raped and I couldn't stand for it" (Pia, single, 17 years old, school student, parents raising the baby).

#### 3.2.2. Feelings and Concerns of Women Facing Unexpected Pregnancy

The results from the focus group discussions with the women in the shelters revealed that most of the young women who had faced this situation recently felt anxiety immediately after suspecting pregnancy. They felt concerned because premarital sex, sex while studying, or pregnancy, without a responsible man, were not socially acceptable. Thus, they were afraid that their parents were angry and disappointed in them. Other more minor issues included acceptance by their relatives, friends, and people in the community. Many women felt that these people would look down on them. Some women disclosed that they cared about their parents' concerns the most. However, if their parents accepted their pregnancy, they would feel relieved and calmer. Also, most of them did not want anyone to know that their partner was an unfaithful man, irresponsible, and had abandoned them. Some women were afraid because their parents would not accept their partner's behavior. The concern about parental worry was due to the women feeling that they cared for their parents the most, and did not want to disappoint them.

With urbanization, the relationships between women and their friends or their communities are bonded weakly. The women who stayed at the shelter for longer periods were not much concerned about the reactions or thinking of society towards unplanned pregnancy. This was because, after the women had passed this crisis situation of unplanned pregnancy that was related to their parents, friends and community, the major concerns were the baby and the future. The most important was the situation in the shelter. They had the opportunity to meet women in the same situation. After the newcomers had interacted with others, they felt that there were many women who were falling into a worse situation than themselves. Thus, they felt more relaxed and happier than staying outside. However, at the shelters they were worried about how to manage their lives with the baby, and what the future would be. Most of them revealed that, as a single mother, they were afraid that they could not raise the baby or provide it with a good future. During pregnancy, some of the young women at the shelter could not make definite decisions to raise the baby or put the baby for adoption. This was because many of them relied on their partner's, parents' or relatives' support to raise the baby. Some of them could make a decision once the baby was born, and they gave the reason that it was because the baby's face looked like them. Moreover, many of the women who took several abortifacient products were afraid of the baby being abnormal. The majority of the women made decisions by themselves about options for solving unplanned pregnancy, which was due to the most-mentioned factors, their partners' responsibility, and support from their parents/relatives. They revealed that if their partner agreed to be responsible for the baby, most would raise the baby instead of putting it up for adoption.

As time passed, most of the participants adjusted themselves and felt more relaxed and comfortable with the situation, especially, women who stayed at the shelter, who were protected from stigmatization because they were among friends faced with the same situation. However, most of the women who kept and raised their baby would worry about it, and this issue would not be easily resolved because of economic problems. The women really needed financial and other supports to raise and adapt themselves to the baby, especially when they went back to their communities. If women lacked the support of their parents, relatives or partners, they would feel reluctant to raise the baby. At the same time, the connectedness between mother and baby from pregnancy made them feel that they should raise the baby instead of putting it up for adoption. These feelings put them into a crisis situation for making decisions about the future.

#### 3.2.3. Choices of the Young Women with Unplanned Pregnancies

Once the participants had realized that they were pregnant, most of them tried to terminate the pregnancy by self-medication, not only to avoid stigmatization, but also because of the convenience, low cost, and because it was easy to do. The reasons for terminating the pregnancy included: starting a new family life; financial problems; having just started a new job or being unemployed; studying; parents/relatives disappointed; premarital sex. Drugstores/grocery stores were the most-mentioned places for purchasing abortifacient products. Most of the women would try as hard as they could to terminate the pregnancy. If some of them failed to terminate the pregnancy themselves, they would visit private clinics. However, many of them took time utilizing abortifacient products and waiting for the results. When the women realized that self-medication was unsuccessful, it was too late for them to have a modern medical practitioner manage an abortion, because the pregnancy term was beyond the medical criteria. So, many of them sought a place to hide themselves and support them during their pregnancy. They maintained the pregnancy to full term because they had no choice. Thus, after delivery, some of them raised the baby by themselves, if they got support from their partners, parents or relatives. In contrast, if there was no support from anyone, the women tended to put the baby up for adoption.

# *3.2.4. Details of Findings for the Reasons of Unplanned Pregnancy and Interactions:*

**1. Individual.** The in-depth interview results showed that many issues caused unplanned pregnancies among young women, including:

1.1. Lack of contraceptive knowledge. Comparing married and unmarried young women, it was found that the knowledge of both groups regarding contraception, and improper utilization of contraception, were not different. Both groups had inadequate knowledge of how to use contraceptive methods properly. Some women lacked knowledge of contraceptive methods. It was surprising that some married women revealed that they did not use any contraceptive methods because they did not know any.

"At that time, I did not protect by using any contraception. I do not know any contraceptive methods. I really knew nothing" (Yam, 19 years old, married).

Some married women indicated that they experienced side effects from using contraceptives, so they decided to stop using them. Subsequently, many of them became pregnant because of discontinuing, or intermittently using, contraceptive pills.

"I know the contraceptive methods. My partner buys them for me, but I cannot take them. When I take them, I get nausea and vomiting" (Pung, 18 years old, married).

"I feel afraid to take contraceptives. I am not sure what will happen if I take it incorrectly" (Air, 20 years old, single).

"When I completed the oral contraceptive course, I thought that it would be OK if I missed it for a while. I felt that missing only one month would not cause pregnancy. So, I would buy it and take it next month. No longer, my belly is getting bigger" (Ploy, single, 17 years old).

"I do not use them (oral contraceptives) continuously. I use them and then stop. As a result, I get pregnant" (Nok, 23 years old, single).

1.2. No time to receive contraceptive services. One reason for unplanned pregnancies was that the women had no time to receive services. Most of the services that they visited were under government authority and operated only during official office hours. Some women said the reason that they could not go was because, if they went, their employers would deduct their daily wages. Some women worked in the factory, and if they could get more work, they could get more money, as well. Thus, they did not really get contraceptives regularly.

"At first I planned to get an injection, but I could not go because I had to work until 10pm. In the morning, I go to work on the factory bus. Everyone needs to get the bus on time. I work as a daily worker, so if I leave I will lose the daily salary" (Kwan, 19 years old, married).

1.3. Inconvenient using contraception. Some married women revealed that they could not choose contraception as they wished because they had no time to get an injection or other form of contraception at government health facilities. If they went to get services from private facilities, it would cost more money, which they did not want to spend. So they ignored using it.

"We used to use natural contraception. He used to use withdrawal, but he does not want to do so lately. I have condoms that I ask him to use, but he does not want to comply" (Pla, 19 years old, married).

1.4. Beliefs and attitudes towards sex. It was found that some unmarried young women believed that having sex only once could not cause pregnancy. Moreover, some single young women acted like their peers, and if their peers did not use contraceptives, they did the same. Some women revealed that they studied at a women-only school, and had not previously learnt about contraception. However, the married women were more independent from their peers and had more experience of sexual intercourse, so their attitudes were different, as mentioned above. The results reflected a lack of inappropriate sex education in school, and peer pressure. Moreover, if they worked in an environment that did not support information, education, or communication about sexual health education, the women had no opportunity to learn about these issues. Thus, having sex once, they did not know how to protect themselves from unplanned pregnancy.

"I never had regular sex with anyone, so I did not realize that I was pregnant. I thought that having sex only once could not cause pregnancy. But one day, my colleague said that I looked fat, which made me aware of my pregnancy" (Oam, single, 18 years old, school student).

"I just pretended to ask my friend whether she used contraception. She told me that she did not use any. Thus, I followed her" (Lek, single, 18 years old, unemployed).

**2. Partner relationship.** The most important reason that caused the young women to feel that they were not ready for pregnancy or to care for the baby was their partner. Many young women revealed that their partners just abandoned them and did not show any responsibility for their pregnancy. Some men left after they had sex without knowing that the woman had become pregnant. Some men abandoned them and left for a new woman. Some men already had wives, so, when the pregnant women knew, they were disappointed and separated from them. However, if the men showed responsibility, the majority of the respondents said that they were willing to reunite. Thus, many women kept the baby to make a new decision, which depended on their partner.

Among the participants who faced violence, some women were beaten by the men because they used drugs, such as amphetamines or alcohol. Once they had taken it, they could not control their behavior or emotions. Many participants were beaten; the married women were more seriously beaten by their partners, while the single women were less seriously beaten, because among the unmarried there was no social bond. So, they just ran away after they were beaten. Some married women accepted the violence with no choice because they needed financial support from the man. Moreover, if they wanted to separate, their parents or relatives did not allow it, and asked them to come back because they did not want anyone to gossip about their family.

2.1. The men used drugs, and/or alcohol and lost control. When the men used drugs, they lost of control and did not take responsibility for the family.

"I did not know that I was pregnant. I am thinking about separating from him because I knew that he is a drug addict. I tried to run away from him; I escaped and went to stay with my friend. However, my parents knew and asked me to come back. I agreed to do so" (Jum, married, 24 years old).

"If he did not beat me, I would live with him. When he beat me the third time, I told him that I would leave him. Then, I ran away from him" (Fon, single, 19 years old, housewife).

"At the beginning of pregnancy, we did not separate; I felt ready for the pregnancy. When he used amphetamines and had another partner, I felt that I did not want the baby" (Fon, single, 19 years old, unemployed).

**3.** Unintentional. Some women were not sure whether to settle with the man as a permanent partner. Because they lived away from parents, they had more freedom in their life. Some of them just wanted to try to have sex like their peers without loving each other. Some of them wanted someone to be their friend because they lived away from home and felt lonely. Thus, when they spent more time together, they quarreled and finally separated because the relationship was bonded weakly (Gay, single, 20 years old).

*3.1. Having sex by accident.* The lifestyle of single young people has more freedom than in the past, which was controlled by their parents or relatives. There are also more places to go out and spend time together. In addition, some of them mentioned that they had more opportunity to have sex without love but by chance.

"He asked me to go to his house to see his parents. I decided to go with him. At his house, there was no one. Then, he forced me to have sex with him" (Pia, 17 years old, school student).

"I like him as an older brother, not like a boyfriend. It is not possible to live together. If I lived with him, our family life may collapse soon" (Oil, single, 20 years old, vocational school student).

#### 4. Parents and close relatives

4.1. Living away from parents or relatives. Many of the women with unplanned pregnancies lived alone or with their friend(s). They had more freedom without the control of their parents or relatives. Nowadays, sex among young people is more accepted than in the past. With peer pressure, many of the participants had premarital sex. However, once they got pregnant, they could not hide it because the symptoms and physical signs started to show. Many of them gave the reason that they were afraid that their loved ones would be disappointed, since they were still students.

"I was desperately depressed. I was so afraid when the provider put the strip into the urine. I wished that I would not be pregnant. I thought my future would disappear because of my pregnancy. All of the efforts of my parents sending me to school would vanish because I got pregnant. I hated the baby very much" (Oam, vocational school student, 18 years old).

"When I knew, I felt very worried. My father would not accept me. Now, he knows and accepts" (Pu, single, 23 years old, unemployed).

4.2. Family problems. Nearly half of the single young women were from broken homes. Some of them lived with a single parent or together with their father or mother-in-law. Some of them could adjust to the new family but some could not. Some of them left their new father or mother-in-law because they disliked them, or their parents disliked the women's partner, or

they could not accept the way the women behaved. Among the married women, many lived separately, so they had fewer family problems.

"There are many reasons. One reason is that I ran away from my mother. I do not want a baby. I only want to work and collect all the money. If I raise the baby, I will have no money left. Then, people I know will look down on me because I cannot survive by myself" (On, single, 20 years old).

4.3. Having relatives with experiences of unplanned pregnancy. Some women had relatives with experiences of unplanned pregnancy. Hence, they tended to have an unplanned pregnancy.

One example was the two young sisters Lek and Nu, who were 20 and 18 years old, respectively, living with their parents. They had a younger brother and sister who were studying at school. After graduating in grade six, both of them worked together at a small factory in Bangkok. They worked for a few periods of time and then guit after that. When the researcher interviewed them, they were both unemployed. Their father was a taxi driver, while their mother was a house worker. Last year, the eldest sister, Lek, got pregnant but her boyfriend who was 4 years younger ran away. She decided to tell her younger sister, Nu. The younger sister decided to tell their mother. After her mother recovered from the shock, she asked her daughter to terminate the pregnancy and she agreed to do so. She took her daughter, who was two months pregnant, to the abortion clinic and signed for her. Nu was very afraid while she was at the clinic because she was afraid of side effects, such as bleeding and pain. However, she was successfully terminated and was happy. The following year, her younger sister, Nu, became pregnant. She kept it a secret from their parents. She was aware of the pregnancy at the second month, because she went shopping at the Mall and fainted. Nu went to check at the clinic and knew that she was pregnant. Her boyfriend put pressure on her to terminate the pregnancy because he was not ready to have a family. She did not want to do so because she realized that her older sister was suffering from guilt after terminating her own pregnancy. Her boyfriend pressed her again and took her to the abortion clinic. Finally, on the day, she decided to tell her grandmother about the pregnancy because she was closer to her grandmother than her mother. Then, her grandmother told her mother. Her mother forgave for the past and continues to support her.

"If I terminated pregnancy, it might come to bother me. Like my elder sister, she terminated pregnancy. Then, she had a bad dream...the baby came into her dream and it wanted to live with her" (Nu, single, 18 years old)

**5. Peer pressure.** The women who had friends with experience of premarital sex and/or unplanned pregnancy tended to behave in the same ways as their peers, since the young people spent more time with their peers than their parents. Some of them lived together with friends to save money while they were working or studying. So, the relationships among them were closer.

As one example of a young factory worker with unplanned pregnancy, the researcher interviewed "Nid", who was 21 years old. Her parents separated when she was young. Her mother remarried, so Nid lived with her grandmother until she moved to work in a factory in Samut Prakan, adjacent to Bangkok. She lived in a room together with 2 other young women

who worked at the same place. They all had premarital sex with their boyfriends. "Nid" met a man who became her partner at the age of 17 years. They had one child but they sent the child to her mother-in-law upcountry. Three years later, she fell pregnant again. She felt very disappointed because she took oral contraceptive pills every day. She thought that the pills might have expired. Both of them were unhappy because they did not have enough income to support the second baby. At the sixth month of pregnancy, her partner disappeared. Her belly was getting big and she was laid off from the factory. Her savings were running short, so she decided to go to the hospital near her apartment to request termination of her pregnancy. The social worker referred her to a shelter because the hospital did not provide abortion services.

She disclosed not only her own miserable memoirs, but also a sad story about her roommate who came from the same province in "Isan" (northeastern part of Thailand), and from a broken home. Her parents had separated. Later, her father passed away and her mother remarried. She came to work in the same factory as Nid. During the economic crisis of the past five years, the factory laid off many workers and she was one of them. Later, she went to work as a waitress in a restaurant and had sex with the customers to earn more income. She fell pregnant later without knowing who the baby's father was. She was very depressed and committed suicide later, using a high dose of pesticide. After the police examined the dead body, Nid and the other roommate took her body to the temple for cremation. They sent her bones back to her mother in her hometown.

Nid was more stressed after her roommate passed away. She was waiting for her partner to return, but it seemed hopeless. Thus, she decided to seek an abortion during the sixth month of pregnancy, but it was not successful because the providers refused to terminate any pregnancy when the term was more than three months. However, it was fortunate that she was referred to the shelter; otherwise she would have suffered very much.

6. Family income. All of the participants in this study were selected based on their incomes as well as other criteria, which were lower than 10,000 Baht per month. Having this condition, it was found that the greatest concern for the participants was dealing with economic problems. Some of them did not go to government health facilities because they found it was not convenient. They had no money to buy oral contraceptive pills or condoms, so they did not use any contraceptive methods, especially married women who had sexual intercourse regularly. Thus, when they were pregnant, many of the young participants were concerned about the baby's future. Both the married and unmarried women were concerned about the same issue, because some of the young people had no job, were studying, or had been laid off because of their pregnancy. In addition, some married women had the baby more than one. Thus, the married women needed more income to afford the expenses of the whole family. However, it was found that they earned the same or less income, but had increased expenditures.

"I was concerned after the second baby was born. How will I raise them both? When the elder and the little one are crying at the same time, what will I do?" (Maew, 18 years old, married, housewife).

"I do not know how to manage it. I just started working less than a month ago. My salary is only 1,500 Baht per month. I do not know what to do" (Lek, single, 23 years old).

"I plan to have an abortion, taking the baby out of my womb. I really do not want to keep it as a burden for my parents, because I have another younger sister who is 11 years old and still studying" (Kwang, single, 15 years old).

**7. Rape.** A few of the cases staying at the shelters had been raped. The proportions of the single women, who had been raped by a man they knew, and by a stranger, were equal. Some women did not know that they had a chance of falling pregnant after being raped, and when they knew of the pregnancy, it was too late to terminate it. However, if they knew earlier, in the first trimester, they could have terminated the pregnancy legally at either public or private health facilities. Most of the women who were raped realized they were pregnant when it was too late, because they had no signs and symptoms of pregnancy. When the pregnancy was confirmed, most of the cases kept it secret because they were ashamed to disclose it to anyone. They waited until there were physical changes, which was too late to solve the problem.

"When I was raped, I felt disgusted because I did not want to have sex with him. If I was willing, it would be another story. So, I decided to be a nun" (Oam, single, 18 years old, school student, raped by a known person).

"At the beginning, the man said he would be responsible if I had sex with him. After he had sex with me, he told me that he needed another two years to complete his education" (Daw, single, 24 years old, raped by a known person).

"I was stressed because my aunt knew. Then, all my relatives at home would know, and they would yell at me and ask me who the father of the baby was. It made me scared. I did not dare to tell them what was happening to me" (Pae, single, 16 years old, out-of-school youth, raped by a stranger).

#### 3.3. Decision-Making Process

The in-depth interview results were used to explain the decision-making processes of the young women with unplanned pregnancies. It starts by examining the interaction process of the women when knowing of pregnancy, their decision towards unplanned pregnancies whether to terminate or continue their pregnancy upon missing menstruation, the definition of pregnancy, consulting popular sector, compromising with self-conflict, and making choices (Figure 3.1).

# 3.3.1. Knowing of Pregnancy, Feelings, and Defining Pregnancy

**Knowing of Their Pregnancy.** The results of the in-depth interviews showed that the women knew their pregnancy and interacted with themselves in the followings ways (please see figure 3.1):

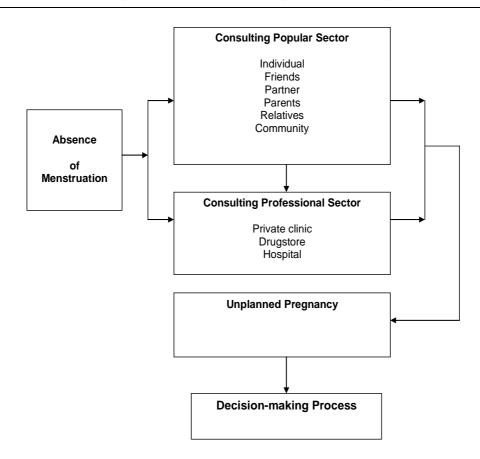


Figure 3.1. Seeking Patterns of Pregnancy Confirmation among Low-income Young Women with Unplanned Pregnancy.

1. Knowledge and experiences based on signs and symptoms. The knowledge and experience of pregnancy made the women aware that they were pregnant. Most of the participants recognized their pregnancy because they knew the signs and symptoms. The most popular sign of pregnancy was missing menstruation. Many of them revealed that if they had sex and then missed menstruation, it was certain that they were pregnant.

"Since missing my menstruation, I did not take a urine pregnancy test because I used to have a baby. I felt confident one hundred percent" (Pla, married, 19 years old, housewife)

To confirm the pregnancy, most of the primigravida visited private clinics, whereas the same proportion visited drugstores/grocery stores to purchase a pregnancy test kit to perform on their own, to confirm the pregnancy. However, some cases did not trust the self-test and would visit a private clinic for final confirmation. A few cases did not make any confirmation because they had experienced pregnancy before. A few cases visited government hospitals or community health centers for pregnancy tests, because they planned to come back again for antenatal care, delivery, and/or post-natal care.

"I missed my menstruation until the third month, when I decided to ask my friend to buy a urine pregnancy test from a convenience store (7-Eleven)" (Pia, single, 17-years old, high school student).

"I tested (urine pregnancy test) myself but I was not sure, so I decided to borrow my friends' money to visit a private clinic to confirm the pregnancy" (Aui, single, 19 years old, vocational student).

**2. Fetal movement**. Some knew that they were pregnant because of the movement of the fetus and physical changes. Some of the young women had had irregular menstruation periods since the onset of puberty, and thus did not realize when they were missing menstruation. They realized when they felt something moving in their belly before others suspected that they were pregnant.

"I felt that there was something moving in my belly. In addition, my relatives suspected, so they took me to the clinic. The results showed that I was pregnant" (Pae, single, 16 years old).

**3. Physical changes.** A few participants did not realize that they were pregnant because they had irregular menstruation patterns since the onset of puberty. One example was a 14-year-old school student who participated in a school camp and overnighted at the school. Someone she did not know raped her that night. She did not realize that it could lead to pregnancy. At the sixth month of pregnancy, her body had changed and was larger. Her mother asked her why she was getting bigger and took her to the clinic.

"My mother asked 'why I am so fat'. Then, she took me to a clinic. The results showed that I was in the sixth month of pregnancy" (Noi, single, 14 years old, school student).

# Feelings and Concerns of Women Facing Unexpected Pregnancy

The results of the focus group discussions and in-depth interviews of the samples revealed that most of them experienced anxiety immediately after they suspected that they were pregnant. They felt concerned because premarital sex, sex while studying, and being abandoned by one's partner, were socially unacceptable. Many women revealed that people would look down on them and their parents. Thus, they were afraid that their parents were angry and disappointed in them. Some women disclosed that they cared about their parents' concerns the most. However, if their parents accepted their pregnancy they would feel relieved and calmer. Other minor issues included acceptance by their relatives, friends and people in the community.

# Self-Defining Pregnancy

After the women had passed the "shock period", they were thinking about the pregnancy and trying to define it. Although they were all in the same situation, with unplanned pregnancies, there were variations in the degree of pregnancy acceptance, which was due to differences in the individuals, partners, parents/relatives, peers, and communities. As described earlier, "Thong mai thong karn", "Thong mai thang jai", "Thong mai prom", or "Thong mai kadkit" meant unwanted, unintended, unplanned, and unexpected, respectively, as the women mentioned during the FGDs and in-depth interviews. The results revealed that all of the women reflected on being physically and psychologically unprepared for pregnancy, because most of them had an ideal husband and/or family. The main reason was that, in Thai culture, premarital sex, or pregnancy before marriage, are unacceptable. Moreover, pregnancy without a responsible man is a shameful situation for a woman and her family. On the other hand, social norms put the pregnant woman into the role of a mother who should be responsible for the baby in the womb. Thus, when the pregnancy situation did not ensue as they expected, most of them intended to terminate their pregnancy. They did not want to keep the baby because most of them were afraid that their parents and people in the community would know that they had had premarital sex with a consequent pregnancy, to avoid their disappointment.

# 3.3.2. Consulting Popular Sector (Other People)

Most of the unplanned pregnancies were caused by the women's partners, e.g., partner leaving them, having another woman, or using drugs. Thus, most of the participants would consult friends at school or at work, or relatives they trusted. They would select the person who could make them feel better or give them some advice. They preferred to consult people with similar experiences who were older or in the same age group, because they could understand their situation easily.

"I was shaking. Then, I ran to see my friend immediately (after knowing of the pregnancy result)" (Rat, single, 22 years old, college student).

Some women who lived alone tended to make decisions by themselves. Many women revealed that they felt stressed because they could not let anyone know about their situation. After they disclosed to someone they trusted, they felt better. To release tension, some of the women who wanted to keep it secret moved away to live in a new place, to avoid questions from close friends or people in the community before their belly started to get big. Some of them moved during the first or second trimester because their bellies were not too big and it was difficult to see from outside. The women who had no one to support them would try to seek a safe place, at their best friend's house or in a public shelter, rather than live alone outside during the last trimester of pregnancy, because they felt worried about the possibility of emergency labor.

"I did not want to tell anyone, so I went to live in Bangkok and worked at Klong Thom" (Lee, single, 21 years old).

Among the single women, some revealed that they did not like to tell their parents about the pregnancy because they were afraid that they would disappoint them. Hence, their parents knew their pregnancy because their physical was getting bigger.

"My mother saw my belly. At that time, it was getting big, at seven months of pregnancy. Then, she asked me whether I had menstruation. I told her that I did not. So, she went to buy a pregnancy test" (Pia, single, 14 years old, school student). The married women preferred to consult their parents. If the relationships of the participants and the parents were more close and friendly, they tended to consult them. However, some married women revealed that they did not want to tell their parents because they already had many problems and they did not want to bother them. However, if their parents asked, they would only release some information, not all the problems.

"My mother, I did not want to consult her because she already had many problems" (Joy, married, 24 years old, factory worker).

"If my parents asked, I would tell them part of the information (not all of it)" (Air, married, 21 years old, factory worker).

After the women gained support and information from the person they consulted, it was surprising that most of them made decisions by themselves. A few cases followed their parents' decision.

"I did not consult anyone, I made decision by myself (put the baby for adoption)" (Porn, single, 18 years old).

"At that time I wanted to terminate pregnancy. My mother made decision (put the baby for adoption) because I did not know the man who made me pregnant" (Noi, single, school student, 14 years old).

#### 3.3.3. Choosing Options

After the women with unplanned pregnancies gained information and recommendations from consulting others, there were two decision-making options, terminating the pregnancy and continuing the pregnancy. The majority needed to terminate the pregnancy, while a few cases needed to continue the pregnancy. In making the decision, both options were painful for the young people. The women who chose to terminate their pregnancy faced self-conflict and other external factors, while the women who continued their pregnancy were insecure because they were unsure how to cope with present and future situations. It was important to note that the young women would change their minds, which depended on their partners and their parents. If these significant people supported them, they tended to keep the baby to term. But if they did not care, or showed no responsibility, the young women tended to terminate the pregnancy. The factors that influenced the choices of the young people are explained below.

**1. Society and community**. Most of the unplanned pregnancies were caused by premarital sex, while studying, or with abandonment by the partner. As a result, they all felt ashamed to have a belly that kept getting bigger. They were afraid that people in the community would gossip and reproach them that they were bad or promiscuous girls. In addition, their family members would be blamed. One example was a young woman named Pae, who was 16 years old. Her parents divorced when she was young. Her mother left her with her aunt, while she was working as a dressmaker in Brunei. After she finished grade 9, she quit school. She liked going out with her friends, and one night at the discotheque, she was drunk and taken away by men she did not know before. After she woke up, she found

herself alone at a motel, and she realized that she had been raped. She kept it secret until the fourth month, when her aunt asked why she was getting bigger. Her aunt took her to a clinic and found that she was pregnant. She wanted to terminate the pregnancy but the doctor at the private clinic could not do as requested because her pregnancy term was over three months, so that if she wanted to do it, it was risky. So, she went to another clinic, but it was too expensive that she could not afford it.

"I was in the fourth month of pregnancy. I was very stressed but I could eat normally. I did not want it (the baby) but I had no way out. I needed to find a new place to hide myself (to avoid gossiping by the neighbors). My neighbors' gossiped that 'I did not study but had a sexual partner instead'. I did not like them to show contempt for my mother and my family members" (Pae, single, 16 years old, out-of-school student).

The women disliked gossip because it would spread by word of mouth, with the addition of the attitudes of the ones who hastened the news. Nowadays, the relationships among people in the community were bonded weakly. Premarital sex was a subject that attacked cultural morals and norms, and therefore was considered a good subject for gossip among community members. However, if the relationships of the women with their neighbors were strong, the gossip could be very useful, because their neighbors would support them and make them feel secure and dare to disclose their troubles. The women who had support from their neighbors tended to keep their babies to term. But the situation where women gained support from their neighbors was rare. However, this is only one of several factors that affected the women's decision.

"My neighbors did not repeat my faults; on the contrary, they supported me. They know my background and understand me. They pity me" (Fon, single, 19 years old).

Another reason that the people tended to support the women is that, at present, the situation of premarital sex is seen more often than in the past, people tend to accept it more and seem to understand the women situation. Consequently, it is important to note that increasing numbers of people view premarital sex as normal in the current situation.

"My neighbor knew. She did not repeat my fault, instead, she said 'it is not serious. It has already happened" (Tuk, single, 21 years old, vocational school student).

"She was sympathetic. She said, 'at present the situation is different from the past when people would repeat your faults. Currently, there are many women like you (having premarital sex and pregnant)" (Tuk, single, 21 years old, vocational school student).

**2. Family members.** If the young women lived with their parents, the parents were the most influential on the young women's choice of terminating the pregnancy or keeping the baby. Most of the parents wanted their daughters to terminate the pregnancy if they had had premarital sex; the women themselves did not want to burden their parents, as well. They felt that they could not support themselves and had no income. For the women who lived with their parents, the mother was the person who played the important role in the women's decision-making. They were the ones who took their daughters to terminate the pregnancy,

because they did not want their daughters or the families to lose social status. One example was Brew, a 16-year-old school student who was living with her parents. She liked going out, and got pregnant with her boyfriend. After the pregnancy was recognized, the boyfriend disappeared. At nearly fourth months of pregnancy, her mother knew that she was pregnant. She took her to many clinics but the providers refused to perform an abortion for her because her pregnancy term was more than three months, and it would be risky. She was referred from the clinic to a shelter, to avoid gossip. After staying at the shelter, her parents decided to let her keep the baby, with their support, so that she could return to school without having to worry.

"When I knew I was pregnant, I only wanted to terminate the pregnancy. I did not want to be a burden. My father and mother are getting old" (father 50; mother 45) (Brew, single, 16 years old, school student).

**3. Partner.** For the women who lived away from parents or lived with their partner, the partner had the most influence on the women's decision whether to terminate the pregnancy or keep the baby. Most of the women with unplanned pregnancies wanted to terminate the pregnancy because of their partners. Many of them were faced with irresponsible men. The reasons for which their partners were influential included disappearing after having sex, using drugs, beating the women, or having another woman. One example was "On", a young woman of 20 years. She ran away from home and lived with her boyfriend. After living together for a while, her boyfriend started to use drugs and did not take responsibility for anything. She worked alone as a waitress in a restaurant to earn income to spend on the expenses of daily life. Her partner took his friends to their room and used drugs. Sometimes, he disappeared for 2-3 days after he got some money from her. After a while, she was using drugs, as well. Without using any contraceptive method, she was pregnant and tried to terminate the pregnancy by using abortifacient products, but it was not successful. One day, policemen came to their room, and all of them were arrested, but because she was pregnant, the policemen referred her to a shelter.

"He was changed after using drugs. He brought his friends to our room. Sometimes, he disappeared for 2-3 days after he got money. One day, I also was arrested after I came back from buying food" (On, single, 20 years old, waitress at a restaurant,).

Some of the young women expected their partners to come back. After their partner left them during pregnancy, some of them waited and expected their partner to come back, so they kept the baby to term waiting and hoping that one-day their partner would come back and responsible for the baby future.

"...before I needed his love, understand, and responsible. But now, the most important is responsible. I do not want him to come back and live with me. I want him to responsible for the baby. This is the only thing that I need from him" (Koi, single, 24 years old).

**4. Friends.** During the adolescent period, the young people tended to follow their peers. The majority of them spent more of their time among friends than with their parents. Once they faced an unplanned pregnancy, they tended to consult their friends. The majority of the

young people tended to terminate their pregnancy because they wanted to avoid follow-on problems, such as quitting school, disappointed parents and relatives, no income to support the baby, being laid off, and lack of acceptance by society and the community. Most of the young people would follow their friends' advice. Moreover, some of them had friends who had experience of unplanned pregnancy and used to terminate the pregnancy. For these reasons, they tended to terminate the pregnancy.

"While I was a student, I found many of my friends were pregnant. Sometimes, I went with them. One of my friends was pregnant but her parents did not like her partner. So, they went to an illegal abortion clinic. They used suction as an abortion method" (Koi, single, 24 years old,).

"My friend said 'if he's no good, then get an abortion'; they told me like it was a normal event" (Nu, single, 20 years old, factory worker).

**5.** Women's situations. Some women were not in crisis situations when they were pregnant, including being a student, workplace policy prohibiting pregnancy, or having a new baby too soon following the previous one. These situations are explained as follows:

*Student status.* In the regular primary-to-high school system, married or other students are not allowed to have a baby. Any woman who fell pregnant while studying was perceived as promiscuous and would be asked to drop out of school. Consequently, young women in this situation would terminate their pregnancy. Some, who could not terminate their pregnancy, would hide themselves and/or quit from school to avoid gossip from others in the community.

"If I keep the baby, first, I will stop studying. Second, how do I avoid disappointing my parents? If I raise the baby, I will quit school. The better way is to terminate the pregnancy. If I go to the clinic on Friday, I can rest on Saturday and Sunday" (Oil, single, 20 years old, vocational school student).

"I wanted to terminate the pregnancy for sure. If I kept the baby, my mother would suffer dishonor. The people in the community would look down us because I was a student" (Jaw, married, 24 years old, housewife).

Quit job or laid off due to pregnancy. In many situations, the women would quit their job because of premarital sex. They would quit the job by themselves because they wanted to avoid gossip. In addition, in some workplaces, they would lay off any worker who fell pregnant, because that was their policy. Moreover, the characteristics of some occupations, such as standing all day long, were not suitable for pregnant women. In some workplaces, the policy was that they would not accept any pregnant women. If a woman was pregnant, she had to resign from the job.

One example was "Nid", who was 23 years old. She came to work in a factory in Bangkok when she was 15 years old, and had her first partner at the same age. Three years later, she had a child with him and lived with her mother-in-law up-country. She moved to work in a new textile factory, because she did not live with the first partner regularly, and later there was another man who was fond of her and became her second partner. Both partners knew about their love affairs with "Nid". Their love affairs were smooth, until one

day "Nid" fell pregnant again. At the second pregnancy, she could not identify who the father was. When her belly was growing large, she quit the job because it was a regulation of the factory not to hire pregnant women. At the second pregnancy, her partners did not take any responsibility and left her alone. She was very depressed and wanted to terminate her pregnancy, but she could not afford the high cost at the fifth month of pregnancy. One of her friends recommended a shelter, so she decided to stay there.

*Having the baby too soon.* Some of the women did not use any contraceptive methods because they thought that the period a few months after delivery was safe. In some women, their fertility resumed very quickly, so that instead of menstruating, they fell pregnant again. This was a stressful situation for most of the women, especially the low-income women, because they needed to consider trying to get more income to save the family. If it was not possible to earn more income, they preferred to terminate the pregnancy, to forestall the problem.

One example was Jum, who was 24 years old. She had married about 3 years previously, and had a one-year-old boy, and a few-months-old baby. She told me that her husband was using drugs and did not take responsibility for the family. She used to run away from him, but her parents asked her to come back because of the children. When she knew of the second pregnancy, she tried to terminate it using abortifacient products, but it was not successful. She had no choice, only to keep the baby to term and raise it without knowing the future. With the economic crisis, their neighbors also looked down on her family because they were poor.

"They said, 'the elder one was still young, and then it is followed by a new pregnancy; one baby grasped in a hand, the elder one walking beside, and another one in the belly'. They talked like I was a promiscuous girl" (Jum, married, 24 years old, small food vender).

#### 6. Women's Experiences

*Having unplanned pregnancy or abortion experiences.* Some of the study participants had had experience of abortion. They tended to terminate the pregnancy because they knew the place and the procedure. They did not panic like the inexperienced ones.

*Inexperience of sex and pregnancy.* Some of the young women did not realize that having sexual intercourse or having sexual intercourse only once could cause pregnancy.

One example was Noi, who was a 14-year-old school student. She participated in a school camp and overnighted at the school. Someone that she did not know raped her that night. She did not realize that it could lead to pregnancy. At the sixth month of pregnancy, her body had changed, and was becoming bigger. Her mother asked her why she was getting bigger, and took her to several clinics for an abortion. All of the providers refused, and one clinic referred her to a shelter. They wanted to terminate the pregnancy but the pregnancy term exceeded the abortion criteria. Hence, they had no choice but to keep the baby to term and delivery. After delivery, they would put the baby up for adoption.

*Rumors and misperceptions.* Since abortion information was not openly disclosed to public, the women needed to seek information themselves. Some of them gained information by word of mouth. Some information was full of misperceptions about abortion, such as that it was a lethal procedure. One of the young women told the researcher that she heard from her friend that once a woman entered the abortion clinic, the provider would give her some kind

of medicine. After she took the medicine, she would feel dizzy and lose consciousness. If the abortion was complete, but the woman still felt dizzy, the provider would take her into a field and leave her there. If she survived, it was only by good luck. If she was bleeding, she might die without anyone knowing or caring. She was told that because abortion was illegal, the providers were afraid of being caught.

"She told me that she used to go there. The provider gave her some medicine. After she took the medicine, she felt dizzy. After the abortion procedure was over, if she were awake and conscious, she would survive...but if she was unconscious or felt dizzy, the providers would take her into a field and leave her there. If she was bleeding, she would die" (Maew, married, 18 years old, housewife).

**7. Access to Information**. Since abortion is illegal in Thailand, women who wanted to access safe abortion places needed to search for the information themselves. They knew the places by word of mouth, but without any evidence to prove that the clinic they visited provided safe abortions. Some of them took a long time to search for abortion information. When they visited the clinic, the provider could not provide the service requested because the pregnancy term exceeded the medical criteria. For this reason, the women kept their babies to term with no choice.

Example of Rat, 22 years old, 3<sup>rd</sup> years college student, she was realized of pregnancy when it was three months of pregnancy. After, the pregnancy was confirmed, she went to drugstore and asked for menstruation inducers, and then the seller asked whether she was pregnant. She told him that she was pregnant with her boyfriend. Then he gave her the medicines to take 2 tablets two times a day. After taking the medicines from drugstore, there was noting happen. So, she sought more information towards abortion and abortifacient products from her friends. She tried several regimens by asking her friends to buy for her. At the fourth month of pregnancy, she realized that the abortifacient products would not help her. Then, at the fifth months of pregnancy she searched for an abortion clinic, when she visited the clinic, it was closed. She decided to keep the baby to term

Some of them, after failing in their visits to the clinics, tried to terminate the pregnancy by themselves, using various abortifacient products, and/or other methods, e.g. massage, or beating the belly. They did not know or realize that there were clinics that could provide abortion services even when the pregnancy term was greater than three months. Moreover, a few cases learned that there were shelters available for pregnant women.

**8.** Affordable. The cost of an abortion was the major concern for the women with unplanned pregnancies. After they had obtained information and knew the cost of an abortion, some of them took time to save and borrow money from people they trusted to pay for it. When they visited the clinic with the money, the provider could not provide the requested service because the pregnancy term exceeded the medical criteria. Some of the women tried hard to save and borrow the money but it was not successful. They could not get enough money to satisfy the fee requested by the clinic. Some clinics requested for more than 10,000 Baht for cases exceeding three months' pregnancy. Many of them, after failing to have an abortion because of the high cost, just kept the baby to term even though they were not ready to have it. Some of them decided to put the baby up for adoption after delivery, whereas some of them raised the baby alone.

"The provider said 'If I want to terminate the pregnancy, it will cost 12,000 Baht'(I was at the fifth month of pregnancy)" (Joy, single, 24 years old, factory worker).

Some women could not afford the high cost of a safe abortion, or even the low cost of an unsafe abortion. One example was a married woman with her partner, who had one six-yearold child. Later, they moved to Bangkok and worked in a gasoline station. Her husband worked as a cashier, while she worked in a small supermarket in the gasoline station. One night, her husband left her, taking all the money in the cash register, which was more than ten thousand Baht. The gasoline station owner asked her to take responsibility for what her husband did, but she could not. So, she quit the job. At that time, she had been pregnant for 4 months. After her husband left, she tried to find abortion services. Her neighbors at her hometown in Isan (north-eastern Thailand) recommended her to an illegal unsafe abortion clinic, where an old lady performed the abortions. When she visited the place, the old woman examined her belly and said that the baby was already formed. She requested 3,000 Baht for the abortion. However, the woman could not afford it because she only had 500 Baht. With such a small amount of money, the provider refused her request. She returned to Bangkok and her friend recommended her to a shelter.

"My friend took me to a house located in a remote area upcountry. I only had 500 Baht. When I arrived, an old lady examined my belly. She said 'the baby was already formed'...I told her that I had 500 Baht. She said that if I gave her 3,000 Baht (1 US\$ = 35 Baht), she would do it for me. I told her that I could only afford 500 Baht. Hence, she did not do as I requested" (Pen, married, 24 years old, unemployed).

#### 3.3.4. Compromise with Self-conflict and Finding a Rationale for Support

Not only the external and internal factors influenced the decision, as explained. In addition, the women also needed to prevail over internal self-conflicts towards terminating the pregnancy, which may be attributed to Thai norms and culture, in which women are taught to be caretakers for family members, and mothers. Society expects that any woman who falls pregnant will be a mother, without looking at their circumstances. Moreover, as Buddhists, many women have been taught that terminating pregnancy is sinful because it is the killing of an innocent life. With the negative consequences of keeping the baby to term, which were due to socio-economic and internal conflict problems, most of the women decided to terminate the pregnancy after weighing up the outcomes and the long-term effects on their lives, which would be those of mothers responsible for their babies' futures. However, some women decided to keep the baby to term. The following are the rationales, based on terminating the pregnancy and keeping the baby to term.

# 3.3.4.1. Rationale for Terminating the Pregnancy

Once the women decided to terminate the pregnancy, they would gain support from a person they trusted, get more relevant information, or interact with themselves to overcome their feelings towards terminating the pregnancy, which included 1) terminating a pregnancy was immoral, 2) terminating a pregnancy was risky for their life, and 3) terminating pregnancy was losing a loved one (please see figure 3.2).

Terminating a pregnancy was immoral. Most of the women understood that terminating a pregnancy was sinful, but it was more shameful to disclose premarital sex and pregnancy to

society. They compromised with self-conflict by saving their parents' status and making their future come true. Moreover, some women said that they terminated their pregnancy because they did not want to be a burden on their parents, because they already had younger brothers and sisters who needed support from their parents. However, many of them felt guilty after terminating their pregnancy because they realized that it was sinful. It was difficult to delete the pain from their memories after terminating the pregnancy.

"I felt it was sinful...sinful. I feel regretful up until now" (2 years after terminating the pregnancy) (Yam, married, 19 years old, a singer).

Many of them told the researcher that after they had successfully terminated the pregnancy, they would try their best to make their parents happy and to take care of them. After they had endured the crisis situation, they realized that their parents were passed through a hard time of taking care of them and they were the only ones who were the most sincere. This was one of the ways the women thought, to make them feel better about their sinful action.

"After the providers completed the process of abortion, I felt relieved. I am not a burden on my parents. My mother was the greatest, she really helped me" (mother took her to the abortion clinic) (Nu, single, 20 years old, factory worker).

Some of them felt that terminating the pregnancy was better than raising the baby without any future, and that the baby would feel bad because of growing up without a father like the other children. So, they made the decision that terminating the pregnancy was better than keeping the baby. One of a vocational school student, who was failing from terminating pregnancy, told that:

"If I keep the baby to term and I want to continue my education, I may not have money to raise the baby. If I want to abandon the baby later, how do I do? Should I leave the baby under the public bridge like someone did and was posted on the front page of newspaper. I felt pity for the baby. If I terminate before it formed to be a baby, it is better to leave until facing problem" (Aui, single, vocational student, 19 years old).

Terminating the pregnancy was risky to their lives. Many of the young women realized that the termination of a pregnancy was a risky process, because they would bleed, and it was very painful. Some of them heard, by word of mouth, that it could cause death because of the bleeding. However, they felt strongly about terminating the pregnancy, without being afraid of what would happen after that.

"In my heart, I did not want to terminate the pregnancy because I was scared of the bleeding and the pain. When I arrived at the clinic, I needed to do it. If I kept the baby, my family and I would be in a difficult situation. Moreover, the child would grow up without a father" (Nu, single, 20 years old, factory worker).

Some of the women felt insecure after they went to the clinic, because the setting and environment of the clinic scared them. Some of the young women said that the way to the abortion room was complicated. Moreover, the staff asked them to provide a signature to approve the provider providing treatment for the uterus instead of an abortion. This process made the women feel unhappy, because they felt that the clinic staff did not behave honestly.

"I felt insecure after the abortion process was completed. The physical appearance of the room was complicated, but the medical equipment was clean. However, I thought that they do not feel responsible for our safety because they asked us to provide a signature to approve uterine treatment instead of an abortion" (Nu, single, 20 years old, factory worker).

Termination was losing a loved one. More than half of the young women had had their first experience of pregnancy. Consequently, at the beginning of pregnancy they did not care much about the baby in the womb compare with the women who had experience of pregnancy. The longer the pregnancy period became, the more attached they felt to the baby. They needed to terminate the pregnancy because of their partner. They were irresponsible, disappearing after knowing about the pregnancy, having another woman, or forcing the woman to have an abortion.

"He begins scolding, beating me. Then, he took me to an abortion clinic. I did not want to go but he forced me to. If I did not go, he would beat me" (Yam, married, 19 years old, singer).

## 3.3.4.2. Keeping the Baby to Term

A few cases among the women with unplanned pregnancies decided to keep the baby to term without attempting to terminate pregnancy. The women who decided to keep the baby to term would weigh the positives and negatives of pregnancy. They tried to adjust themselves to appreciate keeping the baby, by thinking about morality, the mother's role, and the health risks of abortion (please see figure 3.2).

The moral issue was mentioned the most by the women who had decided to keep the baby to term. The longer the pregnancy period, the more the women could adjust their attitudes, because the baby could react to them as the pregnancy term increased. So, they felt attached to the baby, especially the women who lived in the shelters, because they had a chance to care for other people's babies, and thus realized how hard a time their mothers had endured raising them. In addition, they were among other women who were in the same situation. They compared and shared their experiences with others, which made them rethink, based more on logic than the emotions.

"We made it (the baby). It does not know anything that we did. We did it wrong and then we are going to kill it. It was wrong to kill even it was only a bloody shape, because the blood was formed from us. We made a life. If we put it up for adoption, it would be OK. But if we kill it, it is sinful" (Gay, single, 20 years old, out-of-school student).

Some of them did not want to take the risk of terminating the pregnancy because they were afraid of bleeding, which would cause death. Moreover, if there were something wrong, the provider(s) might not take responsibility for their lives.

"Maybe the clinic might not take responsibility for my life, if I was bleeding. When I decided to get an abortion, I did not know what would happen to me" (Ple, single, 16 years old, out-of-school student).

Some of the women did not want to lose the baby even they did not plan to have the baby at the beginning. They lived with their partner happily, but one day their partner left them. They tried to survive and keep the baby to term. After the baby had delivered, the bonding between the mother and her baby became stronger.

"I felt proud of doing the right thing in my life (keeping the baby). Life should be born and grow up. The baby did not do anything wrong. I am happy to have her because I have a friend. I feel happy... I cannot explain to you" (Dao, single, 24 years old, office employee).

Moreover, it was found that the majority of the women who did not put any effort into terminating the pregnancy lived or worked with a religious group. These religious groups provided counseling and information for the women. In addition, they referred the women with unplanned pregnancies to a shelter. However, most of the women did not need to keep the baby to term or raise the baby. Yet they could not overcome the feeling of sinfulness within the religious environment. Thus, some of them put the baby up for adoption after they learned that there was a choice of adoption available.

"I knew sister Vienna. She's one of my relatives. She always supported me, provided counseling, and gave my baby's name...Another sister came to talk to me. Then, she referred me to a shelter" (Kai, single, 20 years old, caretaker for the elderly).

Premarital sex, or pregnancy without a responsible man, represented a shameful situation, and because of this, women would seek other ways to relieve their stress. The following are explanations of the young women's reactions towards keeping the baby to term: 1) seeking a place to hiding during the pregnancy, and 2) putting the baby up for adoption after delivery.

Seeking a place to hide. Most of the young women realized that their pregnancies were not consistent with social norms and culture. Consequently, if they were to keep the baby to term they needed to seek a place to hide themselves or go away and stay at a place where nobody knew them, because their belly was getting big. This reaction was to avoid gossip by the people they knew.

For the women whose lived or worked with the religious group after they knew about their unplanned pregnancy, they would help the women find a shelter to hide their pregnancy, because they did not want the women to terminate pregnancy.

In this study, it was found that the women knew about the shelter(s) from reading the magazine named "Cheewit Thongsue", news stories on television, and from their friends and relatives by word of mouth.

"I decided to visit Ante-natal Care clinic at one hospital in Bangkok (pregnancy term was 5<sup>th</sup> months). I told a provider that I separated from my husband. Then the provider asked me to meet with the social worker. The social worker recommended me to come to Emergency house at Sukothai (Sukhothai Road). At that time, I thought it was located at Sukhothai province" (Lee, married, 21 years old).

"I kept searching information about abortion and pregnancy from various magazine and found information of Emergency house at Dong Muang from Cheewit Thongsue, page 59" (Lee, married, 21 years old).

Put the baby up for adoption. Many of the young women who continued their pregnancies felt more relaxed and comfortable while they stayed at the shelter, because they were among women who had the same problems and could share and tell their stories to each other. In addition, they had more time to reconsider the choices. Since many of them were abandoned by their partners, and some of them were raped, the women tended to put their babies up for adoption after delivery. The women who decided to put their babies up for adoption would feel free after they left the shelter, but some of them missed the baby because they had a chance to raise and feed the baby before departure.

"I definitely put the baby for adoption. If I have money like others, I may not want to do so. Moreover, I have parents who are getting old and still need my support. Also, my family is poor. If I raise the baby by myself, it might face a difficult life. Even, it is clever, but I can support only up to grade 10 or 12. If he/she lives with the adopter, she/he might have a better life" (Lee, married, 21 years old)

Actions. The help- or health-seeking patterns of the young women with unplanned pregnancies are presented in section 3.4 in more detail. The details cover the explanations of the women who decided to keep their babies to term and the women who terminated their pregnancies.

# 3.4. Actions: Help- or Health-Seeking Patterns of the Young Women with Unplanned Pregnancies

In Thai culture, pre-marital sex among young women is stigmatized. Thus, most of the women delayed making a decision about whom to consult to solve their problems and many of them sought help from other people (popular sector) instead of the formal healthcare system, because most of the young people perceived that the formal healthcare sector only provided physical care. When they confronted an unplanned pregnancy, which was not a physical illness, they did not visit healthcare practitioners. Once, they had made the decision to terminate the pregnancy, they sought services from folk or professional sectors when they needed more advanced services, after they had gained information from other people, in order to solve the problem. In this section, the data focus on how and why people choose a particular sector and its patterns.

The results derived are from an analysis of 45 young women with unplanned pregnancies, from both shelters and communities, using an in-depth interview technique. It was found that when the pregnancy was confirmed, most of the women would consult the "popular sector", which included their partners, friends, parents, and relatives. After they gained support and information, most of them visited drugstores for self-medication. A few cases sought help

from traditional healers for abortifacient products. About one third of the women consulted the professional sector about terminating the pregnancy. A few cases managed by using their hands to beat the womb, in order to force the baby out. Some women did not put any effort into terminating the pregnancy.

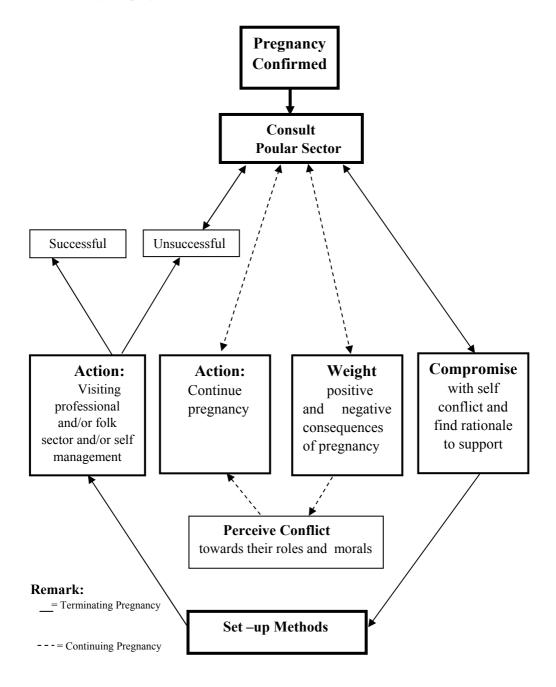


Figure 3.2. Decision Making Process among the Young Women with Unplanned Pregnancies.

Some of the women attempted using various ways to terminate the pregnancy, and if it were successful, the process of seeking help would end at that point. However, it was found that the majority of the women were not successful as planned after attempting to terminate the pregnancy. Thus, some of them would stop all efforts and continue the pregnancy to term because they had no choice. Some of them would stop because they felt guilty about hurting the baby. However, the majority of them would try other ways to solve the problem. Many of them went to private clinics, some visited drugstores/grocery stores again, to try as hard as they could to terminate the pregnancy (please see figure 3.3).

#### Help –or Health Seeking Patterns Model

The model of health-seeking behavior of the young women with unplanned pregnancies can be explained by using the most effort, and then explaining the patterns by which the women made decisions to solve their problems. There were four different patterns, as follows:

# 3.4.1. Did not Put any Effort into Terminating the Pregnancy

Once the pregnancy had been confirmed, the women in this group would consult their partner, friends, relatives, or parents, to gain information and support. All of them accessed a shelter or a place to hide their pregnancy because they got information from the people they consulted. To avoid gossip and embarrassment, the women in this group would stay in the shelter or a new place. After delivery, the majority of them raised the baby by themselves. Some of them put the baby up for adoption because of partner abandonment, study, or financial problems.

When comparing the women who consulted the popular sector (partner, friends, neighbors, relatives, or parents) and the women who did not consult the popular sector at the onset of knowing about the pregnancy, it was found that the women who consulted the popular sector were more ready to raise the baby by themselves.

# Examples of Women with Unplanned Pregnancy who did not Put any Effort into Terminating the Pregnancy

Miss Gay (1) was 20 years old. She was an orphan because her parents divorced when she was about 10 years old. Her mother took her to an orphanage at Pattaya, which belonged to the Catholic Church. She never returned to visit her. She was referred to Bangkok for further study supported by the Srichumpaban Unit, which was a nuns' group. While she was studying in grade 12; she met a man who became her boyfriend later. He was a final year vocational school student. After a few months they had sex and she fell pregnant. At the onset of pregnancy, she did not realize that she was pregnant, even though she felt nauseous and liked eating sour fruits. Moreover, missing her menstruation did not make her suspect pregnancy because she had had irregular menstruation periods from the onset of puberty. When the signs and symptoms were more obvious; she decided to visit a private clinic with a friend. At the clinic, the urine pregnancy test showed positive. She was in shock and dared not to tell the sisters, who were her caretakers. Some of her friends recommended an abortion, but her best friend told her not to do it because it was a sin. She followed her close friend's advice. She told and discussed her pregnancy with her boyfriend, and moved to live with him to gain a sense of family and get support from him. When she had lived with him for 2 months, she learned that he had another girlfriend before her. They started quarreling and she was

seriously beaten. She returned in deep depression and stayed at a dormitory belonging to the Catholic Church.

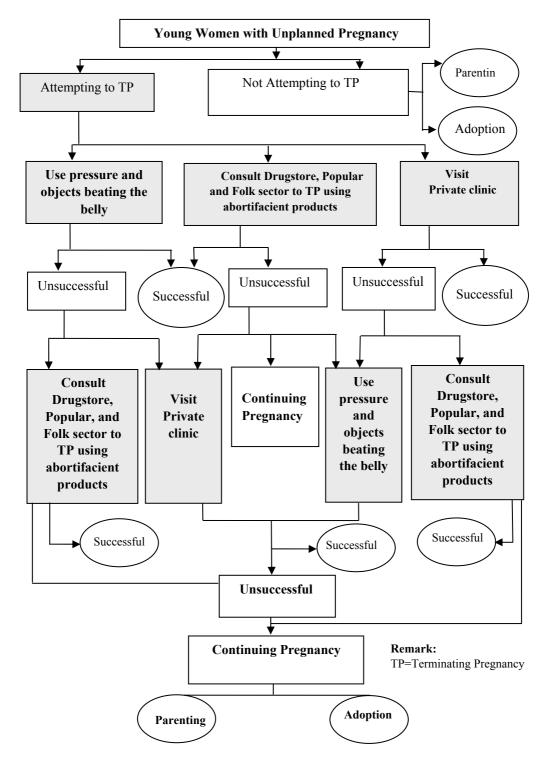


Figure 3.3. Patterns of Help- or Health-Seeking among Young Women with Unplanned Pregnancy.

She planned to keep the baby to term and raise the baby herself because she did not want the baby to be an orphan like she was. Moreover, because she was a Christian, she believed that one life was valuable. She quit school and started a new class at a vocational training center that belonged to the Catholic Church. After completing the course, she intended to find a job to earn income to raise her baby.

Miss Kai (2) was 20 years old. Her mother had died when she was young. Her father, who was a public driver, had remarried. Kai lived with her step-mother and her father until she graduated in grade 12; her mother-in-law hit and beat her quite often. Before she left home to find a job, she was hit and was hurt on the head with a big piece of wood. She got a new job as a caretaker at a Catholic house for ageing people. A few months after working at the house she met a construction worker and they fell in love with each other. After two months, she unknowingly fell pregnant, until the sister who was in charge of the house asked her to visit a clinic for testing. She suspected that Kai might be pregnant because her belly was getting big. At the clinic, it was found that she was 4 months' pregnant. She was in shock. She decided to tell her boyfriend about the pregnancy. He showed no responsibility for the outcome and told her that he had a wife. Kai was very stressed and depressed. After the sister knew about her situation, she referred Kai to a shelter that also belonged to the Catholic Church.

Kai did not want to terminate the pregnancy because she received counseling from the Catholic sisters. She planned to put the baby up for adoption after delivery because she could not support or raise the baby. Without support from her parents, she could survive on a salary of only 2,500 Baht per month, and it was difficult for her to raise another baby.

#### 3.4.2. Put Effort to Terminate Pregnancy

There were 3 patterns of help-or health seeking patterns of the young women with unplanned pregnancy who put effort to terminate pregnancy, which were explained as follows:

#### Terminating the Pregnancy Using Pressure and Objects to Beat the Belly

The women in this situation were both single and married. Most of them had a very negative attitude towards their partner because the partners had abandoned them and had not taken responsibility. However, none of the women dared consult anyone else at the beginning. Thus, they were stressed and tried to find some relief from the stress. When they were at their utmost level of stress, they would beat their belly to make the baby out from the womb. If it were successful the women would feel relieved. The women who could not solve the problem would seek further help through consultation, to relieve their stress. However, after beating their bodies, they felt guilty and afraid that the baby would be deformed if it did not come out. Some of them visited private clinics for terminating pregnancies but the pregnancy term exceeded the medical criteria, and some of them did not visit private clinics because they had no money. So, they tried using abortifacient products, but it was not successful. Once these efforts had failed, they tried to seek a place to hide their pregnancy, to avoid gossip. After the baby was born, if the partner was not responsible, the women tended to raise the baby up for adoption. However, if the partner showed responsibility, the women tended to raise the baby themselves.

#### **Examples of Women with Unplanned Pregnancies Beating the Belly**

Miss Koi (1) was 24 years old. She was the youngest of four brothers and sisters. Her father had died two years previously, and her mother had remarried. She met her boyfriend, who was a musician, while studying at a vocational school. Without the approval of her parents or relatives, she lived with him after graduation. When she fell pregnant, she decided to keep the baby because her partner promised to take responsibility for the family. While she was pregnant, her partner frequently came home late. One day, he left her to go to Bangkok, by claiming that there were more jobs in the city, but lost contact after that. At the sixth month of pregnancy, she followed him to Bangkok and found that he had another woman, who had just delivered a baby boy. She was furious and very depressed. At the seventh month of pregnancy, she beat herself whenever she felt depressed.

When the researcher met and interviewed her, she had delivered a one-month-old baby girl. She had reunited with her parents and her own family and her boyfriend's family gave her support. She decided to keep and raise the baby, and was living happily with the baby.

Miss Oam (2) was 18 years old and studying at a vocational school. Her parents divorced because her father had many wives. She could not communicate with her new mother-in-law. Thus, she decided to live alone at the dormitory with her father's support. She also worked as a waitress in a restaurant in the evening after classes ended, to earn more income to support her studies. Later, she met a man who was a customer in the restaurant; they became friends, and soon had sex. She realized that she was pregnant at the fourth month of pregnancy because her friend asked her to get a pregnancy test. The test result shocked her because it was positive. She did not expect that she would be pregnant because she only had sex with him once. She contacted him and told about the pregnancy, but he showed no responsibility and told her that he had a wife. She was very distressed. She said that, at that time, she hated the baby and tried to hit her belly using her hands and beer bottles to drive the baby out of her womb. Finally, she delivered a baby girl and put her up for adoption.

When the researcher met her, the baby was already living with her new parents. She told me that she felt very sorry when she saw others holding a baby. She cried and showed the researcher the baby's picture.

#### Consult Popular and Folk Sectors, and Drugstores for Abortifacient Products

The women in this group were both married and unmarried. Once the pregnancy was confirmed, they were unprepared to have it. So, many of them consulted their friends, relatives/parents, or their partner who was the most important. If their partners were prepared to be responsible for the baby and accept the women as his wives, the women did not terminate the pregnancy and raised the baby by themselves. However, it was found that their partners abandoned many of them. Thus, they tended to terminate their pregnancies. Since many of them had financial problems and could not afford the cost of an abortion at a clinic, they tended to manage by themselves using various abortifacient products.

The women tended to purchase abortifacient products from drugstores or grocery stores. Few of them sought abortifacient products from traditional healers. Less than half were successful, and the majority was not. The outcomes of their efforts varied greatly because of external factors, which were separate from the physical conditions of the women, including pregnancy term, abortifacient product regimen, and duration utilizing the product. Moreover, the psychological condition, the acceptance level of the baby and the relationship with the partner, were the main pressures causing the women to persist in doing as planned, or canceling the effort.

The women in this group tried as hard as they could to terminate the pregnancy with varying results, as mentioned above. After trying for a while, some of them were successful, but the majority of them were not. The women who failed to use abortifacient products successfully tended to visit private clinics, and/or hospitals to seek further services. As a last resort, some of them might beat the fetus or jump up and down on the floor, hoping that it would help to terminate the pregnancy. Once there was no choice, they decided to continue the pregnancy, and needed a place to hide themselves during the pregnancy, until delivery. After delivery, some of them decided to raise the baby by themselves, but some of them put the baby up for adoption.

However, some of the women who had information about shelters came to stay at a shelter immediately after they had failed to terminate the pregnancy using an abortifacient product. They got information about the shelters from their friends, relatives, neighbors, or printed material. The women who did not know about the shelters would persist in trying to terminate the pregnancy until they felt that there was no way left to help them.

It is important to note that the women who were seeking help/health in this pattern took a long time and used various abortifacient products regimens. Thus, when they realized that it was not successful, it was too late for them to try another solution. Also, it was found that the women who failed to terminate their pregnancies by using various abortifacient product regimens without the support of their partners, friends, relatives, or parents would put the baby up for adoption rather than raise the baby themselves. This was because of the financial aspect, which was the main problem. The women who did not have information of availability of shelters or adoption services just kept the baby to term without knowing the future. Some of them decided to give the baby to their relatives/parents as an adopter after delivery.

# Examples of Women with Unplanned Pregnancies who Consulted Popular, Folk, and Professional Sectors using Abortifacient Products

Miss Pung (1) was 18 years old. She had married about 3 years previously with parental approval. She could not use contraceptive pills due to the side effects and her husband did not want to use condoms. Thus, after 3 years of living with her husband, she already had her third pregnancy. The first pregnancy occurred a few months after living together. At the second month of that pregnancy, her husband became addicted to amphetamines, so she decided to terminate the pregnancy at a private clinic, with a successful result. A few months after the abortion, she had her second pregnancy. She decided to keep the baby to term and raise the baby by herself because her husband had quit amphetamine use and was back to normal. After the baby was 4 months old, she fell pregnant again. At the third pregnancy, she was worried and not ready to have the baby, because she already had one child to take care of and their family income was not stable. She had no money to buy milk powder, and the baby got only sweetened condensed milk, for which infant consumption was forbidden. With the family crisis situation, she decided to terminate the pregnancy by utilizing various abortifacient products. She went to drugstores and took 2-3 times higher doses of the products than the regular doses recommended by the providers to regulate menstruation. However, after she took the product it was not successful. Thus, she asked friends to buy the abortifacient product for her again, with the same outcome; the product could not push the fetus out. Finally, she went to see a traditional healer in the community for uterine massage but was refused.

When the researcher met her, it was her seventh month of pregnancy; she was very sad, her face showed worry and depression because she had no choices left. She still wanted to terminate the pregnancy; she asked me whether any places would agree to perform an abortion in the seventh month of pregnancy. If there were no choice, she said that she would give the baby to relatives, because she could not afford to raise another baby.

Miss Koi (2) was 21 years old. During her last years at vocational school, she ran away from home. She moved to live with her boyfriend who was also a student. She fell pregnant a few months after living together because her boyfriend did not use a condom and she was scared of using contraceptive pills because they made her get fat. So, her partner used withdrawal as a contraceptive method, but it failed. When she realized she was pregnant, it had only progressed two months. At that time, she quarreled with her boyfriend because he had another woman and he showed irresponsibility and was unconvinced that Koi was pregnant to him. Koi felt very sad and stressed, and decided to consult her friends who used to terminate pregnancies while studying. She went with her friend to an unsafe abortion place, which looked like a house. After her friend completed the process, she went to rest at her apartment with heavy bleeding for many days. She felt exhausted and looked pale. With that bad experience, Koi decided to use abortifacient products instead of the services at that house. She took one bottle of "Ya Satee" and one sachet of "Ya Tanjai" together. She felt dizzy and drunk, just like drinking alcohol. The next morning, she had heavy bleeding and a big bloody tissue mass came out that looked like a bees' nest. She told me that the big bloody tissue mass might be a baby, which had just formed. She felt scared when she saw the bloody tissue. It was difficult to make up her mind that she did not kill the baby, because she tried to compensate by thinking that it was just formed; it was just blood. She felt that terminating the pregnancy while the baby was just formed was better than terminating it when the blood had formed into a baby-like shape. After that horrible morning, she suffered from bleeding for another two months. She had heavy bleeding, which required her to use a sanitary pad every day. She was very weak and looked very pale. She went to see a doctor at a private clinic after one month of bleeding. She told the researcher that she was afraid to be blamed by the doctor if he examined her vagina, because he would know that she had undergone an unsafe abortion. However, the doctor did not blame her and he ordered iron and vitamin tablets for her to take every day.

When the researcher met her, it was nearly a year after she had the experience of terminating the unplanned pregnancy. At present, nobody in her family member was aware of her pregnancy because she kept it secret and there was no evidence, because the product of pre-marital sex had already come out and gone down the toilet.

#### • Visiting the Professional Sector using Modern Medicine

The women with unplanned pregnancies in this group were both married and unmarried. They could afford the cost of terminating a pregnancy at a private clinic. More than half were successful, while the remainder failed because the pregnancy term exceeded three months. If they wanted to terminate the pregnancy, the cost was very high, and they could not afford it. Some of them went to a clinic and changed their minds because they saw scenes of women bleeding after finishing the process and waiting to go home. Thus, they changed their minds from terminating the pregnancy to continuing the pregnancy, because of the consequences of terminating the pregnancy.

Some of the women who failed to visit a private clinic or hospital would visit drugstores or grocery stores to try again, utilizing abortifacient products, especially women who were afraid of the consequences of abortion. However, the majority of those who had experienced failure to terminate their pregnancy by visiting a private clinic or hospital, which was their first choice, kept trying.

Finally, most of the women in this group tended to raise their babies after delivery because they were more financially ready than the women who visited drugstores or grocery stores at the first attempt. The results implied that the women who visited a private clinic as their first choice were financially better off than the women who visited drugstores or grocery stores. Thus, when faced with unsuccessful termination of pregnancy, more tended to raise the baby than put it up for adoption.

# Examples of common types of unplanned pregnancy, visiting the professional sector to terminate the pregnancy:

Miss Yam (1) was a 19-year-old singer. Her parents worked in a massage parlor that belonged to her aunt. She became a singer after completing grade 8 because her friend asked her to quit school. By the age of 19 years, she had had three partners. At the age of 17 years, she met her first boyfriend. While living with the first boyfriend, she was beaten. A few months after living with him, she fell pregnant because she did not know about contraception. She was not ready to have a baby at all. However, she did not consider terminating the pregnancy. During her pregnancy, she was still beaten by her boyfriend. She decided to leave him and return to live with her parents. One day in the fourth month of her pregnancy, she lifted a Pepsi crate and there was heavy bleeding. She found out that it was a miscarriage after visiting a private clinic the next day. Some time later, she met her second boyfriend, and moved in to live with him. She fell pregnant again. Her second boyfriend forced her to terminate the pregnancy. He divulged that he had a wife with a 3-year-old child. He beat her and forced her to go to a clinic even though she did not want to. She could not refuse to give in to him because he beat her harder and harder. Finally, he took her to a private clinic when she was in the fourth month of pregnancy. The clinic requested 7,000 Baht but they only had 5,000 Baht. After her boyfriend negotiated, they paid 5,000 Baht. She felt very bad after completing the abortion with the second boyfriend because he showed no responsibility for her or the baby. Moreover, her parents did not like him, so they subsequently separated.

She told the researcher more about the scene at the clinic, which was still fresh in her memory even though it was nearly a year. The process began with waiting in the waiting area for registration. After that, she was called to change clothes upstairs for preparations to starting the procedure. After she had finished changing her clothes, she was asked to lie down on a bed. Her eyes were closed with big eye pads, so that she could not see anything. She was given an injection after she had lain down, without any conversation. She felt scared because she did not know what would happen next. She wanted to change her mind and go back home because her parents did not know that she would terminate the pregnancy. If they knew, they might be upset because they did not want her to do it, but she dared not to tell them because her partner forbade her. If she told her parents he would beat her. After the injection, she fell asleep a few minutes later. When she woke up, the procedure was completed. She had been asleep for about 2 hours. After she had rested for a while, they allowed her to go home without recommending anything.

When the researcher met her, she had the third partner, who was a mechanic. They were living happily together. However, she still felt bad for terminating the pregnancy. She believed that it was a sin to kill the life of a baby.

Miss Noi (2) was 14 years old. She studied in grade 8 in a school. She did not realize that she was pregnant until her mother asked, "Why are you getting fat?". She could not tell her mother what was wrong with her. Her mother decided to take her to a clinic for a check-up. They were in shock and crying when the doctor told them she was in her sixth month of pregnancy. She recalled one night in the past six months when she went camping at school, and a stranger came into her room and raped her. She kept it secret and did not divulge it to anyone. Moreover, she did not realize what the consequences would be. Her mother decided to tell her aunt and grandmother. They were all miserable. Her mother taught her that the baby was disgusting because it was not her child. Then she took her to many clinics to seek an abortion. All of the providers refused to terminate the pregnancy because her pregnancy term exceeded the medical criteria. At the last clinic, the staff recommended she stay at a shelter while she was pregnant. She felt better after staying at the shelter because she was ashamed to live in the community, and ashamed for her parents.

When the researcher met her, she was eight months pregnant. She had arrived at the shelter a few days beforehand. The researcher felt depressed, seeing a young, innocent student with a big belly. She told me that she did not see the man's face or know who he was. She and her parents decided to put the baby up for adoption after delivery.

Miss Brew (3) was 16 years old. She studied in grade 9. She lived with her parents; her mother was a housewife and her father was a guard. She liked to go out shopping with her friends. She had many friends, both boys and girls, who did the same. Later, she had sex with her boyfriend unintentionally. However, she realized that she was pregnant when she missed her menstruation for two months, because she had a regular menstruation cycle. She kept it secret from her parents, but told her close friends. Her friend warned her not to have an abortion and she agreed. At nearly fourth months of pregnancy, she had an abnormal discharge from her vagina, and told her mother about it, but did not tell her about missing her menstruation. Her mother took her to a clinic and found out that she was pregnant. She apologized to her mother and told her the truth. Her mother told her father, and instead of being angry, he asked her to keep the baby. They forgave and understood her situation. However, her mother did not want her to keep the baby because she wanted her to complete studying at school. She took her to an abortion clinic, but the providers refused because she was in her fourth month of pregnancy. They referred her to a shelter. Her mother came to visit her often at the shelter.

"After I could not terminate pregnancy (because the pregnancy term was more than 3 months). Then, my mother called BUG 1113 and talk to a social worker at a hospital. Then she recommended me a shelter. After that my mother took me to the shelter" (Brew, single, school student, 16 years old).

When the researcher met her, her mother was coming to see her. They decided to keep and raise the baby themselves. Brew told me that her mother would raise the baby for her while she went to school. Also, they would tell their neighbor that the baby was their nephew.

# 3.5. Women-Proviver Interaction

Information about the interaction of the women and the providers was derived from indepth interviews during Phase I of data collection with 45 young women who had experienced unplanned pregnancy. This information on reproductive health services for young people will aid understanding of the status of services available for young people, so that gap(s) can be identified, and the health system and personnel can be strengthened, to serve the needs of the target population. When the young women made their decisions, they had to interact with themselves to find the rationale to support their decision, especially the decision to terminate the pregnancy. The first choice for the majority of the young women with unplanned pregnancies was to terminate the pregnancy. Most of the women felt that the clinic that they visited was run by professionals, providing safe abortions, because of their friends', neighbor, relatives, and parents hearsay. Once it was not successful, the women would interact with themselves and others, in order to keep the baby. Some women decided to go to an Antenatal Care Clinic for a check-up and/or try to find a place to hide their pregnancy. Premarital sex or sex among students are not socially accepted, and therefore most young students try to keep it confidential. The larger the belly gets, the more stress they feel. The following illustrate the feelings and interactions when women visit the professional sector, including drugstores, abortion clinics, Antenatal Care Clinics, and shelters. In addition, a few cases visited the folk sector or traditional healers for abortifacient products.

# Interaction with Drugstore Personnel

Most of the women who purchased abortifacient products at drugstores had particular product(s) in mind, and therefore visited them in a short period of time. For this instance, they rarely had a chance to have any dialog with drugstore personnel. They felt that this service sector was convenient and comfortable for them just to walk in, unlike other health facilities where they met the providers and had to answer questions.

"At that time, I did not know about abortifcient products. Thus, I bought Ya Satree at the drugstore because it indicates that pregnant women should not use. Hence, I took 2-3 tablespoon at a time as indicated in the leaflet. But nothing had happen" (Lek, single, 23 years old)

Some of the respondents revealed that if they wanted to repeat another abortifacient regimen at the same drugstore, they would ask a friend to purchase the products for them because they were afraid that the drugstore personnel would be suspicious and ask them about the purpose of using them.

#### Interaction with Providers at the Abortion Clinic

Once the young women with unplanned pregnancies made the decision to terminate the pregnancy, they interacted with themselves and other persons they could trust to support the choice. If there was at least one supportive person, and they had enough money, they felt confident about terminating the pregnancy. Most of them thought that it was the best solution, because they would return to their prior status, without their parents or others knowing. Not every woman who decided to go to an abortion clinic to terminate her pregnancy was as

successful as planned, due to the clinic's conditions and the women themselves. Most of the respondents who visited an abortion clinic revealed that the providers performed abortion for women who were pregnant for less than three months. If the pregnancy were over three months, the providers would refuse or refer the client to a clinic that agreed to terminate pregnancies at any term.

"The doctor examined by belly, and asked me whether I had sex. I said yes, but only once. Then, he asked me to check for urine. It was found that I was pregnant (6 months). I told him that I was not ready to have a baby. He asked me to talk with my partner. I told him the truth that my partner did not know that I was pregnant. He warned me that it was dangerous to do abortion at this period because it was risky to my life. However, he recommended another 2 clinics but they refused to do so" (Oam, single, vocational school student, 18 years old).

However, the respondents revealed that it was very expensive, and depended on the month of pregnancy; it ranged from 8,000–20,000 Baht. With the very high cost, many respondents gave up and decided to keep the baby to term without knowing their future.

"The provider told me that they would do it (5 months of pregnancy) but it cost 20,000 baht. At that time I had only 10,000 baht. So, I asked the provider to give me a special rate. They said the last price would be 18,000 baht" (Oil, single, vocational school student, 20 years old).

Low-priced clinics were also not well accepted by the young people. One respondent revealed that her friend recommended a clinic that charged 2,000 Baht for terminating a three-month pregnancy. She hesitated to go there because she had heard that women suffered from bleeding after visiting the clinic. However, she went to the clinic with her friend, but gave up after seeing the clinic from outside, because she was unsure that the providers would provide a safe abortion.

#### A Difficult Situation for Young Women Walking Into an Abortion Clinic

However, walking into an abortion clinic was a crisis situation for many young women, especially students. They felt unsure about doing as planned because of their inexperience, but they needed to confirm what the abortion clinic really looked, even though they had some information about abortion from friends and various other sources. At first, they decided definitely to terminate the pregnancy, but many of the young people were uncertain about what would happen to them. The idea that it was sinful to kill the life of the baby also made the women distressed. Some of them made a new decision after visiting the clinic, because they were not sure that their lives would be safe.

One example was a respondent who was a pregnant vocational school student. When she realized her pregnancy, she spent a few days for gathering information on abortion and collecting money. She drove a motorcycle alone to a clinic recommended by her friend, and drove around the clinic three times before entering it. She was afraid of seeing anyone she knew while visiting the clinic. Moreover, she was not sure what would happen to her once she went in.

"I felt nervous. I drove the motorcycle around the clinic three times. The fourth time, when only a few people were passing by, I decided to stop the motorcycle and walk inside" (Oam, 18 years old, vocational school student).

Having entered the clinic, most of the young women felt shy and dared not to ask any questions. However, if the clinic provided a private place for history taking and counseling, the women would feel relaxed, warm, and comfortable enough to divulge their confidential information. If the situation were opposite, the women would feel bad and be more stressed.

"At first, I felt shy walking into the clinic. But I thought that there are others who have the same problem as me. This made me feel better. However, I was not alone, my mom went with me" (Brew, 16 years old, secondary school student).

"When I arrived at the clinic, I didn't know how to start. The staff at the clinic was not nice to me. She did not pay attention to me. She did not ask me about my problem or offer any service. I decided to talk to her first. That area was an open area. Luckily, there were no other people, otherwise I would not have dared to talk " (Oil, 20 years old, vocational school student).

# The Scene and the Cost of the Abortion Made the Women Change Their Minds

Some of them returned from the clinic after entering because they had seen women who had just had their pregnancies terminated and they looked exhausted. This reaction implies that the women did not get counseling so that they might not have had an opportunity to explore every dimension of abortion and the other choices that were available.

"I went to the clinic by myself. While I was waiting, after agreeing upon on the price of the abortion, I saw a few women with pale faces come out of the room after the procedure was complete. I felt bad; I was scared when I saw the blood. At that time, I was not sure about doing as planned. I told the service provider that I changed my mind. The provider said 'don't worry, it's up to you'. Then I went home and did not want to do it any more" (Porn, single, 18 years old, factory worker).

Mostly, the women returned from the clinic without terminating their pregnancy because of the high cost. If the pregnancy term exceeded three months, pregnancy termination was very expensive. Many women revealed that the clinic asked for more for than 10,000 Baht. Thus, some women returned to get more money, but some could not afford the high price and gave up and continued the pregnancy. (Please see section 4.3.3 under affordability, which provides more detail).

#### Women Need Counseling Before the Final Stage

Once the women met the providers, most of them needed counseling, and many women went to the clinic with the most common doubt being what would happen to them. Moreover, many women revealed that they had used abortifacient products or menstruation-regulation products during pregnancy and they were afraid that the products had affected the fetus. If the fetus had formed into a baby-like shape, most of the women would hesitate to do so. They felt guilty about doing it because it was like taking the life of their baby. Moreover, the Buddhists believed that taking lives was sinful, especially the lives of innocent babies who have done nothing wrong.

However, most of the women revealed the same story, i.e., they did not get any counseling from the providers and they dared not ask for it. However, a few clinics provided counseling when the pregnancy term exceeded three months, and the providers would refer them to other clinics or to shelters. The women were afraid to ask the providers questions addressing the issues of concern to them, because it was more difficult for them to start asking questions than to walk into the clinic. They were afraid that if they did ask, the providers might ask more questions that they did not want to answer. Hence, they were growing increasingly gloomy during the process of terminating the pregnancy. Without proper counseling before terminating the pregnancy, some women felt distressed, and this feeling would leave a black scar on their hearts forever.

"I felt afraid. I felt it was sinful. I wanted to get out of the bed when the provider started the process and go home" (Yam, married, 19 years old).

"After the injection, there was a pressure in the lower abdominal area. I felt that I might not survive and I was not sure that it would be successful" (Jaew, married, 24 years old). (During this conversation, the interviewee was crying. The researcher stopped the interview for a while, waiting until she felt better before starting again).

With doubts about what would happen to them and their babies, most of the women needed psychological support from the provider and their significant person, because during the pregnancy termination process, physical pain and psychological trauma commonly occur to young people.

"The provider in my home town does not care about the clients, if you want to do it then you do it. If you are unsatisfied and do not want to do it, you go home. When I felt pain and was nervous, the provider scolded me and told me to stay still or go back home" (Pung, married, 18 years old).

In addition, one 18-year-old respondent lived in a low-income community. She believed that if a woman did not recover after terminating the pregnancy, the staff at the clinic would take her out into a field. They were afraid that if someone came and saw women who had just had their pregnancies terminated at the clinic, they would be arrested, because it is illegal to terminate pregnancies, and the owner(s) and the provider(s) would end up in jail. If the women were left in a field after the abortion, some would recover and some would die of blood loss.

"One of my friends, who had experience of abortion, told me that the provider gave her some medicine that made her felt dizzy. During this time, if a woman felt dizzy for a long period of time, the clinic's staff would take her out into a field and leave her there. If she were lucky, she would be safe. If she had bad luck, she would die due to blood loss" (Maew, married, 18 years old, housewife).

#### Interaction With the Provider at the Antenatal Care Clinic

When faced with an unplanned pregnancy, most of the women wanted to terminate it. However, some young women knew only after the first trimester. Thus, this group had fewer choices than others who knew in the early months of pregnancy. As discussed earlier, terminating a pregnancy is very expensive after the first trimester. Thus, most of the women in this situation gave up and continued their pregnancy to full term. However, a few cases decided to keep the baby to term and visited the Antenatal Care Clinic (ANC) because of stigmatization. The cases who visited the ANC were less sensitive to social stigma, e.g. the married women.

#### Young Single Women Did Not Dare to Visit the ANC Clinic

Once the women realized that they could not terminate the pregnancy, some decided to go to the Ante Natal Care Clinic (ANC), but most of the young women did not go to the ANC clinic because they were afraid that they would meet someone they knew. Lack of money was also a main reason for many women deciding not to go to the ANC clinic. Married and out-of-school adolescents tended to go to the ANC clinic more than others.

"I visited the ANC clinic at the fourth month of pregnancy after the abortifacient products were not successful" (Tan, married, 18 years old, housewife).

#### Factors Deterring Young Women from Visiting the ANC Clinic

The cost of the services was not the main concern of the young people because this was their first experience of pregnancy. Some of them could utilize the thirty-Baht scheme. Their main concern was stigmatization, because some of them were students or unmarried. They were afraid that the providers would blame them or have contempt for them.

"I paid 30 Baht for services apart from the medical fee. The thirty-Baht card helped me at the time of delivery" (Lek, single, 18 years old, factory worker).

The distance from home to the ANC clinic was also an issue for concern, because many of the young women with unplanned pregnancies had separated from their partners. Moreover, most of the women lived alone. Some lived with their parents, relatives, or friends, but at the beginning, most of them hide their pregnancy from others. Consequently, they preferred to visit the clinic by themselves. For this reason, if the clinics were near their residences, it would be helpful for married women. For the unmarried women, distance was not a problem because stigmatization was the main concern.

"Far or near is a concern because I consider that it is possible for me to go alone" (Maew, married, 18 years old, housewife).

Apart from the external factors that deterred young women from seeking ANC services, the interaction with providers at the clinic was a concern. The women would have negative or positive impressions, which depended on the provider's approach. If the providers provided services with a positive attitude and approach, the young women would feel good. However, if the providers had negative attitudes towards premarital sex, they would feel bad.

"Before visiting the clinic, I thought they would ignore me. When I visited, I found that they cared because they came and asked whether the baby was hungry. I felt good" (Lek, single, 18 years old).

At the ANC clinic, if the young women received a positive service approach, they would feel warm and dare to consult the providers. Most of the women were afraid to ask questions because they did not want to be blamed by the providers. They would keep the questions to themselves and try to find the answers on their own. They might seek information from friends, relatives, and various media. With unreliable sources of information, they might get inappropriate answers.

### Interaction with Friends and Providers at the Shelter

The shelter is a place where women with unplanned pregnancies can hide themselves from others. In this place, they meet other women with the same problem. They have a chance to interact with themselves, other women, and providers. Women are assigned to stay together in the big or small room depending on the design of each shelter. Some shelters have small income-generating activities for the women, such as sewing and cross-stitch. One shelter had a cooking group to generate a small income for the women. These activities made the women work together as a team, and this created a friendly atmosphere, where it was easy to get together, share experiences, and care for each other.

### Sadness and Happiness are Common at the Shelter

When many people stay together, it is common for them to quarrel. Sometimes, the women argued with one another because of stress, because many of them still had no solution to how they could manage their lives in the future. Some of them felt that they had nowhere to go and nobody supported them. In these stressful situations, some women could control their tempers, but some could not. Women stayed at the shelter for periods ranging from a few days up to many months. One case stayed at the shelter for nine months; she had nowhere to go because her parents had divorced and they had both remarried. She used to live with her mother but her father-in-law raped her. So, she ran away to seek help from the shelter. However, women who stayed for a long time got together as a group and became powerful. Thus, it was difficult for the newcomer to join and communicate with the earlier group. Sometimes, the newcomers who could not adjust themselves would suffer from isolation. Some women tried to get together with other newcomers. Thus, there was conflict at the shelter when there were larger numbers of women staying together. However, this conflict lasted for a while, and then they got together and later became friends. One respondent compared the situation to teeth and tongue, which sometimes fight and sometimes are friendly. However, some women realized that the women were different in many ways, including background, thinking, and age. This rationale made them apologize to the others.

"I felt good at the beginning when I came here. When I stayed on longer, we started quarreling and fighting" (Tan, married, 18 years old, housewife).

"Women who stay here are separated into groups. As a newcomer, I felt that it was difficult" (Lek, single, 23 years old).

"It should be (Fighting with one another). Women came from different families and were staying together in the same place. They had different backgrounds, ways of thought, and age. Some were very young, some old, so they did not get along well" (Aui, 19 years old, vocational school student).

Some of the women even disclosed a conflict situation among themselves at the shelter. However, there were many positive outcomes of staying at the shelter. Many women said the same thing; that they felt better when they decided to stay there. They stayed at the shelter with other women who were in the same situation, so they could compare their situation with others that were worse than theirs. Thus, they motivated themselves into believing that they needed to fight for themselves and the baby. At the shelter, they were not alone; they had a person they could walk in and consult at any time. Moreover, the activities arranged by the providers at the shelter included a workshop for mother and baby, recreational activities, handicraft work, and other assigned group work. These activities made them feel relaxed, happy, and enjoy living with the other people. Once they became friends, they helped each other by encouraging, telling them about their experiences, sharing their knowledge, and supporting the routine activities of the shelter.

"Sister always emphasizes every week that everyone has their own problems, so we must not let problems assail us or make us blue. Let beautiful thoughts come into our minds" (Su, single, 19 years old, babysitter).

"When I lived outside, there was social pressure because I was pregnant without a responsible partner. Living at the shelter, there are women who all have problems. We share our history and experiences. I also had a chance to ventilate by sharing my experiences. I feel better" (Ploy, single, 17 years old).

"It made me realize (by staying at the shelter) that there are other people who are in worse situations than me" (Brew, student, 16 years old).

"There are many reasons that made me feel better. Some of my friends have more serious problems. For me, I already have a solution. After I deliver, I can still raise my baby but others cannot. They need to give the baby up for adoption" (Pu, single, 21 years old, unemployed).

# Encouragement and Arranging Activities are Provider Roles

Most of the women with unplanned pregnancies who visited and stayed at the shelter had failed in the use of abortifacient products or had financial problems going to an abortion clinic. Thus, they would come and stay at the shelter, to hide until delivery. At the beginning of their stay at the shelter, some of them did not have any solution for their future. Some women were depressed, however, interaction with providers through counseling, workshops, and assigned routine activities, helped them to feel better because they started building relationships with the providers and with others. The women trusted them and felt that they could help them solve their problems. Whenever they had problems their friends could not help them solve, they would turn to the providers. In addition, group counseling helped them gradually to make them stronger by learning and sharing experiences with each other. By learning from other people's experiences, the women could reflect on themselves and understand themselves better. Once they could understand themselves, they could establish their goals and be ready to go back into the community. One example was Porn, who was an 18-year-old, single factory worker. She was alone after her parents had separated. She did not know any of her relatives. She was pregnant because a man she knew raped her. She came to stay in the shelter because her friend told her about it. At the shelter, she could make the decision to put the baby up for adoption and go back to work in the new factory. The second example was Su, who was a babysitter. She was pregnant at 19 years of age to her boyfriend without the knowledge of their parents. She decided to raise the baby but she did not know how to manage her life because her boyfriend lived in another province and she lived alone in Bangkok. Afterwards, she consulted a provider and got a better idea, which was to put the baby into the nursery provided by Ban Tantawan, which was a non-governmental organization.

"I decided (to put the baby up for adoption) while I was staying here. If I took the baby with me, how could I raise it? I have no money, which I need to pay for the costs of accommodation, milk powder, baby sitter, and daily living costs. I do not know how to earn an income to pay for these costs. Moreover, I have no experience of raising a baby. If I put the child up for adoption, the baby and I will be better off. The provider gave me information on how to raise the baby. She also asked me how I could raise the baby by myself. If I put the baby up for adoption, the baby will have a future. The providers will check the adopters before giving the child to them" (Porn, single, 18 years old, factory worker).

"First of all, I consulted my close friends, but my friends had no ideas. Then, I consulted providers regarding a temporary place for raising the baby. The providers suggested that I needed to wait in a queue" (Su, single, 19 years old, baby sitter).

# **DISCUSSION AND RECOMMENDATIONS**

Decision Making Process and Help – or Health-Seeking Pattherns of Young Women with Unplanned Pregnancies

As evidenced in the above discussion, the in-depth interviews indicated that once the pregnancy was confirmed, they would feel worried and seek help or consultation from the popular sector first, which included the partners, parents, and friends. After they had gained more information and interaction with themselves, their partners and/or parents, they would make the decision by themselves, which was based on their status, attitudes, relationships and the consultation outcomes. If they decided to keep the baby to term, they would not seek further service from the professional sector but might consult the popular sector to seek support. The majority who decided to terminate their pregnancies would seek further professional services, especially at drugstores, which were most popular among low-income young people seeking abortifacient products for self-terminating the pregnancy. It was popular because there would be less interaction with the providers. Few of the young people

utilized government health facilities due to the abortion laws, beliefs and stigmatization that deterred the individual and the family member from seeking services there.

However, if the women failed in their use of abortifacient products, they tended to seek more efficient facilities, such as private clinics. Some of them were successful, but some were not. The unlucky women would try as hard as they could to seek ways to terminate their pregnancies until they felt that it was not possible to do as planned, and then they would abandon the attempt with no hope for the future.

Regarding the health-seeking patterns of the low-income young women with unplanned pregnancies, it was observed that the women started with simple self-medication, then moved towards more complex efforts i.e. taking higher doses or mixing abortifacient regimens, beating, jumping up and down, or visiting private clinics to terminate the pregnancy. However, because of premarital sex, which is a major cause of unplanned pregnancy, it was different from other health problems. The women's ability to seek and utilize certain help- or health-services is socially controlled by laws, regulations, culture, and norms. However, none of these factors were studied, but they do impact upon the choices of young women with unplanned pregnancies.

To examine this research question more extensively, it is important to compare the help or health-seeking patterns among young women with unplanned pregnancies. According to Kleinman's (1980) healthcare system framework, the professional sector consists of modern medicine, while the folk sector is, in general terms, traditional medicine. The popular sector is the sector where health problem are first defined and help- or health-seeking behaviors are initiated. It is also the sector where multiple layers of factors (individual, family, society/peers, and community) interact to shape and influence these help, or health, beliefs and behaviors. Hence, the findings for help or healthcare seeking in this study are consistent with Kleinman's theoretical framework of the healthcare system.

The Characteristics of Service Facilities that Meet the Needs of Young Women with Unplanned Pregnancies

# *Expectations of Women with Unplanned Pregnancies of the Types and Characteristics of Services Needed*

The provision of sexual and reproductive health services (SRH) to young women with unplanned pregnancies need to be sensitive to the characteristics of service provision, because values, norms, and culture control their help- or health-seeking behaviors. Thus, to have them change their utilization to the existing service interventions, they need to be involved in all of the processes of service provision. According to the in-depth interviews and observations at the shelter, unplanned pregnancy was not a physical illness; hence, women in this situation need more psychological support and counseling. While they were in a crisis situation, they had nobody to counsel them, and many of the young women mentioned that if they had got counseling they might have arrived at a better solution and pass through the crisis more smoothly. They also said that they needed shelter to hide in while their belly was getting big. Moreover, they needed support setting up with the baby as a single mother, because many of them faced an unfaithful man and were abandoned. The following are details of each type of service needed: **Counseling once pregnancy is confirmed**. The in-depth interviews indicated that before knowing about the shelters, many women expressed the same view, that they had nobody to advise them how to manage when they were in trouble. If they had someone to counsel them they might have better solutions instead of living under a dark cloud. They said that telephone counseling was more convenient and comfortable for them before they came in to utilize the services. Most preferred female counselors with a friendly manner.

Shelter to hide in while the belly was getting big. During the first trimester of pregnancy, most of the women still hid their pregnancy in large clothes because their belly was not very big. However, during the second to third trimester, their belly was getting big. The in-depth interviews indicated that the bigger the belly was, the more stress the women felt. As there are a limited number of shelters throughout the country, shelter is available only in the big cities and Bangkok. Moreover, few women know about the shelters, thus, some of the young women with unplanned pregnancies ran away from their community to new places where nobody knew them. The women with bigger bellies who still lived in the community would feel embarrassed. Hence, if there were more shelters available, the women would endure less stress and could cope with their problems. Moreover, the in-depth interviews with the women in the shelters revealed that the shelter is a good place where women who have the same problem can share, support, and learn from each other's experiences. It helps them feel stronger and ready to return to their communities.

**Nursery care a few months after delivery.** The in-depth interview results revealed that the women who failed to terminate their pregnancies would be very worried about the baby's future, because they had no income and had been abandoned. Hence, many of them felt worried about how to manage their lives after they delivered. Many women said that they needed a place to nurse their baby at least for 3 months, so that they could earn income and set up the new life first; then they would be ready to take care of the baby. However, the results from the in-depth interviews with the social workers at the shelters revealed that the women needed to wait in a queue for a nursery room for their baby, because there were limited numbers of free nursery care services available.

**Support for setting up a new family**. The in-depth interviews indicated that the majority of the women had had their first unplanned pregnancy. Moreover, due to inexperience, youth, immaturity, and lack of income, many of them needed support at the beginning of setting up the new family.

**Counseling to prevent recurrent unplanned pregnancies.** Apart from the services and activities provided at the shelters, explained above, the in-depth interviews indicated that the women expected counseling, which was very important. If they were in a shelter, they preferred to have face-to-face counseling. The counseling would support them and prevent them from recurrent unplanned pregnancies.

The in-depth interviews of the young women with unplanned pregnancies indicated that friendly administrative processes included: 1) a private area for history taking, 2) anonymity, 3) no identity card required, 4) no need for the parents or caretakers to approve of using the service. The minimum requirement was confidentiality, to make the young women feel relaxed and comfortable to utilize the services with a good impression.

Apart from the types and characteristics of services needed, public relations and information dissemination to potential users about the services should be widely available, so that the women would have more alternatives to taking abortifacient products or visiting a private clinic to terminate the pregnancy, even when they did not want to.

#### Recomendations

Based on the in-depth interviews and FGDs, solutions and recommendation for various health systems emerge from this study, which are listed as follows:

# Pharmacist Association, Pharmacy Association, Drugstore Club of Thailand, and other related Pharmacy Organizations

Local business/drugstore as a focal point for increasing access to prevention and care for young women with unplanned pregnancies. Based on the in-depth interview study of the help-or health-seeking patterns of the young women with unplanned pregnancies, most of the women did not seek help or services from the formal government healthcare system. Thus, the provision of comprehensive unplanned pregnancy prevention and care services should be established through collaboration from both private and public sectors. Drugstores should be an initial source of collaboration and referral, since the majority of the young women with unplanned pregnancies sought abortifacient products, menstruation regulators, and other essential products from this source.

The most important is that drugstore should be the best option for providing primary prevention services, education, and information for both unmarried and married youths. Information towards emergency contraception, and other related contraceptive technologies to the potential users to prevent unplanned pregnancy should be provided through this channel.

#### Health Department, Ministry of Public Health

Healthcare providers in all levels of health facilities need to screen for unplanned pregnancies. In the in-depth interviews with the providers of obstetric services, it was suggested that nurses or clinicians should assess young people's pregnancies, and make more detailed decisions in determining planned or unplanned pregnancies. In addition, in assessing the women decision-making on the choices for unplanned pregnancies, they should consider whether to continue or terminate the pregnancy, so that the provider can provide appropriate counseling and support, which will increase women's decision-making capacity in areas of life, sex, and pregnancy.

**Establishing comprehensive sexual and reproductive health systems.** Health Department should play as an umbrella's role for all organizations dealing with sexual and reproductive health (SRH) provision. The organization should collaborate by segmentation to provide a comprehensive SRH package, to make a broader contribution in delivering SRH services and preventing unplanned pregnancies among young women, such as counseling and testing for pregnancy, STI/HIV; diagnoses for abnormal vaginal discharge or menstruation; treatment and care for STI/HIV; post-abortion care; and contraceptive technologies.

Clinics or small hospitals may be better targeted to provide secondary services and care, including antenatal care, delivery, postpartum care, and well-baby care. Tertiary care, including general hospitals, university hospitals and large private hospitals that are well-equipped and have medical specialists, should provide more complex health services, including diagnosis, delivery, treatment services, and any complications or abnormal pregnancy or delivery.

Moreover, all organizations should provide couple counseling and information to prevent unwanted pregnancies in the future. Most important is that the services should ensure that all women who in different situations are effectively referred to facilities with services and encourage the use of such services. For example, youths who visit drugstore personnel for abortifacient products might be given a referral to see a counselor(s) at a women's organization or at a shelter. In this way, "friendly service" could be linked to several types of healthcare facilities.

The referral network should include drugstores, clinics, community health centers, hospitals and shelters. Moreover, the beneficiary group, the school and home, should be strengthened to network with the service facilities, so that it can be assured that women with unplanned pregnancies and other SRH problems are fully assisted. All the organizations need to work together to develop formal acceptable practice guidelines or protocols to help ensure that each organization provides specific services as indicated in the protocol and have clear criteria for referral.

**Training of providers.** The Health Department should provide training of service providers at all level of service facilities on delivery of gender-sensitive, gender equality, and quality of services (i.e., positive interpersonal relationship, confidentiality, privacy, in counseling and management, and referral).

**Provision of emergency contraception, i.e., emergency contraceptive pills, and postcoital insertion of IUD.** Based on the in-depth interviews majority of the sexual intercourse had happen by chance. So, all level of health care facilities need to provide and made available of the emergency contraception to the person who are in needed.

Strengthening counseling services, which are essential for unplanned pregnancies women. In the in-depth interviews, most of the women mentioned that they needed counseling once the pregnancy was confirmed. Counseling services should provide both hotlines and face-to-face. Moreover, it should add on male and couple counseling to their routine services, which emphasize on female. They should have specially trained personnel to provide counseling to young people about unplanned pregnancy. They should also be able to provide comprehensive counseling to respond to women's emotional and physical needs, and their other concerns related to their trouble.

Adoption services. As shown in the in-depth interviews, abortion was most popular among the young people with unplanned pregnancies. One underlying reason was that many young people did not know that an adoption agency was available. For this reason, information about this agency should be provided to women who will potential want to have an abortion, so that they have one more choice for consideration. Moreover, the government needs to support and strengthen this type of service and make it available nationwide.

**Male involvement in reproductive health.** The in-depth interviews indicated that the majority of unplanned pregnancies were attributable to males, so that it is important to develop programs targeting males, aiming to change their attitudes and behaviors, to be more responsible for sex and its consequences.

**Involve young people in all stages of program development.** It is important to involve young people in all stages of program development because they can help ensure that the programs are relevant to young people's real needs. Moreover, they can help identify communication channels, activities, and messages to other young people in meaningful ways.

**Evaluation.** The umbrella organization should be responsible for evaluation of all components of the program, to determine the effectiveness of preventing current and future pregnancies among young people, and expanding upon those shown to be effective.

#### School Masters, and Teachers

Enhancing sex education in the school by collaboration with the home. The in-depth interview and structured interview results showed that unplanned pregnancy could occur at the age of 12, or older. Even though they are at a low risk compared to those who engage in risky behavior, such as drug taking, alcohol consumption, or smoking, it is important to prepare them to make responsible decisions when the rare opportunity to have sex occurs. Life-skills decision-making regarding sex, unplanned pregnancy, its consequences and prevention, should be taught in school by collaboration with the home before completion of grade 6. Sex education should be comprehensive and consistent. This will help the young people make decisions based on rationality, responsible choices made with an awareness of the available options. Strategies for school training are as follows:

**Expanding the training of teachers.** Teachers who train in life skills should be trained on a regular basis in transaction and teaching techniques.

**Involving young people.** Since peers have more influence on sexual behaviors, the school program should promote more active involvement of young people in all stages of program development to increase the effectiveness and marketing of its activities. Feedback should be sought from the young people regarding what kinds of support would be helpful for them to prevent unplanned pregnancies, and what things the providers should keep in mind when providing services or addressing issues of sexuality. All the responses would help improve the program and increase its acceptability among young people.

**Encouraging the school system to work with the home.** Since parents also impact upon their children's sexual behavior, the school program should support parents developing and nurturing their children. Moreover, the school system should work with the parents to assess the developmental needs of the young people and identify programs, resources, and tools to assist them in making responsible choices. In addition, they could share information to prevent substance abuse at all levels of childhood development, because substance abuse is likely to be associated with unplanned pregnancy.

**Encourage the school system to work with the community.** The community environment also impacts upon the sexual and reproductive health of young people. Thus, the school system should work with community leaders to assess and develop a youth-friendly environment.

#### Family

**Parent-child relationship**. As indicated in the findings that family played important roles on unplanned pregnancy among young women. Parents are the key people who play in shaping the perceptions, beliefs, and behavior of the young people towards adolescent sexual and reproductive health. If the relationship between the young people and parents is weak, then they will turn to their peers and follow their advises. However, if the relationship is strong the young people will turn to their parents whenever they have problems.

To create a good relationship among them, parents need to change of attitude, including understanding their own sexuality in the a way that enabled them to relate positively with their children on sexuality; and understanding of sexuality and adolescent psychology, communication, and counseling skills.

#### Representative of Shelters

**P.R. the services to the potential users.** Based on the findings from in-depth interviews, there were a few young women who knew about the shelters services. Majority of them were refereed from private clinics, churches, or words of mouth. If they knew about shelters' services, it would help them released from the crisis, and had more choice.

**Create shelter's image as a second home for the women.** The shelters should create images as a second home for the women. Also, it should be a place for all young women who want to get information, education, counseling, and cares towards women issues.

**Continue to support the women who want to raise the baby themselves.** The women with unplanned pregnancies who want to raise the baby themselves need a follow-on support at least for 3 months. Due to they are young and less experience on caring the baby, thus, the shelter needs to create a systems to help support and follow-up after the women go back to their family, or community.

Create the self-help group in the shelter and continue to strengthen the network. During the women who face the same crisis are at the shelter, it is important to create a friendly environment by having them get together, and sharing experiences. Counselor at the shelter should play as a facilitator at the beginning in establishing group activities. After they are settled, the facilitators should let them manage the group by themselves.

#### Decision Makers, Policy Development, and Women's NGO(s)

**Provision of one stop comprehensive services.** The health care facilities should provide a comprehensive one-stop service for the young women with unplanned pregnancies to reduce crisis that they may encounter at the long processes.

**Established youth friendly shelter nationwide.** At present, all the shelter s provide services for the women at all age group. However, majority of the unplanned pregnancy women are in the adolescent stage. Thus, it is worth to create youth friendly environment and services for the young women while they are staying at the shelters.

**Relief regulation in educational system.** Based on the findings from in-depth interview, the main reason of terminated pregnancy was that the young women were student. So, to relief regulation in educational system would help those adolescents having more choices rather than terminating pregnancy.

**Revise of business or labor rules and regulations.** The business, labor rules, and regulation, which are unfair, and discriminatory, i.e., deterrence of pregnancy during employment, no paid for maternity leave need to be revised.

**Revise of the abortion law.** The most important in issue of unplanned pregnancy is that the abortion law, which is unfair to the women, need to be revised. Also, the law needs to decrease in restrictions for doctors to perform induce abortion and less stringent conditions than the existing one. Moreover, it should allow the women to access to induce abortion if needed. This effort will decrease risks of unsafe abortion especially those performed by "quacks" and self-management using several of abortifacient product regimens.

### ACKNOWLEDGMENTS

We would like to express our sincere gratitude and deep appreciation to all the young women in the five shelters and communities of Bangkok Thailand who patiently, candidly answered, and shared more on their grief story.

Last, we are greatly indebted our family who are always shares their love, support, and encouragement.

## REFERENCES

- Asian Institute for Development Communication (AIDC). (1999). "Module 1: Knowledge and Information on Adolescent Reproductive Health." Promoting Advocacy for Adolescent Reproductive and Sexual Health Through Effective Communication: A Training Program for Media Practitioners, Bangkok, Thailand, November, 1999.
- Boonmongkon, P. et al. (2000). "Thai Adolescent Sexually and Reproductive Health: Implications for Developing Adolescents' Health Programs in Thailand." Center for Health Policy Studies, Faculty of Social Sciences and Humanities, Mahidol University at Salaya, Thailand.
- Burack R. (1999). "Teenage Sexual Behavior: Attitudes Towards and Declared Sexual Activity." *Br J Fam Plann. Jan; 24 (4)*: 145-8.
- Chuamanocharn, P. et al. "Community Drug Utilization and Health Service System Development Project: Situation of Community Drug Utilization, Problems, Obstacles and Ways to Solve Through Drugstores Services in Payao, Lampang, and Nan." Pharmaceutical Science faculty, Chiengmai University, 2000.
- Ford, N., and Kittisuksathit, S. (1996). "Youth Sexuality: The Sexual Awareness, Lifestyle and Related-Health Services Needs of Young, Singles, Factory Workers in Thailand." Institute for Population and Social research, Mahidol University at Salaya.
- Gray, A. and Punpuing, S. (1999). "Gender, Sexuality and Reproductive Health in Thailand." Institute for Population and Social Research. Mahidol University, Thailand.
- Gramzow, R and Major, B (1999). "Abortion as Stigma: Cognitive and Emotional Implication of Concealment". *Journal of Personality and Social Psychology*, v. 77 no 4, (oct. 1999), p.735-45.
- Kanchanaraj, T. et al. (2001) "Situation of Community Pharmacy Services: Problems of Drug Utilization in Khon Kaen and Udornthani and Recommendation for Promoting Rationale Drug Use in the Community" Part of Community Drug Utilization and Health Service System Development Project Phase I, Khon Kaen." Community Pharmacy Division, Pharmaceutical Science Faculty, Khon Kaen University.
- Kleinman, A. (1980). "Patients and Healers in the Context of Culture: An Exploratory of the Borderland between Anthropology, Medicine and Psychiatry." Comparative studies of health care systems and medical care. London England: University of California Press Ltd.
- Koetsawang, S. (1993). "Illegally Induced Abortion in Thailand." Paper Presented at IPPF SEAO Regional Programme Advisory Panel Meeting on Abortion, Bali, Indonesia, 29-30 October.

- Nelson, K. (2000). "The Effect of Youth Friendly Services on Service Utilization among Youth in Lusaka, Zambia." A dissertation submitted to Tulane University in conformity with the requirements for the degree of Doctor of Public Health, New Orleans, Louisiana, USA.
- Naravage W. et al. (2005). "Factors Affecting Decision Making of Low-Income Young Women with Unplanned Pregnancies in Bangkok, Thailand." Southeast Asian J Trop Med Public Health, Vol 36, No.3, p 775-782.
- Manisha G.; Sunita B.; and Hemlata P. (1997)." Abortion Needs of Women in India: A Case Study of Rural Maharashtra." *Reproductive Health Matters, no. 9*, p.77-86.
- PATH (2001). "Adolescent Health Center: Ban Wai Sai." Paper present at the Asia Pacific Conference on Reproductive Health (APCRH), Manila, Philippines, February.
- Pisulbutr S. (1997). "Thai Economic Classification." Statistics Association Journal, 23: 20.
- Punyawuthikrai, P. et al. (2001). "Community Drug Utilization and Health Service System Development Project (subproject: Bangkok)." Bangkok: Pharmaceutical Council (Thailand).
- RHO (2002). Internet site: http://www.rho.org
- Ratchukul S. (1998). "An Unwanted Pregnancy and Decision to Terminate Pregnancy." A dissertation submitted to Srinakarindtrawiroj University in conformity with the requirements for the degree of Doctor of Social Development, Bangkok Thailand.
- Sainsbury, J. (1997). "The New Inequality-Women Workers's lives in Thailand and the Philippines." A Catholic Institute for International Relations Briefing.
- Sertthapongkul, S and Phonprasert, P. (1993). "Causes of Unplanned Pregnancy and Decision Process for Abortion: A Case Study at Bangkok Private Clinic." Journal of Development Administration, Bangkok, Thailand.
- Soonthorndhada, A., (1996). "Sexual Attitude and Behaviors and Contraceptive Use of Late Female Adolescents, Bangkok: A Comparative Study of Students and Factory Workers." Institute for Population and Social Research. Mahidol University, Thailand.
- Warakamin, S. and Boonthai N. (2001). "Situation of Induced Abortion in Thailand." Paper Presented at the Seminar on Problems of Terminated Pregnancy, Bangkok, Thailand, 6 August, 2001.
- WHO (1997). Internet site: http://www.who.int/reproductive-health/publications/MSM/97.16
- UNESCO Regional Clearing House on Population Education and Communication Asia and the Pacific Regional Bureau for Education Bangkok, Thailand. (2001) "Package of Research Briefs Series 2: Adolescent Reproductive Health". December, 2001.

Chapter 17

# PROJECT TOWARDS NO DRUG ABUSE (TND): A NEEDS ASSESSMENT OF A SOCIAL SERVICE REFERRAL TELEPHONE PROGRAM FOR HIGH RISK YOUTH<sup>\*</sup>

Steve Sussman<sup> $\dagger 1,2$ </sup>, Silvana Skara<sup>1</sup> and Patchareeya Pumpuang<sup>1</sup>

<sup>1</sup>Institute for Health Promotion and Disease Prevention Research, University of Southern California, Department of Preventive Medicine, Keck School of Medicine, Los Angeles, California, USA <sup>2</sup> University of Southern California, Department of Psychology, Los Angeles, California, USA

# ABSTRACT

This chapter describes a quantitative and qualitative needs assessment of a potential social service resource telephone program component among high risk youth who received the Project Towards No Drug Abuse (TND) classroom-based program (approximately 1-year earlier). Information was obtained to determine whether the targeted youth would be interested and receptive, or even need the information available from such a program. Results supported youths' overwhelming receptiveness of a social service referral program. The vast majority of respondents indicated a strong desire for resource and referral information on vocational, educational, recreational, transportation, and mental health and drug counseling. Participants' responses will be used to better structure and tailor our booster program. Further research is needed to investigate the effectiveness of the provision of social service resource information on drug use among emerging adults.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: Professor of Preventive Medicine and Psychology,Institute for Health Promotion and Disease Prevention Research'University of Southern California, Department of Preventive Medicine, Keck School of Medicine,1000 S. Fremont Avenue, Unit 8, Alhambra, California 91803,Email: ssussma@usc.edu

# INTRODUCTION

#### Drug Use Consequences: Teens and Emerging Adults

Many negative consequences befall substance users during adolescence or in emerging adulthood, a period defined here as extending from approximately 16-25 years of age (coinciding with the developmental period when young persons transition to independent roles; Chassin et al., 1999; Kandel et al., 1986; Krohn et al., 1997; Maggs et al., 1997; Newcomb and Bentler, 1988; 1988b; O'Leary et al., 2002). These consequences include adverse immediate events (e.g., overdoses, accidents, and physical health problems), early involvement in family creation and divorce, and crimes such as stealing and vandalism. Additional negative consequences include dropping out of high school or college, seeking less skilled employment, developing unusual beliefs that interfere with problem-solving abilities (e.g., thinking about the "quick fix"), diminished adaptive coping and achievement behavior, and greater social isolation and depression. Finally, some adolescents and emerging adults suffer a sufficient number of negative consequences that classify them as suffering from substance abuse or dependence disorders (Sussman et al., 1997).

#### Developmental Transitions during Emerging Adulthood

Young people leaving high school are expected to seek new opportunities (Arnett, 2000; Bachman et al., 1997). These may include: (a) assuming career avenues and financial independence, (b) learning skills of independent living (e.g. buying or renting a place to live apart from one's parents; Bachman et al., 1997), (c) growth in self-care skills (e. g. cooking, cleaning, grooming, buying goods, and traveling), and (d) social adventures (e.g., love and young adult groups). Social adventures lead eventually to commitment in relationships (e.g., marriage and children). Youth may also transition from a relatively high level of family conflict in adolescence to the reduction of such conflicts in emerging adulthood, as one achieves an emotional distance from parental demands and begins to associate more as a junior peer (Aseltine and Gore, 2000). One may argue that youth enter a protracted life phase. Within this phase, there is a "trade-off" between what one commits to and new areas of exploration that must be abandoned. For example, getting married would tend to preclude further dating. Beginning a full-time job would tend to preclude taking on another full-time job with a different directional emphasis. Events characteristic of young and middle adulthood, such as taking on the role of a parent, economic provider, and nurturer, leads to new experiences but often in sacrifice for others and also leads to increasing law abidance, diligence, and conservatism (Stein et al., 1986).

Youth who receive parental support are academically and socially competent, strongly bonded to school, attend church, and hold normatively popular attitudes are relatively likely to transition smoothly into normatively defined young adult roles (Brook et al., 1995; Guo et al., 2001; Krohn et al., 1997; Maggs et al., 1997). Normatively popular attitudes include such sentiments as valuing one's health, affirming the importance of hard work, and expressing respect of family. On the other hand, youth that exhibit unconventional behavior (e.g., cheating, having a child out of wedlock), unconventional attitudes (e.g., tolerance of deviance

and preference for sensation seeking), poor emotional control, anger, intrapsychic distress, and interpersonal difficulty are relatively likely to use drugs in emerging adulthood (Aseltine and Gore, 2000; Brook et al., 1997). These youth tend to enter adult roles early ("precocious development"), prior to being prepared to take them on. They tend to drop out of high school or attend part-time education, get married and quickly divorced, become a parent while relatively young, and take on relatively undesired full-time employment. Those teens who exhibit precocious development are at particularly high risk for drug use in emerging adulthood (Krohn et al., 1997). Successful marriage (often forestalled until later in emerging adulthood) is inversely related to drug use, possibly because social opportunities to use decrease, and relationship commitment and consideration of the other person may reduce one's desire to use. This effect applies to males and females, though more strongly for females. Pregnancy and parenthood, in the context of happy marital relationships, are statuses that are also inversely related to drug use. On the other hand, cohabitation, which is relatively strongly associated with holding nontraditional beliefs, is positively associated with drug use. Job stability in young adulthood is negatively related to drug use, although participation in the military or in hourly jobs may be relatively strongly associated with use of cigarettes and alcohol (Bachman et al., 1997). Another major predictor of drug use in emerging adulthood is drug use in adolescence (Brook et al., 2002). The stability of cigarette smoking from high school graduation over the next 10 years is very high, is moderately high for alcohol and marijuana use, but decreases dramatically for other illicit substance use (Bachman et al., 1997; Rohrbach et al., 2005). Finally, age is a major curvilinear predictor of drug use. Experimental drug use tends to peak during this period of emerging adulthood (Bachman et al., 1997), and then tends to decline later in young adulthood (around 25 years of age). Prevention programs focusing on emerging adulthood must not only acknowledge the features of emerging adulthood (e.g., increased exploration of various experiences) but also must link these features to feasible strategies of prevention (e.g., enhancing pro-social skills, providing job information).

## Prevention Programming Relevant to Emerging Adulthood

Relatively few prevention programs have been completed among persons in emerging adulthood. A recent review by Sussman and colleagues (Sussman et al., 2004) examined all evidence-based (model or promising) targeted (indicated or selective) drug abuse prevention programs identified by the Blueprints program, SAMSHA, and in a literature search of Medline and PsycINFO. A total of 29 programs were identified (as of 2004). Of these programs, 16 included youth that were 16 years old at baseline. Of these 16 programs, any effects on drug use were reported for 13 of them. Of these 13 programs, effects that occurred or persisted into emerging adulthood (18-25 years of age) were reported in only four programs: Nurse-Family Partnership (N-F P), Project PATHE, Reconnecting Youth (RY), and Project TND. N-F P provided home visits and resource advocacy for first-time mothers with no income from pregnancy to 2 years after childbirth. This program emphasized environmental advocacy, counseling by highly trained counselors, and decision making instruction, and led to decreased smoking and alcohol use and reduced rates of child abuse. Project PATHE provided programming to youth 12-18 years of age who were of disadvantaged socioeconomic status, low achieving in school, and disruptive. School-based

activities, peer counseling, academic skills, development of school-pride, and job seeking skills were provided. This program decreased drug involvement, school alienation, and discipline problems and increased graduation rates. RY was implemented with regular high school youth at risk for drop-out from high school. This program involved 90 class sessions and use of small student groups, and led to decreased hard drug use, and less perceived stress and improved grades at approximately a one-year follow-up. Finally, Project TND was implemented with alternative high school youth, as well as regular high school youth. This program involved 12-sessions, which helped motivate youth to change their perspectives on and perceptions of drug use, learn social and life skills to bond to pro-social institutions, and decision making to help them plan good solutions to complex, problem situations. Decreased use of cigarettes, alcohol, marijuana, hard drugs, and weapons carrying was found in this study at one-year follow-up. Effects on cigarette smoking, hard drug use, and marijuana use were maintained at a two-year follow-up. Together, these programs encouraged several changes. They sought to help youth change their drug use motivations, learn new skills (communication, self-control, academic, job seeking), and make good decisions. In addition, they sought to make the older teens' or emergent adults' current life situation more tolerable through providing leads on how to obtain new environmental resources. Extended programming appeared desirable.

# Social Service Resource Telephone Program (SSRTP) and Booster Programming Considerations

Relatively little research exists documenting the influences of the provision of social services resource information on drug use among youth, as was indicated in the Sussman et al. review (Sussman et al., 2004). Still, availability of environmental resources has been proposed as a potential elaboration of the DSM axes (Scotti et al., 1996), and the importance of providing access to social-environmental resources is part of the fundamental justification within fields such as social work and urban planning. There is some suggestion of effectiveness on drug use in single-group studies by attendance of high risk preteens and teens at community-schools. These programs involve bringing in an intensive and costly in-house network of social service resources to teens under one roof (Coalition for Community Schools, 2004).

Availability or perception of availability of environmental resources such as access to jobs, education, recreation, transportation, or drug/counseling services in one's community could be enhanced among emerging adults by receiving such information through a telephone education service. If provided as booster programming following receiving drug prevention educational programming in school, this strategy could enhance hope for lifestyle stability with satisfactory self-fulfillment (Lesser and Escoto-Lloyd, 1999; Ward et al., 1982) in a time and location efficient way. However, the effects of provision of resources to emerging adults as an adjunct to drug abuse prevention education has not been attempted to our knowledge. Telephone education has been becoming a more popular modality of delivering information to teens and adults in such venues as smoking prevention (Elder et al., 1994; Skara and Sussman, 2003), smoking cessation (Leed-Kelly et al., 1996, with recovering alcoholics; Mermelstein et al., 2003; Miguez et al., 2002; Whelan et al., 1993; Zhu et al., 2000); at-risk drinking (Curry et al., 2003); prenatal substance abuse treatment (Laken and Ager, 1996);

medication adherence (Aubert et al., 2003; Tutty et al., 2000); pap screening behavior (Hou et al., 2004); mammography screening (Champion et al., 2000); exercise programming (Nies et al., 2003; Pinto et al., 2002); and nutrition behavior (Pierce et al., 2004), with demonstrable results across target behaviors in the vast majority of studies relative to no telephone education. Telephone education could be provided as a means of booster programming (of classroom-based drug use prevention programming) and to help refer emerging adults to nearby social-environmental resources.

Several conceptual models relevant to emerging adulthood have been advanced that provide a language to organize the interaction between developing persons and their environment (Mason et al., in press). Among these include psychosocial process models focusing on the role of social integration and perceived availability of support (Cohen, 1988) and "ecologically-relevant approaches" that seek to understand functional changes of individuals as they are constituted by particular settings and contexts (Bronfenbrenner, 1977; Green, et al., 1996; Mason et al., in press). Consistent with features of these models, the SSRTP component attempts to provide information to subjects to help network them with elements of their surrounding environment. Emerging adults are provided awareness that positive opportunities are "out there" in these systems and youth are motivated further to not select drug use as an alternative. As proposed herein, we believe that an SSRTP component will protect emerging adults from engaging in drug use by helping (Cohen, 1988; Eccles et al., 2003; Mason et al., in press; Millstein, 1994) in at least four ways which correspond to a MACH (mastery, attachment, cue, hope) model of helping early-stage emerging adults to transition successfully during emerging adulthood. First, participants will receive instruction to help them develop a sense of mastery and autonomy (e.g., planning how and where to get assistance in independent living). Receiving general instructions on how to accomplish a life function does provide a vicarious means of anticipating barriers and surmounting them. Further, just being prompted to learn by experience through accessing opportunities in one's environment may assist participants in achieving developmental goals of emerging adulthood (Arnett, 2000; Eccles et al., 2003; Millstein, 1994). Second, by providing information on available resources (jobs, education, transportation, recreation, and mental health or drug abuse counseling), participants might access alternative sources of institutional attachments (emerging adults linking to institutions that support healthy development), that might serve as constructive alternatives to a drug use-based lifestyle. This notion of alternative attachments is familiar within the "bonding" drug abuse prevention, social support, and ecological literatures (Bronfenbrenner, 1977; Aubert et al., 2003; Cohen, 1988; Green et al., 1995; Hawkins et al., 1987; Hawkins et al., 1992; Mason et al., in press). Third, the mere availability of the SSRTP is likely to cue participants to previously learned classroom-based program material; particularly the message that drug use/abuse is risky (Stacy et al., 1995; 1996; 1990a; 1990) this is consistent with suggested drug abuse prevention applications of the memory association literature (Stacy et al., 1996; Sussman, 2001). Since memory is sensitive to a variety of associated cues, links of classroom-based programming to the participants' experiences (prompted by the proposed SSRTP component) may subsequently lead to these life experiences cuing memories of program material. Finally, merely providing participants with the knowledge that there are numerous life options in their community or in southern California that they might think of tapping could instill hope that satisfying and stable life opportunities are available in the future. A future-directed orientation may lead emerging adults to desire to protect their health (Eccles et al., 2003; Millstein, 1994; Sussman and Dent, 2004).

Previous research shows that school-based drug use prevention programs have much greater chance of long-term success, especially for those youth who began using drugs before the prevention program was implemented, if such programs are "boosted" (Skara and Sussman, 2003). Booster programs typically involve implementation of one-to-three lessons or contacts six months to one or two years after the main program implementation. Booster lessons: (1) summarize previously taught material and (2) encourage discussion of how program material was utilized in post-program daily living. Phone boosters are feasible and promising in this population and are probably one of the only ways to reach the participants after high school. A randomized comparison is critical for future prevention efforts that consider implementing booster programs among these emerging adults. If relatively inexpensive (and easily implemented) boosters substantially enhance or maintain effects of a school-based program, then major progress would be made in understanding how to best improve effects of school-based prevention into emerging adulthood.

# CHS Youth Subjects: Risk, Ethnic Diversity, and Contexts for Booster Programming.

Since 1992, we have investigated drug abuse prevention among youth across at least 80 continuation high schools (CHSs) in southern and northern California. CHS youth are at relatively high risk for drug abuse. When reaching high school age, those youth who are unable to remain in the regular school system for functional reasons, including lack of credits and consistent use of substances, are transferred to a CHS. These schools were established in 1919 pursuant to the California Educational Code (Section 48400), which requires continued (part-time) education for all California youth until reaching 18 years of age. Every school district that has an enrollment of over 100 students in 12th grade must have a CHS program; there are approximately 600 CHSs in the State of California. Usually students who are experiencing life difficulties when beginning comprehensive (regular) high school transfer to continuation high schools where hours are more flexible and the teacher-to-student ratio is twice as high (i.e., 1:15 versus 1:30). Drugs are used by a greater percentage of youth at CHS than at regular high schools, though. Drug use rates in the last month among CHS average: cigarettes, 57%; alcohol, 63%; marijuana, 54%; stimulants, 21%; hallucinogens, 13%; and all other drugs, 5-8% (e.g., Sussman et al., 2002; n=1861). These data are similar to that reported in a national cross section of alternative high school youth (Grunbaum et al., 2000). On the other hand, among 10<sup>th</sup> graders assessed concurrently at comprehensive high schools (RHS) from overlapping school districts (n=1208), use in the last month of these substances average: cigarettes, 24%; alcohol, 36%; marijuana, 22%; stimulants, 2%; hallucinogens, 2%, and all other drugs, 1-3%. In addition, there is rich ethnic heterogeneity in continuation high schools, particularly Latino, white, Asian American, and African American youth. Project TND was developed on this diverse population. For example, names of persons in the "talk shows" used in Project TND reflect this cultural diversity. The program was found to be generalizable among youth from different backgrounds (Sussman et al., 2002).

To reach emerging adults with effective programming, one should consider the contexts within which they are likely to spend their time. We followed up persons who received project TND-1 while attending continuation (alternative) high school five years later, while they were in emerging adulthood. We noted that only 25% of the sample was still in CHS at a 1-year follow-up, 5% at a 2-year follow-up, and none by 3-year follow-up. We generated descriptive statistics for subjects' demographic characteristics and living situation at the 5year follow-up. These subjects did not differ on these variables from those at baseline that had left the pool, aside from baseline living situation (McCuller et al., 2002). Subjects were a mean age of 22 (sd=0.8); 57% male. Half of the subjects were Latino (51%), followed by 35% white, 6% African American, 4% Asian American, and 4% other ethnicity. A total of 67% of the sample reported having completed high school. During the prior 12 months, 42% of the sample had been in school or job training, with 17% in trade school (i.e., trucking, appliances, dent repair, machine operator, plumbing, financial/management, medical assistant, art, computer operator, electronics, insurance, beauty, cooking). Another 14% were in junior college, 7% in a four-year college, and 4% in adult education high school completion classes. Most (80%) longed for additional education training.

The majority (77%) was currently employed. Of those employed, 69% were working fulltime (40 or more hours per week). Their job positions included skilled laborers (27%), semiskilled laborers (16%), clerks or salespersons (14%), small business owners or managers (14%), and professionals (e.g., teacher or nurse; 11%); with 27% reporting they were unskilled laborers, unemployed, or housewives/husbands. A total of 75% of the working sample wanted to get a new job in the future. In terms of marital status, 65% were single, 30% were married or engaged, and 5% were divorced or separated. A total of 42% had at least one child. Many CHS emerging adults hope to continue educational and vocational training. In addition, their lives are busy, although not particularly satisfying. Programming for them should be concise and directed in contexts (i.e., at home by telephone) in which they have time to reflect on their behavior.

# **THE PRESENT STUDY**

This chapter reports the results of a preliminary needs assessment study that was designed to assess interest in a potential social service referral telephone program (SSRTP) component for youth. A needs assessment is a very important step to any program development, implementation, and/or improvement because it allows for the identification of potentially beneficial interventions, as well as gaps in provision that health services could meet (Billings and Cowley, 1995; Stevens, 1998; Wright et al., 1998; Gilbody et al., 2002). This approach has been used more frequently in recent times especially due to limited resources. Needs assessments studies have been conducted in many different settings such as in the medical treatment of patients in hospitals (Man et al., 2004; Rainbird et al., 2004; Wood et al., 2004) or the provision of health promotion and prevention information to various populations, including methadone users at clinics (Morrow and Costello, 2004) and adolescents in peer-led sex education programs (Forrest et al., 2004). A legitimate needs assessment is particularly crucial in developing prevention or intervention programs and in effectively utilizing program resources. The current needs assessment follows a consumerbased perspective, in that it provides identification of the needs of health care and/or improvement in health or quality of life based on the perceptions of the target population rather than on the 'normative' or professionally defined needs. In other words, a consumer need approach is necessary because there is a potential that individual needs may be ignored, while the perceived needs of the health professionals are heeded; thus, there is a danger that the planning and provision of health services would follow a top-down approach, which may end up not matching the needs of the target population.

The participants for this needs assessment study are former students who received the Project TND 12-session classroom-based program (approximately 1-year earlier). Project TND is a school-based drug abuse prevention program funded by the National Institute on Drug Abuse since 1992 (e.g., Sussman et al., 2002). This program is considered a model or exemplary program by the U.S. D.O.E., SAMSHA/CSAP, NIDA, Health Canada, Colorado and Maryland Blueprints, and numerous other agencies. The target population of our ongoing research has been primarily youth from the alternative school system (continuation high schools, CHSs) in California. The average age at baseline of these youth is approximately 17 years old (93% of the sample is 16-18 years of age). While promising short-term program effects (Dent et al., 2001; Sussman et al., 1998; 2003) were obtained for the classroom-based program, and some effects were maintained over several years (see Sun et al., under review), it is believed that program effects could be better sustained during early emerging adulthood through use of booster programming (e.g., delivered by telephone). However, before implementing this social service referral telephone program, it is first necessary to determine, through a formal needs assessment, whether the targeted youth would be interested and receptive, or even need the information available from such a social service referral program.

# **METHODS**

During the period from August to September 2004, we undertook a cross-sectional study involving a convenience sample of subjects who received the Project Towards No Drug Abuse (TND) classroom-based program (approximately 1-year earlier). Subjects were randomly selected from 5 southern California urban and suburban continuation high schools (CHSs). A total of 102 subjects (54 males and 48 females) completed telephone interviews (75 % response rate). The mean age of interview participants was 18.5 (range=17-22, SD=0.93). The sample included 40% Latino, 40% white, 10% African American, and 10% other ethnicity.

Following a standardized questionnaire format, the interviewer asked for the subject's opinions about a potential, new type of free telephone-based referral service for people who participated in Project TND while in high school. Specifically, the subject was told that well-trained telephone educators would call subjects multiple times to provide desired information on resources that people out of high school typically need such as jobs, work training, educational opportunities, transportation, recreational activities, and mental health or drug counseling resources. The educator would attempt to find resources within the subject's local area and zip code. For example, if the subject were interested in finding a job, the educator could discuss the subject's interests and provide some advice and ideas, and then give the subject information on where to go to either get a list of jobs available in the subject's area or, if needed, information on where to go to find out about training for the job. Further, the

educator could also print and send mailed information to the subject about the resources (such as telephone numbers and maps).

Finally, the subjects were told that they could call our resource center as many times as wanted for information on our 1-800 telephone line from 10 a.m.-through 3 p.m., on Mondays through Fridays.

After the subjects were given this brief description of the social service referral telephone program, subjects were asked 14 questions in a 5-minute interview. On scales of 1 (not at all useful)-to-10 (very useful), they were asked about their perceptions of usefulness of the service to people who are recently out of high school; how useful job, education, recreation activities, transportation, mental health counseling, and drug counseling sources of information would be; overall interest in such a service; and perceived helpfulness in preventing people from abusing drugs (including cigarettes, alcohol, marijuana, or harder drugs). Subjects were also asked that if a counselor from the telephone resource center called to give resource information, would they talk to them (Yes/ No/ Not Sure), what an adequate number of calls from the educator might be within a 6 month period (1 time /2 times/3 times/4 times+), whether the subject would ever call this service on their own (Yes/ No/ Not Sure) and, if so, how often (1/2-5/6+ times in 6 months). Finally, subjects were asked if and why this service would help to prevent drug abuse (open-ended responses, multiple responses were permitted, and responses were coded into categories by two raters, with a 95% agreement).

# RESULTS

Subjects were extremely receptive to the concept of the SSRTP. The overall usefulness mean score was 8.67 (SD=1.57). Usefulness by categories ranged from means=8.90 and 9.05, for job and education, respectively (SDs=1.58 and 1.29, respectively), to means of 8.36 and 8.63, for drug counseling resources and transportation, respectively (SDs=2.10 and 1.83, respectively), to mental health and recreation (means=8.00 and 8.06, SDs=2.30 and 2.01, respectively). Subjects reported being very interested in the service (mean=7.83, SD=2.54). Most subjects (76%) said that they definitely would be willing to talk with the telephone educator, and 22% said they were not sure (only 2% said they would not). Interestingly, subjects desired a mean of 2.53 calls every 6 months (SD=1.09). Also, 65% said that they definitely would call the telephone educator on their own as well (on the 800-line), and that they would call an average of 2-5 times, themselves, in a 6-month period.

Finally, 82% of the subjects reported that this service would help prevent drug abuse among themselves and their peers. A total of six categories were coded based on student responses to the open-ended question that assessed why this service might help to prevent drug abuse: (1) it would educate participants on risks of drug abuse (38%); (2) it would help those who are unsure how to ask for or obtain advice, and assist them in self-improvement, autonomy, or mastery (31%); (3) it can provide access to alternative attachments (e.g., jobs) that keep one busy (20%); (4) it can provide access to help for smoking or drug abuse cessation (6%); (5) it would help participants project their life into the future (6%); and (6) it would help participants to think more about others (2%). (Twenty subjects gave more than one response.) Subjects anticipated that the phone calls would take an average of 20 minutes to complete, although they desired having up to an average of 45 minutes of telephone time. All proposed 4 functions of the SSRTP component appeared to be represented; that is, attain mastery, access alternative attachments, cue program information on drug abuse risks, and future orientation/hope (MACH).

Approximately two weeks following the first telephone interview, we telephoned a random selection of 27 of these persons (13 males; 14 females) to assess if our merely contacting previous attendees of the Project TND classroom program regarding social services would prompt them to begin looking for services. Surprisingly, 8 of these subjects had taken action in the last 2 weeks (and attributed this to our call) by either calling a number after looking at job wanted ads (7 subjects), or visiting the local college to pick up a catalogue (2 subjects; 1 person looked for a job and school).

#### DISCUSSION

The results of this needs assessment study identified youths' strong interest and positive feedback regarding the provision of a potential social service resource telephone program (SSRTP) that was designed to offer information to individuals recently out of high school to help integrate them into emerging adulthood. This chapter is the first description, in the form of a needs assessment, of the perceptions and opinions of youth on the information and services that a telephone resource and referral center should provide. The findings are thus a major step forward in using the views of interested parties to provide a new service that may empower people who are at high risk for drug use.

The vast majority of respondents indicated a strong desire for resource and referral information on vocational, educational, recreational, transportation, and mental health and drug counseling services. However, the highest value was placed on vocational and educational information, which is not surprising given their recent transition from the high school context to a young adult "real world" context where they face increasing challenges to become financially independent from their parents. Subjects also reported being very interested in actually utilizing the service, with most indicating that they definitely would be willing to talk with the telephone educator if called, as well as call the telephone educator (on the 800-line) on their own. This reported interest may be due to an important aspect of SSRTP, which is the provision of a tailored service to emerging adults. Information on local and current resources will be provided to each subject based on where they live and what they want to achieve. This localized resource component decreases perceived barriers and increase self-efficacy (Hochbaum, 1958; Rosenstock, 1960). Findings from past research have shown that systematic barriers (i.e. time, location of resources, and transportation) have been a major problem in accessing these resources and hence, compliance (Flores et al., 1998; Fitzpatrick et al., 2004). In addition, providing emerging adults with information on how they can obtain their own information increases the feeling of self-efficacy or the perception that one can perform a certain task or tasks. Programs designed around the self-efficacy framework have been shown effective in enhancing compliance to particular behaviors in diverse fields such as in decreasing use of alcohol and tobacco among adolescents (Ellickson & Bell, 1990; Bell et al., 1993), taking medication (Buchmann, 1997), weight control in patients with end-stage renal diseases (Tsay, 2003), increase in patient-initiated partner notification for curable sexually transmitted infections among adolescents (Fortenberry et al., 2002), and in having healthy diets among adolescents (Ebbeling et al., 2003; Woo et al., 2004).

Another explanation for the strong enthusiasm about this service may be due to the delivery modality. Research has shown that telephone counseling can be a very favorable means in reaching emerging adults (Chen et al., 1995). In general, telephone services have been shown to be effective in improving health care delivery (McBride and Rimer, 1999; Ploeg et al., 2001; Wilkes et al., 2004). Using telephones to deliver interventions allows individualized services to broard sections of the target population as well as minimize logistics and systematic barriers (McBride and Rimer, 1999). Telephone services have also been shown effective when working with teens and tobacco cessation (Chen et al., 1995; Zhu et al., 2000). For example, telephone HELPLINES provided counseling to participants living in California in six different languages. It provided information on tobacco cessation, counseling and self-help kits. Compared to those who did not use the HELPLINES, participants who did were more likely to have tried to quit smoking recently and were ready to try to quit again if the first attempts were not successful. Using peer counselors to deliver important information and support have also been useful (Horton et al., 1997; Goldschmidt and Graves, 2003).

Needs assessment studies for establishing telephone health services have been conducted in the past. For example, Poncia et al. (2000) conducted a needs assessment for a next day telephone follow-up of the elderly after being discharged from the hospital. The results of this assessment revealed a need for improvements in care of elderly patients before and after discharge and thus, the importance of this telephone service (Poncia et al., 2000). The telephone assessment provided an inexpensive alternative to assessing the needs of patients and therefore, potentially decreasing unnecessary return visits to the emergency room. Baker et al. (1999) conducted a needs assessment via an after-hours telephone triage on the advice in private and non-private pediatric populations and found very important implications for improvements of such services (Baker et al., 1999). Thus far, it can be seen that the telephone needs assessment conducted in this current study plays a very critical role in development of the SSRTP program, especially among our target population. The SSRTP has the potential to provide time- and cost-efficient dissemination of imrtant information to emerging adults in a manner most suitable for them. Since many will be out of high school at the time of the telephone intervention, they are not easily reached and followed. Mailing information is difficult as students tend to move out of their parent's home and establish their own residence elsewhere. Mailing can also be costly when considered over long periods of time and is prone to getting lost or misplaced. Needs assessment of other young adults have been in favor of telephone interventions (Chen et al., 1995).

Interestingly, another indication of the great potential of this telephone service to influence high risk youth to engage in productive activities has been evidenced in the results of this study. That is, a follow-up survey to the initial needs assessment of the SSRTP has shown almost immediate positive feedback towards the telephone service. Our findings indicate that by just implementing the initial needs assessment study, several students (almost one-third of the small follow-up sample) were prompted to seek job and educational information on their own. These strong, albeit limited, data provide promising implications for significant program effects in the future.

Furthermore, participants in this study also provided open-ended, qualitative responses indicating that they believed that the Project TND prevention education offered in the

classroom in high school plus the addition of this new continuing social resource service (i.e., the SSRTP) might be helpful in preventing people from using drugs (including cigarettes, alcohol, marijuana, or harder drugs). These data lend support to our hypothesis that a telephone booster component to Project TND will help emerging adults to integrate successfully into main stream adult lifestyles, while avoiding substance use. Specifically, it is feasible that including an SSRTP booster component could bolster preventive effects on drug use over a longer period of time, through review of drug abuse prevention program material, and by assisting former CHS youth better prepare themselves during emerging adulthood by helping with at least four ways which correspond to the MACH (mastery, attachments, cue previous material, and hope) model introduced earlier.

# LIMITATIONS AND FUTURE DIRECTIONS

Our results should be interpreted in the context of a few methodological limitations. First, data in this study were generated from self-report surveys, the accuracy of which could not be independently verified. Thus, it is impossible to assess the extent to which such data may be biased, particularly in regard to social desirability. It is also important to understand that due to difficulty in contacting students for surveys, it eventually became a selective group of individuals who answered the survey questions. That is, only those who were willing to talk to us were the ones providing responses. Thus, it is likely that they would provide positive responses in support of our service. Those that may provide lower scores were those that declined, and thus, were not surveyed. Second, the results of this study are generalizable only to subjects who are similar to those examined in this study. Continuation high school students differ in many important ways from general population youth (Sussman et al., 1995; 1995) Also, this sample was quite heterogeneous ethnically. It is possible that these results differ from other, more ethnically homogeneous populations of youths.

Third, there was no comparison group for the findings presented in this study. However, results from this study were primarily intended to represent a preliminary needs assessment of our targeted sample to be used for the first phase of development of the SSRTP.

Despite these limitations, this study provides new evidence which suggests that continuation high school students have a strong interest in utilizing a potential social service resource telephone program (SSRTP) that was designed to offer information to individuals recently out of high school to help integrate them into emerging adulthood. The results provide feedback on how to better structure and tailor such a prevention intervention program. For example, the results reveal exactly for which topics (e.g., vocational and educational) the participants are interested in receiving information. Other programs with similar missions of developing a tailored social service referral and resource intervention may benefit from these findings.

Finally, it should be noted that he SSRTP intervention cannot stand alone; it is only a component of Project TND and serves as a supplement to the program. However, it is expected that this component will boost program effects and serve to enhance the lifestyles of emerging adults by reminding them of what was learned while in school and assist them in applying it to real life situations. Thus, once the SSRTP is developed, further experimental research is needed to investigate the effectiveness of the provision of the social service resource

information component, as well as the Project TND classroom-based component on drug use among emerging adults.

# REFERENCES

- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55, 469-480.
- Aseltine, R. H., & Gore, S. L. (2000). The variable effects of stress on alcohol use from adolescence to early adulthood. *Substance Use & Misuse, 35,* 643-668.
- Aubert, R. E., Fulop, G., Xia, F., Thiel, M., Maldonato, D., & Woo, C. (2003). Evaluation of a depression health management program to improve outcomes in first or recurrent episode depression. *American Journal of Managed Care*, 9, 374-380.
- Bachman, J., Wadsworth, K., O'Malley, P., Johnston, L., & Schulenberg, J. (1997). Smoking, Drinking, and Drug use in Young Adulthood: The Impacts of New Freedoms and New Responsibilities. Mahway, NJ: Lawrence Erlbaum Associates.
- Baker, R. C., Schubert, C. J., Kirwan, K. A., Lenkauskas, S. M., & Spaeth, J. T. (1999). After-hours telephone triage and advice in private and nonprivate pediatric populations. *Archives of Pediatrics & Adolescent Medicine*, 153(3), 292-296.
- Bell, R. M., Ellickson, P. L., & Harrison, E. R. (1993) Do drug prevention effects persist into high school? How project ALERT did with ninth graders. *Preventive Medicine*, 22, 463-483.
- Billings, J., & Cowley, S. (1995). Approaches to community needs assessment: A literature review. *Journal of Advance Nursing*, 22(4), 721-730.
- Brook, J. S., Balka, E. B., Gursen, M. D., Brook, D. W., & Shapiro, J. (1997). Young adults' drug use: A 17-year longitudinal inquiry of antecedents. *Psychological Reports*, 80, 1235-1251.
- Brook, D. W., Brook, J. S., Zhang, C., Cohen, P., & Whiteman, M. (2002). Drug use and the risk of major depressive disorder, alcohol dependence, and substance use disorders. *Archives of General Psychiatry*, 59, 1039-1044.
- Brook, J. S., Whiteman, M., Cohen, P., Shapiro, J., & Balka, E. (1995). Longitudinally predicting late adolescent and young adult drug use: Childhood and adolescent precursors. *Journal of the American Academy of Child Adolescent Psychiatry*, 34, 1230-1238.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*, 513-531.
- Buchmann, W. (1997). Adherence: a matter of self-efficacy and power. *Journal of Advance Nursing*, *26*(1), 132-137.
- Champion, V. L, Skinner, C. S., & Foster, J. L. (2000). The effects of standard care counseling or telephone/in-person counseling on beliefs, knowledge, and behavior related to mammography screening. *Oncology Nursing Forum*, 27, 1565-1571.
- Chassin, L., Pitts, S. C., & DeLucia, C. (1999). The relation of adolescent substance use to young adult autonomy, positive activity involvement, and perceived competence. *Development and Psychopathology*, *11*, 915-932.

- Chen, S. P., Dallas, C., & Chen, E. H. (1995). Teen voices on a proposed telephone health service. *ABNF Journal*, 6(1), 19-23.
- Coalition for Community Schools. (2004). *Evaluation of Community Schools: An Early Look.* Retrieved September 2, 2004 from http://www.communityschools.org/ evaluation/evalbrieffinal.html.
- Cohen, S. (1988). Pyschosocial models of the role of social support in the etiology of physical disease. *Health Psychology*, 7, 269-297
- Curry, S. J., Ludman, E. J., Grothaus, L. C., Donovan, D., & Kim, E. (2003). A randomized trial of a brief primary-care-based intervention for reducing at-risk drinking practices. *Health Psychology*, 22, 156-165.
- Dent, C. W., Sussman, S., McCullar, W. J., Stacy, A. W. (2001). Project Towards No Drug Abuse: Generalizability to a general high school sample. *Preventive Medicine*, 32, 514-520.
- Ebbeling, C. B., Leidig, M. M., Sinclair, K. B., Hangen, J. P., & Ludwig, D. S. (2003). A reduced-glycemic load diet in the treatment of adoelscent obesity. *Pediatrics & Adolescent Medicine*, 157(8), 773-779.
- Eccles, J., Templeton, J., Barber, B., & Stone, M. (2003). Adolescence and emerging adulthood: The critical passage ways to adulthood. In M. H. Bornstein, L. Davidson, C. L. M. Keys & K. A. Moore (Eds.), *Well-being: Positive Development Across the Life Course* (pp. 383-406). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Elder, J. P., Woodruff, S. I., & Eckhardt, L. (1994). Participation in a telephone-based tobacco use prevention program for adolescents. *American Journal of Health Promotion*, 9, 92-95.
- Ellickson, P. L., & Bell, R. M. (1990). Drug prevention in junior high: a multi-site longitudinal test. *Science*, 247, 1299-1305.
- Fitzpatrick, A. L., Powe, N. R., Cooper, L. S., Ives, D. G., & Robbins, J. A. (2004). Barriers to health care access among the elderly and who perceives them. *American Journal of Public Health*, 94(10), 1788-1794.
- Flores, G., Abreu, M., Olivar, M. A., & Kastner, B. (1998). Access barriers to health care for Latino children. Archives of Pediatrics & Adolescent Medicine, 152(11), 1119-1125.
- Forrest, S., Strange, V., Oakley, A., & The RIPPLE Study Team. (2004). What do young people want from sex education? The results of a needs assessment from a peer-led sex education programme. *Culture, Health & Sexuality, 6*(4), 337-354.
- Fortenberry, J. D., Brizendine, E. J., Katz, B., P., & Orr, D., P. (2002). The role of selfefficacy and relationship quality in partner notification by adolescents with sexually transmitted infections. *Archives Pediatric & Adolescent Medicine*, 156(11), 1133-1137.
- Gilbody, S., House, A., & Sheldon, T. (2002). Routine administration of Health Related Quality of Life (HRQoL) and needs assessment instruments to improve psychological outcome - a systematic review. *Psychological Medicine*, 32(8), 1345-1356.
- Goldschmidt, R. H. & Graves, D. W. (2003). The National HIV Telephone Consultation Service (Warmline): a clinical resource for physicians caring for African-Americans. *Journal of the National Medical Association*, 95(2 Suppl 2), 8S-11S.
- Grunbaum, J. A. Kann, L., Kinchen, S. A., Ross, J. G., Gowda, V. R., Collins, J. L., & Kolbe, L. J. (2000). Youth risk behavior surveillance National Alternative High School Youth Risk Behavior Survey, United States, 1998. *Journal of School Health*, 70, 5-17.

- Green, L. W., Richard, L., & Potvin, L. (1996). Ecological foundations of health promotion. *American Journal of Health Behavior*, 10, 270-281.
- Guo, J., Hawkins, J. D., Hill, K. G., & Abbott, R. D. (2001). Childhood and adolescent predictors of alcohol abuse and dependence in young adulthood. *Journal of Studies on Alcohol*, 62, 754-762.
- Hawkins, J., Lishner, D., Jenson, J., & Catalano, R. (1987). Delinquents and drugs: What the evidence suggests about prevention and treatment programming. *Youth at High Risk for Substance Abuse*. Superintendent of Documents, U.S. Government Printing Office, DHHS ADM-D87-1537 (pp. 81-131). Washington, DC: National Institute on Drug Abuse.
- Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance use prevention. *Psychological Bulletin*, 112, 64-105.
- Hochbaum, G. (1958). Public Participation in Medical Screening Programs: A Sociopsychological Study. PHS Publication No. 572. Washington, DC:U.S. Government Printing Office.
- Horton, R., Peterson, M. G. E., Powell, S., Engelhard, E., & Paget, S. A. (1997). Users evaluate lupusline a telephone peer counseling service. *Arthritis Care &Research*, 10(4), 257-263.
- Hou, S. I, Fernandez, M. E., & Parcel, G. S. (2004). Development of a cervical cancer educational program for Chinese women using intervention mapping. *Health Promotion Practice*, 5, 80-87.
- Kandel, D. B., Davies, M., Karus, D., & Yamaguchi, K. (1986). The consequences in young adulthood of adolescent drug involvement. *Archives of General Psychiatry*, 43, 746-754.
- Krohn, M. D., Lizotte, A. J., & Perez, C. M. (1997). The interrelationship between substance use and precocious transitions to adult statuses. *Journal of Health and Social Behavior*, 38, 87-103.
- Laken, M. P. & Ager., J. W. (1996). Effects of case management on retention inn prenatal substance abuse treatment. *American Journal of Drug & Alcohol Abuse*, 22, 439-448.
- Leed-Kelly, A., Russel, K. S., Bobo, J. K., & McIlvain, H. (1996). Feasibility of smoking cessation counseling by phone with alcohol treatment center graduates. *Journal of Substance Abuse Treatment*, 13, 203-210.
- Lesser, J. & Escoto-Lloyd, S. (1999). Health-related problems in a vulnerable population: pregnant teens and adolescent mothers. *Nursing Clinics of North America*, 34, 289-299.
- Maggs, J. L., Frome, P. M., Eccles, J. S., & Barber, B. L. (1997). Psychosocial resources, adolescent risk behaviour and young adult adjustment: is risk taking more dangerous for some than others? *Journal of Adolescence*, 20, 103-119.
- Man, D. W. K., Lee, E. W. T., Tong, E. C. H., Yip, S. C. S., Lui, W. F., & Lam, C. S. (2004). Health services needs and quality of life assessment of individuals with brain injuries: a pilot cross-sectional study. *Brain Injury*, 18(6), 577-591.
- Mason, M. J., Cheung, I., & Walker, L. (in press). Substance use, social networks and the geography of risk and protection of urban adolescents. *Substance Use & Misuse*.
- McBride, C. M. & Rimer, B. K. (1999). Using the telephone to improve health behavior and health service delivery. *Patient Education & Counseling*, *37*(1), 3-18.

- McCuller, W. J., Sussman, S., Holiday, K., Craig, S., & Dent, C. W. (2002). Tracking procedures for locating high risk youth. *Evaluation and the Health Professions*, 25, 345-362.
- Mermelstein, R., Hedeker, D., & Wong, S. C. (2003). Extended telephone counseling for smoking cessation: does content matter? *Journal of Consulting & Clinical Psychology*, 71, 565-74.
- Miguez, M. C., Vazquez, F. L. & Becona., E. (2002). Effectiveness of telephone contact as an adjunct to a self-help program for smoking cessation: a randomized controlled trial in Spanish smokers. *Addictive Behaviors*, 27, 139-44.
- Millstein, S.G. (1994). Health promotion and disease prevention in the 21<sup>st</sup> century: Time for a paradigm shift? Paper presented at the State of the Art Symposium: Innovative Prevention Approaches for the 21<sup>st</sup> Century, American Psychological Association 1994 Annual Meeting, Boston, MA.
- Morrow, K. & Costello, T. (2004). HIV, STD and hepatitis prevention among women in methadone maintenance: a qualitative and quantitative needs assessment. *AIDS Care,* 16(4), 426-433.
- Newcomb, M. D. & Bentler, P. M. (1988). Consequences of Adolescent Drug Use. Newbury Park, CA: Sage Publications, Inc.
- Newcomb, M. D. & Bentler, P. M. (1988b) Impact of adolescent drug use and social support on problems of young adults: A longitudinal study. *Journal of Abnormal Psychology*, 97, 64-75.
- Nies, M. A., Chruscial, H. L., & Hepworth, J. T. (2003). An intervention to promote walking in sedentary women in the community. *American Journal of Health Behavior*, 27, 524-535.
- O'Leary, T. A., Brown, S. A., Colby, S. M., Cronce, J. M., D'Amico, E. J., Fader, J. S., Geisner, I. M., Larimer, M. E., Magges, J. L., McGrady, B., Palomer, R. S., Schulenberg, J., & Monti, P. M. (2002). Treating adolescents together or individually? Issues in adolescent substance abuse interventions. *Alcoholism: Clinical and Experimental Research*, 26, 890-899.
- Pierce, J. P., Newman, V. A., Flatt, S. W., Faerber, S., Rock, C. L., Natarajan, L., Caan, B. J., Gold, E. B., Hollenbach, K. A., Wasserman, L., Jones, L., Ritenbaugh, C., Stefanick, M. L., Thomson, C. A., & Kealey, S. (2004). Telephone counseling intervention increases intakes of micronutrient- and phytochemical-rich vegetables, fruit and fiber in breast cancer survivors. *Journal of Nutrition*, 134, 452-458.
- Pinto, B. M., Friedman, R., Marcus, B. H., Kelley, H., Tennstedt, S., & Gillman, M. W. (2002). Effects of computer-based, telephone-counseling system on physical activity. *American Journal of Preventative Medicine*, 23, 113-120.
- Ploeg, J., Biehler, L., Willison, K., Hutchison, B., & Blythe, J. (2001). Perceived support needs of family caregivers and implications for a telephone support service. *Canadian Journal of Nursing Research*, 33(2), 43-61.
- Poncia, H., Ryan, J., & Carver, M. (2000). Next day telephone follow-up of the elderly: a needs assessment and critical incident monitoring tool for the accident and emergency department. *Journal of Accident Emergency Medicine*, 17, 337-340.
- Rainbird, K. J., Perkins, J. J., & Sanson-Fisher, R. W. (2004). The Needs Assessment for Advanced Cancer Patients (NA-ACP): A measure of the perceived needs of patients with

advanced, incurable cancer. A study of validity, reliability and acceptability. *Psychooncology*, *14*(4), 297-306.

- Rohrbach, L. A., Sussman, S., Dent, C. W., & Sun, P. (2005). A five-year longitudinal study from adolescence to emerging adulthood. *Journal of Drug Issues*, *35*(3), 333-356.
- Rosenstock, I. (1960). What research in motivation suggests for public health. *American Journal of Health Promotion*, 50, 295-301.
- Scotti, J. R., Morris, T. L., McNeil, C. B., & Hawkins, R. P. (1996). DSM-IV and disorders of childhood and adolescence: can structural criteria be functional? *Journal of Consulting & Clinical Psychology*, 64, 1177-91.
- Skara, S. & Sussman, S. (2003). A review of 25 long-term adolescent tobacco and other drug use prevention program evaluations. *Preventive Medicine*, *37*, 451-474.
- Stacy, A. W. (1995). Memory association and ambiguous cues in models of alcohol and marijuana use. *Experimental and Clinical Psychopharmacology*, 3, 183-194.
- Stacy, A. W., Ames, S., Sussman, S., Dent, C. W. (1996). Implicit cognition in adolescent drug use. *Psychology of Addictive Behaviors*, 10, 190-203.
- Stacy, A. W., Dent, C. W., Sussman, S., Raynor, A., Burton, D., & Flay, B. R. (1990a). Expectancy accessibility and the influence of outcome expectancies on adolescent smokeless tobacco use. *Journal of Applied Social Psychology*, 20, 802-817.
- Stacy, A. W., Widaman, K. F., & Marlatt, G. A. (1990). Expectancy models of alcohol use. Journal of Personality and Social Psychology, 58, 918-928.
- Stein, J. A., Newcomb, M. D., & Bentler, P. M. (1986). Stability and change in personality: A longitudinal study from early adolescence to young adulthood. *Journal of Research in Personality*, 20, 276-291.
- Stevens, A. & Gillam S. (1998). Needs Assessment: from theory to practice. British Medical Journal, 316(7142), 1448-1452.
- Sun, W., Skara, S., Sun, P., Dent, C. W., & Sussman, S. (2006). Project Towards No Drug Abuse: Long-term substance abuse use outcomes evaluation. *Preventive Medicine*, 42, 188-192.
- Sussman, S. (2001). *Handbook of Program Development for Health Behavior Research and Practice*. Thousand Oaks, CA: Sage.
- Sussman, S. & Dent, C. W. (2004). Five-year prospective prediction of marijuana use cessation among youth at continuation high schools. *Addictive Behaviors*, 29, 1237-1243.
- Sussman, S., Dent, C. W., & Galaif, E.R. (1997). The correlates of substance abuse and dependence among adolescents at high risk for drug abuse. *Journal of Substance Abuse*, 9, 241-255.
- Sussman, S., Dent, C. W., Simon, T.R., Stacy, A. W., Galaif, E. R., Moss, M. A., Craig, S., & Johnson, C. A. (1995). Immediate impact of social influence-oriented substance abuse prevention curricula in traditional and continuation high schools. *Drugs and Society*, 8(3/4), 65-81.
- Sussman, S., Dent, C. W., & Stacy, S. (2002). Project Towards No Drug Abuse: A review of the findings and future directions. *American Journal of Health Behavior*, 26, 354-365.
- Sussman, S., Dent, C. W., Stacy, A., & Craig, S. (1998). One-year outcomes of Project Towards No Drug Abuse. *Preventive Medicine*, 27, 632-642 (Erratum: 27(5, Pt.1), 766, 1998).

- Sussman, S., Earleywine, M., Wills, T., Cody, C., Biglan, A., Dent, C. W., & Newcomb, M. D. (2004). The motivation, skills, and decision-making model of "drug abuse" prevention. Substance Use & Misuse, 39(10-12), 1971-2016.
- Sussman, S., Stacy, A. W., Dent, C.W., Simon, T. R., Galaif, E. R., Moss, M. A., Craig, S., & Johnson, C. A. (1995) Continuation high schools: youth at risk for drug abuse. *Journal of Drug Education*, 25(3), 191-209.
- Sussman, S., Sun, P., McCuller, W. J., & Dent, C. W. (2003). Project Towards No Drug Abuse: Two-year outcomes of a trial that compares health educator delivery to selfinstruction. *Preventive Medicine*, 37,155-162.
- Tsay, S. (2003). Self-efficacy training for patients with end-stage renal disease. *Journal of Advance Nursing*, 43(4), 370-375.
- Tutty, S., Simon, G., & Ludman, E. (2000). Telephone counseling as an adjunct to antidepressant treatment in the primary care system. A pilot study. *Effective Clinical Practice*, 3, 170-8.
- Ward, D. A., Bendel, R. B., & Lange, D. (1982). A reconsideration of environmental resources and the posttreatment functioning of alcoholic patients. *Journal of Health & Social Behavior*, 23, 310-7.
- Whelan, V., McBride, D., & Colby, R. (1993). Public health department tracking of high-risk drug users. *Public Health Reports*, 108, 643-645.
- Wilkes, L., Mohan, S., White, K., & Smith, H. (2004). Evaluation of an after hours telephone support service for rural palliative care patients and their families: A pilot study. *Australian Journal Rural Health*, 12(3), 95-8.
- Woo, K., Chook, P., Yu, C. W., Sung, R. Y. T., Qiao, M., Leung, S. S. F., Lam, C. W. K., Metreweli, C., & Celermajer, D. S. (2004). Effects of Diet and Exercise on Obesity-Related Vascular Dysfunction in Children. *Circulation*, 109(16), 1981-1986.
- Wood, I., Douglas, J., & Priest, H. (2004). Education and training for acute care delivery: a needs analysis. *Nursing in Critical Care*, 9(4), 159-166.
- Wright, J., Williams, R., & Wilkinson, J, R. (1998). Health needs assessment: Development and importance of health needs assessment. *British Medical Journal*, 316(7140), 1310-1313.
- Zhu, S. H., Anderson, C. M., Johnson, C. E., Tedeschi, G., & Roeseler, A. (2000). A centralised telephone service for tobacco cessation: the California experience. *Tobacco Control*, 9(Suppl 2), II48-II55.
- Zhu, S. H., Tedeschi, G., Anderson, C. M., Rosbrook, B., Byrd, M., Johnson, C. E., & Gutierrez-Terrell, E. (2000). Telephone counseling as adjuvant treatment for nicotine replacement therapy in a "real-world" setting. *Preventive Medicine*, 31, 357-63.

Chapter 18

# CAUSES, IMPACT, PREVENTION AND MEDICOLEGAL ASPECTS OF CHILD ABUSE AND NEGLECT<sup>\*</sup>

# **B. R.** Sharma<sup> $\dagger$ </sup>

Department of Forensic Medicine and Toxicology, Govt. Medical College and Hospital, Chandigarh, India

# ABSTRACT

Abuse and neglect of children and adolescents continue to be major public health problems all over the world. It has been found to be correlated with emotional and behavioral problems, suicidal behavior, substance abuse, and many psychiatric disorders including post-traumatic stress disorder. However, despite the high prevalence and the mental health consequences of childhood abuse for the victims, insufficient efforts are being made to recognize these consequences and there are too few resources available to provide the needed mental health services to these victims and their families. On the other hand, exciting advances have taken place regarding care of sexual abuse victims. Clinical research has demonstrated the effectiveness of psychotherapies and suggested the efficacy of medications in ameliorating the behavioral and emotional consequences of childhood abuse. Specific clinician interviewing strategies have been developed to avoid suggestion and interference with sexual abuse-associated litigation. However, psychiatry and allied mental health professionals have much to contribute in terms of advocacy for abuse victims and an increased understanding of the need for provision of rehabilitation opportunities to these children. Medical professionals can also contribute to lower the high acquittal rate in cases of child sexual abuse by proper documentation of the history and examination findings.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence: Dr. B. R. Sharma, #1156 – B, Sector – 32 B, Chandigarh – 160030. India. E-mail: drbrsharma@yahoo.com and drbrsharma@gmail.com

#### INTRODUCTION

Child abuse and neglect is a terrible misfortune for millions of children and families, for communities, and for society. It affects children of all colors, social classes, ethnic groups, and religions. It damages young people in infancy, childhood and adolescence. The term child abuse refers to any act or failure to act that violates the rights of the child that endangers his or her optimum health, survival or development. Definitions of child abuse inevitably vary according to the uses for which they have been devised. There ought to be differences of emphasis between definitions that are intended for legal purposes, clinical diagnostic purposes or to suit the needs of a research study. Cultural relativity may be a significant issue, for certain behaviors that are construed as abusive by some cultures or subcultures may not be regarded in the same light by others. Definitions of abusive behaviors that are essentially the acts of commission may need to be distinguished from neglectful care taking, which is characterized by the acts of omission. Neglect may refer to deprivation or non-provision of necessary and societally available resources due to proximate and proscribed human actions that create the risk of permanent impairment to development or functioning. Child abuse or maltreatment is commonly divided into five categories; physical abuse, emotional abuse, sexual abuse, neglect and exploitation. Although any of these forms may be found separately, but more often they occur together.

According to a World Health Organization estimate about 40 million children aged 0 - 14 around the world suffer form abuse and neglect and require health and social care [1]. In 2000, nearly 2 million reports of alleged child abuse or neglect were investigated by child protective services agencies, representing more than 2.7 million children who were alleged victims of maltreatment and who were referred for investigation [2]. Of these children, approximately 879,000 were found to be victims of maltreatment. This reflects a national rate of approximately 12.2 children per 1,000 children younger than 18 years of age in the general population who were found to be substantiated or indicated victims of maltreatment. Nearly two-thirds of child victims (62.8%) suffered neglect, including medical neglect, while nearly one-fifth (19.3%) suffered physical abuse and approximately 10 percent suffered sexual abuse.

The 2000 Annual Report from the National Child Abuse and Neglect Data System (NCANDS) indicates that national child abuse incidence rates increased in each year from 1990-1993, and decreased in each year through 1999. The rate of victimization was 13.4 per 1,000 children in 1990. The rate peaked at 15.3 children per 1,000 in 1993, then decreased to 11.8 per 1,000 in 1999, while increasing slightly to 12.2 children per 1,000 in 2000 [3]. Meanwhile, the Third National Incidence Studies (NIS-3) found that rates of child maltreatment under the *Harm* Standard increased 149 percent from the time the first NIS study (NIS-1) was conducted [4]. In the 2003-2004 financial year, Kids Help Line counselors responded to 452 contacts from children and young people in Western Australia with concerns about child abuse. Among those seeking help, 50% cases were that of physical abuse, 38% of sexual abuse, 6% emotional abuse, and 5% of neglect [5].

However, available figures on the incidence of child abuse are undoubtedly underestimates for three main reasons. Firstly, there is often failure to detect the signs and symptoms that are indicative of abuse, especially if these are psychological rather than

477

physical in nature. Secondly, even if the signs and symptoms are recognized it may not be realized that they result from abuse. Thirdly, even if the true etiology is suspected, the case may not be reported and so cannot enter into any official statistics. Furthermore, recognition and diagnosis of child abuse are compromised by denial that pervades all aspects of the phenomenon: perpetrators routinely deny that they did it; abused children sometimes deny that any thing had happened to them; otherwise concerned adults, such as neighbors, teachers, and emergency department physicians, may overlook and underreport abuse for any number of personal reasons; complex legal requirements may make it hard to convict the abusers; and generalized societal denial protects people from awareness of an unpleasant side of life.

Child abuse and neglect affects the healthy and normal course of development. It causes deviation from an expected trajectory, preventing the developing child's negotiation of sequential tasks and disrupting the normal transaction between different facets of development. The evaluation and management of child abuse require cooperation of diverse professional groups, including pediatricians and other primary care physicians, emergency department staff, radiologists, pathologists, law enforcement agencies, social service personnel, mental health professional, and forensic experts. Therefore child abuse and neglect is the very antithesis of adequate childcare and rearing posing a major threat to human development.

## NEGLECT

Child neglect has been difficult to define [6]. Definitions have varied across disciplines agencies and states, in accordance with differing goals and thresholds [7]. Some authors have proposed broad definitions that incorporate not only caregiver acts and omissions, but also societal and institutional conditions (for example, hunger, lack of health insurance) that affect children [8]. More commonly, child neglect has been narrowly defined as parental or caregiver acts of omission or failure to provide for a child's basic needs. However, this perspective is limited and may not afford adequate protection to children since societal factors (e.g., poverty) that compromise the abilities of parents to care for their children also impair children's health and development and as such 'neglect' must be evaluated within a societal context. Any definition of child neglect must take into account the heterogeneity within the phenomenon. Different forms of neglect, of varying severity and chronicity, and within differing contexts, require a range of responses tailored to the individual situation. For example, a parent's ignorance of a child's nutritional needs requires a very different response from protecting children from lead or any other poisonous substance in the environment. On the other hand, it is also important to recognize the variability among children and their response to specific situations. Being a latchkey child might be appropriate for a mature child with adequate neighborhood supports, but endangering for another [9].

Child neglect, thought to be the most common of the five categories of maltreatment and probably the most life threatening may be further classified as: 1) Physical as for example, lack of appropriate supervision or failure to provide necessary food, shelter, or medical care. 2) Educational as for example, failure to educate a child or attend to special education needs. 3) Emotional as for example, inattention to a child's emotional needs or exposure to domestic violence, however, these situations do not always mean that a child is neglected. Sometimes

cultural values, the standards of care in the community, and poverty may be contributing factors, indicating that the family is in need of information or assistance. When a family fails to use information and resources, and the child's needs continue to be unmet, then further child welfare professional intervention may be required.

In considering the impact of a situation on a child and the possibility of neglect, several factors need to be evaluated: whether actual or potential harm occurred, the severity of harm involved, and the frequency / chronicity of the behavior / event(s). In current practice, neglect is often considered only when actual harm has occurred, however, in some situations, the potential for moderate to serious harm needs to be an important consideration, as many forms of neglect have no immediate physical consequences, although there may be substantial and long-term psychological harm [10]. Severity of neglect is typically based on the degree of harm involved. A serious injury is apt to be seen as more severe neglect than a minor injury and any injury is likely to be seen as more severe than a potential injury. Pattern of omissions in care has been an important criterion of neglect, although single or occasional lapses in care are often considered 'only human' and are not regarded as neglect. But, even a single omission in care can have devastating results, such as an unattended toddler drowning in a swimming pool. In contrast, some omissions in care are unlikely to be harmful unless they are recurrent. For example, for a child to miss asthma medications occasionally may involve little risk, but that risk is far greater if medications are missed repeatedly.

# **PHYSICAL ABUSE**

Physical abuse is the physical assault of a child by any person having custody, care, or charge of that child. It includes physical injury (ranging from minor bruises to severe fractures or death) as a result of punching, beating, kicking, biting, shaking, throwing, stabbing, choking, hitting (with hand, stick, strap, or other object), burning, or otherwise harming a child. Such injury is considered abuse regardless of whether the caretaker intended to hurt the child or not. Cruelty towards children has been described throughout history, with children viewed as parental property, however, the problem, was unveiled by Silverman (1953) and Wooley and Evans (1955) in its exact shape, magnitude and significance, who established the deliberate trauma character of certain specific types of pathological lesions [11], earlier detected by J. Caffey, a pediatric radiologist (Caffey's Syndrome) [12].

Physical abuse occurs at all ages, although biological sequelae are more severe in infancy. There is no association with ethnic groups, but a strong one with low socioeconomic status among the under fives, becoming weaker throughout the childhood and disappearing by adolescence [13]. Children with developmental disabilities have a raised risk [14]. Families in which physical abuse occurs are more likely to support mutually abusive coercive communications and interactions than controls. Partner abuse and domestic violence is relatively more common, combined with evasive hostility and decreased cohesion. Discussion, positive display of affection, and encouragement of prosocial behaviors are less common than in non-maltreating families [15].

The Third National Incidence Study which sampled child protective services, law enforcement, juvenile probation, public health, hospital, school, day care, and mental health and social services agencies for the child maltreatment prevalence data, reported that physical abuse had increased to 5.7 per 1000 children as compared to 4.3 per 1000 and 3.1 per 1000 reported by Second and the First National Incidence Study respectively [4]. According to a U.S. Department of Health and human Services report, there were 903,000 victims of child maltreatment nationwide, of which, neglect involved 53.5%, and physical abuse involved 22.7% of cases [3].

Physical abuse is detected through the observation of physical injuries without an alternative non-abusive explanation [16]. Usually, the diagnosis is based upon a discrepancy between the physical findings and the history provided. When an explanation is forthcoming, trigger events or developmental challenges are common – for example, persistent crying in infancy, problem of toileting or feeding among toddlers, or issues of discipline in later childhood. There may have been previous episodes of similar or lesser concern.

In the usual course of prognosis, some physically abused children may have neurological and other physical sequalae as a result of their injuries [17]. Educational difficulties are consistently found on follow-up. These children are less sensitive to social cues and less skilful at managing personal problems. Their capacity for emphatic concern with the everyday problems of their peers becomes blunted. Not surprisingly therefore, chronic oppositional and aggressive behavior is the most consistently documented outcome. These children may range from the socially withdrawn and avoidant to those who demonstrate fear, anger, and aggression [13]. Approximately 20 to 30% of physically abused children become delinquent in their teenage years. They are at increased risk of running away from home, alcohol and illicit drug use, self-destructive behavior, suicide, teenage pregnancy, etc [15].

#### **SEXUAL ABUSE**

Child sexual abuse was once regarded as the 'hidden pediatric problem' that Kempe referred to in 1977. By the latter half of nineteenth century, three French physicians, Tardeu, Bernard, and Brouardel, had each tried to expose the problem. Near the turn of the twentieth century, Sigmund Freud initially tried to uncover it, but like children who attempt to disclose, he recanted and created a myth that children's accounts of sexual abuse were the result of childhood desires and fantasy. This myth was accepted more readily than was the truth. In the course of last three decades, we have witnessed awakening concern, increasing public awareness, and organized backlash about child sexual abuse. Statements like "children never lie" and "children always lie" have been shown to be myths. Management of sexual abuse began as an art, and the scientific basis for the art continues to be laid down.

There is no universal definition of child sexual abuse. Most definitions cover a wide range of sexual activities, including situations with no physical contact. Kemp [18] defined child sexual abuse as the involvement of children and adolescents in sexual activities they do not understand, to which they cannot give informed consent, or that violate social taboos. Some definitions emphasize the unwanted, manipulative and exploitative factors while recognizing the importance of age and developmental level differences between participants [19, 20]. A simple definition of child sexual abuse, however, should include activities by a parent or caretaker such as fondling a child's genitals, penetration, incest, rape, sodomy, indecent exposure, and commercial exploitation through prostitution or the production of pornographic materials [21]. The definition of child and adolescent sexual abuse used in the

Third National Incidence study [4] involves a child younger than 18 years of age having experienced one of the following types of sexual acts: intrusion, defined as evidence of oral, anal, or genital penile penetration or anal or genital digital or other penetration; molestation with genital contact, but without evidence of intrusion; or other acts that involved contact with non-genital areas of a child's body (e.g., fondling of breasts or buttocks, exposure) or inadequate or inappropriate supervision of sexual activities when the perpetrator was a parent, parent substitute, or other.

Child sexual abuse is the product of complex factors involving the victims, the perpetrator, and the environment. Intrafamilial sexual abuse or incest is fostered by an environment that allows poor supervision or poor choice of surrogate caretakers, or fails to set appropriate sleeping and role boundaries [22]. Likewise, extra-familial sexual abuse is fostered by factors that increase the child's vulnerability. Children with unmet emotional needs are easier to victimize. Adolescence defiance and peer pressure place some children at increased risk of sexual abuse. The risk-taking child is even less likely to disclose the abuse out of the fear of blame or punishment, or concern about angering or upsetting the peer group [23]. Finkelhor [24] described four preconditions for sexual abuse. The first precondition is an abuser whose motivation to abuse children comes from 'emotional congruence', often secondary to abuse a child, sexual arousal by children, and the inability to have appropriate sexual relationships with peers. The second precondition is the ability of the abusers to overcome their own internal inhibitions or moral standards to abusing children. The third precondition requires overcoming external inhibitors to abuse, such as the protective parent or normal boundaries between family members or between children and adult in general. The fourth precondition is overcoming the resistance of the child through use of pressure, seduction, or coercion.

The process of sexual victimization usually occurs in the context of relationship and is accompanied by behaviors designed to engage the child in the sexual activity and permit the abuse to go on over time. Initially, the child is targeted for victimization, and a non-sexual relationship is established. Typically, the sexualization of the relationship appears to take place gradually. The child's cooperation and silence are maintained through a variety of forms of coercion, often by exploiting a child's normal need to feel valued [22]. Summit [25] used the term 'child sexual abuse accommodation syndrome' to describe the process by which the perpetrator gains access to the child, initiates the abuse, and assures cooperation and secrecy using threats or rewards. The child accommodates to the increased sexual demands with an increasing sense of betrayal and feelings of guilt, and may develop behavioral symptoms. The children who disclose often find an unsupportive response, which may lead to repeated attempts at partial disclosure. If the unsupportive reactions continue, the child retracts the complaint of abuse. The child's false retraction is quickly accepted, leaving the child vulnerable to continued abuse.

Victims of child sexual abuse may present in many ways. Some children are brought for medical evaluation after the disclosure of abuse. Masked presentations, however, are more common. Some children present initially with physical or behavioral complaints and, on further investigation, the history of sexual abuse is obtained. Common masked complaints include genital symptoms, abdominal pain, constipation or rectal bleeding, straddle injury, and various other somatic and behavioral problems [26, 27]. Chronic or recurrent urinary tract infections may be another masked presentation [28], however, symptoms mimicking a urinary

tract infection are more common than actual urinary tract infection in sexually abused children [29].

A child's direct statement describing sexual abuse is the most definitive historical indicator of abuse. The child may also provide a history of abuse during evaluation of a particular physical problem or behavioral problem. Both types of disclosures should be taken seriously and should be investigated. Physical indicators of abuse may be used to corroborate the history of abuse. Common behaviors, often categorized as behavioral indicators of child sexual abuse, although may reflect the child's response to any emotional conflict, but should not be considered diagnostic of sexual abuse. However, they may be considered supportive evidence of sexual abuse [30, 31].

In the course of prognosis, a wide range of psychological sequlae in childhood and adult life are associated with prior childhood sexual abuse [32]. However, these are linked with the effects of both the quality of the family environment at the time of abuse, and the nature of the subsequent life events [33]. In particular, factors such as family disharmony and violence, existence of other forms of abuse and neglect, and parental mental health difficulties in addition to subsequent events, such as losses through death or separation, combined with the child's own method of coping with the abuse and ameliorative effects of positive school or social relationships all contribute to the outcome. According to reports, about one-third of children are symptom free. Approximately 10% of children show worsening symptoms over time, including depression and post-traumatic symptoms. While in some cases, the effects on personality and social relationships can be disabling during development, other children are unaffected [34].

### **EMOTIONAL ABUSE**

Preferably referred to, as 'psychological maltreatment' is any pattern of behavior that impairs a child's emotional development or sense of self-worth. It encompasses the acts of omission and commission, which are considered potentially psychologically damaging, by conveying to children that they are worthless, flawed, unloved, unwanted, endangered, etc [35]. It may occur in institutional settings as well as in families. Psychological maltreatment can be direct or indirect, for example, being terrorized by witnessing domestic violence directed towards a loved parent, or observing parental involvement in anti-social activities such as drug abuse.

In all six forms of psychological maltreatment have been described, though most cases involve mixture of the following: spurning (hostile rejection or degradation) terrorizing, isolating, exploiting or corrupting, denying emotional responsiveness (ignoring) and mental health, medical and educational neglect [13]. 'Spurning' includes both verbal and non-verbal acts that reject and degrade the child through demeaning, belittling, degrading, ridiculing, shaming or consistently singling out for criticism or humiliation. 'Terrorizing' involves behavior that threatens the child directly or indirectly through threatening the child's attachment figures. 'Isolating' means the consistent denial of the child's opportunities to meet and interact, for example through confining or unreasonably limiting the child's freedoms. 'Exploiting and corrupting' incorporates those acts, which encourage the child to develop inappropriate or antisocial behaviors, for example criminal activities, or sexually illegal

activities such as involvement in pornographic productions or prostitution. 'Denying emotional responsiveness' includes inattention to the child's emotional needs, requirements for nurturing, or emotional well-being. It represents failure to give both love and affection, as well as attentiveness to the child's emotional cues. 'Medical care neglect' refers to the failure or denial to provide the necessary medical treatment (including immunizations, prescribed medications, surgery and other interventions) for serious diseases or injuries. 'Mental health neglect' refers to parent's refusal to comply with the recommended therapeutic procedures in situations where children have serious, treatable psychiatric disorders. Although less well recognized, it can be an issue in anorexia nervosa, pervasive refusal syndrome, or chronic fatigue syndrome. 'Educational neglect' is normally linked with regulations and laws concerning the responsibility of parents to ensure that their children attend school. It is not uncommon, however, for cases involving physical and psychological neglect to also include educational neglect.

Reports of emotional abuse in children become more frequent throughout childhood into adolescence. According to a report, psychological maltreatment constituted 0.5 per 1000 in the United Kingdom, accounting for 17% of the caseload of child protection agencies in 1999 [36]. Reported cases are more likely to be linked with lower socioeconomic status. There is no particular link with racial or ethnic groups. Psychological maltreatment is frequently integral to other forms of maltreatment and so distinguishing different etiological factors and consequences is difficult.

Recognition of emotional abuse usually follows detection of other kinds of child abuse, or when domestic violence or parental substance abuse is uncovered. It also occurs when children are referred to child developmental or mental health clinics, or through reported observations of neighbors or professionals (e.g., teachers). Diagnosis requires detailed history, direct observations of parent-child interactions, and interviews with older child.

### **CAUSES OF CHILD-ABUSE AND NEGLECT**

Abuse and neglect of children occurs in families from all walks of life, and across all socioeconomic, religious, and ethnic groups. There is no single, identifiable cause of child maltreatment; rather, it occurs as a result of an interaction of multiple forces impacting the family. For example, lack of preparation or knowledge of critical issues surrounding parenting, financial or other environmental stressors, difficulty in relationships, stress of single parenting, and depression or other mental health problems can all lead to abusive or neglectful behavior. Parents may lack an understanding of their children's developmental stages and hold unreasonable expectations for their abilities. They also may be unaware of alternatives to corporal punishment or how to discipline their children most effectively at each age. Parents also may lack knowledge of the health, hygiene, and nutritional needs of their children. These circumstances, combined with the inherent challenges of raising children, can result in otherwise well-intentioned parents causing their children harm or neglecting their needs. However, while certain factors related to parents, children, families, and the environment are commonly associated with a greater incidence of child maltreatment, the presence of these factors alone is not sufficient for abusive situations to develop. Stated differently, the presence of known risk factors does not always lead to family violence, and

factors that may cause violence in one family may not result in violence in another family [37].

There is no single cause of child abuse and neglect, rather multiple and interacting factors at the individual (parent and child), familial, community and social levels contribute to child maltreatment. Maternal problems in emotional health, intellectual abilities and substance abuse have been reported to be associated with child neglect. Emotional disturbances, including depression, have been a major finding among mothers of neglected children [38]. Polansky [39] described the apathy-futility syndrome in mothers of neglected children, characterized by an emotional numbness, loneliness, interpersonal relationships that involve desperate clinging, a lack of competence in many areas of living, a reluctance to talk about feelings, the expression of anger through passive aggression and hostile compliance, poor problem-solving skills, a pervasive conviction that nothing is worth doing, and an ability to evoke a sense of futility in others. Mothers of neglected children have been described as more bored, depressed, restless, lonely, and less satisfied with life than the mothers of non-neglected children [40] and more hostile, impulsive, stressed and less socialized than either abusive or non-maltreating mothers [41]. Intellectual impairment, including severe mental retardation and a lack of education, also have been associated with child neglect [42, 43].

Theories of child development and child maltreatment emphasize the importance of considering children's characteristics because caregivers respond differently to these characteristics. For example, parents of children who are temperamentally difficult report more stress in providing care and the situations that lead to parental stress may contribute to child maltreatment. This association is supported by research that has, for example, found increased depression and stress in parents of chronically disabled children [44]. Several studies have found low birth weight or prematurity to be significant risk factors for abuse and neglect [45, 46], because these babies usually receive close pediatric follow-up as well as other interventions. In addition, medical neglect might be expected to occur more often among children who require frequent health care because their increased neglect, but not abuse, among disabled children who had been hospitalized [47].

Problems in parent-child relationships have been found among families of neglected children. Research on dyadic interactions indicated less mutual engagement by both mother and child [48] and frequent disturbances in attachment between mother and infant [49, 50]. Several studies reveal that the mothers of neglected children may have unrealistic expectations of their young children with a lack of knowledge concerning child developmental milestones [51 - 52]. Kadushin [54] described chaotic families with impulsive mothers, who repeatedly demonstrated poor planning and judgment. He further described the negative relationships that many mothers of neglected children have with the fathers of their children, particularly because the fathers had often deserted the family or were incarcerated. Most of the studies on child neglect and high-risk families focuses on mothers and ignores fathers. This bias probably reflects the greater accessibility of mothers, and suggests that the frequently modest involvement of fathers in these families might be an additional contributor to child neglect, and a type of neglect in and of itself. An investigation of families with 5-year old children found that father involvement protected children against neglect [55]. Neglect has been associated with substantial social isolation [56]. Single parenthood without support from a spouse, family members, or friends poses a risk for neglect. In one large controlled study, mothers of neglected children perceived themselves as isolated and as living in unfriendly neighborhood [57]. Crittenden [58] described three types of neglect associated with deficits in cognitive processing, affective processing or both: a) disorganized, b) emotionally neglecting and c) depressed. Families, who respond to the immediate affective demands of situations, with little regard for the cognitive demands, characterize the first type – 'disorganized'. These families operate in crisis mode and appear chaotic and disorganized. Children may be caught in the midst of this crisis, and consequently, their needs are not met. The second type, 'emotionally neglecting', includes families in which there is minimal attention to affect or to the emotional needs of the child. Parents may handle the demands of daily living, but pay little or no attention to how the child feels. The third type, 'depressed', represents the classic image of neglect. Parents are depressed and therefore unable to process either cognitive or affective information. Children may be left to fend for themselves emotionally and physically. Although the intervention for these three different family patterns differs, they all rely on a family system approach in which parent and child behavior are considered in context.

The community context and its support systems and resources influence parent-child relationships and are strongly associated with child maltreatment [59]. A community with a rich array of services, such as parenting groups, high quality and affordable child care, and a good transportation system, enhances the ability of families to nurture and protect their children Informal support networks, safety, and recreational facilities also are important in supporting healthy family functioning. Garbarino and Crouter [59] described the feedback process whereby neighbors may monitor each other's behavior, recognize difficulties, and intervene. This feedback can be supportive and diminish social isolation, and may help families obtain necessary services. Studies have reported that the families in a high-risk environment are less able to give and share and might be mistrustful of neighborly exchange. In this way, a family's problems seem to be compounded rather than ameliorated by the neighborhood context [60]. Socioeconomic factors (poverty) and parent's negative perceptions of the quality of life in the neighborhood have been reported to be strongly associated with increased child maltreatment [59]. In summary, communities can serve as a valuable source of support to families or they may add to the stresses that families are experiencing.

# IMPACT OF CHILD ABUSE AND NEGLECT

It has been reported that the impact of child maltreatment is far greater than its immediate, visible effects. Abuse and neglect are associated with short- and long-term consequences that may include brain damage, developmental delays, learning disorders, problems forming relationships, aggressive behavior, and depression [61 - 72]. Survivors of child abuse and neglect may be at greater risk for problems later in life—such as low academic achievement, drug use, teen pregnancy, and criminal behavior—that affect not just the child and family, but society as a whole. Social and emotional handicaps are perhaps the most serious long-term consequences of maltreatment. Physically abused children are generally more aggressive with their peers than children who have not been abused, have more troubled interpersonal relationships and have more depressive symptoms and affective disorders. Sexually abused children, in addition to their depressive and aggressive symptoms

have an increased frequency of anxiety disorders and problems with sex roles and sexual functioning [73]. Furthermore, abused children when they become adults have 2 or 3 times as many problems with substance abuse and depression. Still further, abused children, when they become parents have a tendency to abuse their own children. However, the physical, behavioral and emotional manifestations will vary between children, depending on the developmental stage at which the abuse occurs and its severity.

Walker et al. [74], reported a broad array of adverse adult health outcomes with which child sexual and physical abuse are associated, including disturbances in emotional, social, and role functioning; mental health problems; risky health practices including substance abuse; a broad array of physical symptoms and health disorders reaching clinical threshold. Felitti et al. [75], found a strong graded relationship between exposure to abusive or neglectful conditions in childhood and risk of certain adult disorders, including ischemic heart disease, cancer, chronic lung disease, skeletal fractures, and liver disease.

Findings of these studies call for an explanation as to how do abusive and neglectful family environments in childhood affect mental and physical health decades later? An important route involves alterations in autonomic and endocrine responses. The consequences of early exposure to abuse or neglect appear to be potentially irreversible disruptions in biological systems that may, perhaps also in conjunction with genetic predisposition, produce significant differences in susceptibility to stress, to biological markers of the cumulative effects of stress, and to stress related physical and mental health disorders [76]. For example, alterations in sympathetic nervous system reactivity in response to stress appear to begin in childhood, in part as a result of exposure to abusive or neglectful circumstances. These include elevated blood pressure and heart rate responses, as well as prolonged sympathetic response to stress. Alterations in hypothalamic-pituitary-adrenocortical reactivity in response to stress also appear to be related to adverse family history. In families characterized by conflict, anger, and aggression, for example, there may be disruptions of the normal rise and fall of serum cortisol levels in response to acutely stressful circumstances, such that a protracted cortisol response, or in extreme cases a flattened cortisol response, may develop. In contrast, a warm and supportive family environment may actually foster a less elevated cortisol response to potentially stressful circumstances or one that rapidly habituates to stress [77]. There is evidence too, for serotonergic dysregulation in offspring of abusive and neglectful families. Difficulty in moderating aggressive impulses, problems in developing and maintaining social relationships, and risky health-related behaviors, especially substance abuse, are among the outcomes consistently seen in response to the abusive or neglectful childhood family environment, and serotonergic dysfunction may be implicated in these patterns.

The adverse long-term health effects of childhood abuse cannot be understood solely in terms of biological dysregulations and behavioral pathways that follow from, exacerbate, and interact with biological dysregulations are at least as important. Two important psychosocial mechanisms whereby the long-term adverse effects of abuse may be maintained into adult hood include: deficits in emotion regulation skills and low social competence. In the warm and supportive household, children learn emotion regulation strategies, including how to recognize their emotions when they occur, and how to control those emotions when interacting with others. Children from families in which there has been abuse or neglect have difficulty monitoring, identifying, reacting to, and regulating their emotions, including emotional intensity, range, and duration. Among the consequences of poor emotional regulation are social rejection and isolation. Children who are unable to express or control their emotions or respond appropriately to the emotions of others are more likely to be bullied by others, act as bullies themselves, or be socially ignored or isolated – problems that lay the groundwork for difficulties in adulthood [78].

Deficits in appropriate emotion regulation and social regulation skills take their toll on adult health in many ways. People who lack these vital skills interact poorly in the social environment. As adults, they may have difficulty forming long-term social relationships, such as romantic relationships and friendships, and compromise their work-related relationships as well. Social isolation and lack of social support have been related to a greater risk of mortality from all causes, as well as more protracted courses of illness and recovery from several acute and chronic disorders. Emotion dysregulation and low social competence, in turn, increase the likelihood of substance abuse, aggression, risky behavior, and other poor health habits. Poor control over emotions and the social fallout that results may thus be essential components of the sustaining pathways that explain the long-term effects of childhood abuse on social, mental and physical health. Emotional dysregulation as a risk factor for mental illness is implicated in almost half of the DSM-IV axis I psychiatric disorders and in almost all of the axis II disorders. Emotion regulation is implicated in physical health, in that people who are emotionally reactive are also likely to be physiologically reactive. For example, emotions such as anger and hostility have been related to the development and exacerbation of hypertension and coronary heart disease. In summary, child neglect and abuse can justifiably be viewed as a public health problem with immediate and long-term social, mental and physical health consequences.

An association between abuse and suicide has been found in studies of abused children and adolescents, of adolescent suicide attempters, and of mothers who attempt suicide. Selfmutilation has been reported in abused and neglected children [79]. Adolescents who attempt suicide have more often been reported as abuse victims [80]. High rates of suicide attempts have been reported in adolescent runaways [81]. Kaplan et al., [82] reported that conduct disorders, substance abuse, depression, family discord, and family histories of suicide, and other exposure to suicide appear to be risk factors for adolescent suicide. Child-abusive behaviors were also found more often in mothers who attempted suicide than in a comparison group of non-suicidal mothers [83]. Studies have reported that physically abused adolescents have significantly more risk factors for suicide than do non-abused adolescents [84]. Though the origin of the association between suicidal behavior and physical abuse remains to be studied, it may be secondary to modeling of aggressive behavior within the family or to exposure to the suicidal behavior of family members. [85, 86]

#### PREVENTION

Prevention of child abuse and neglect has taken on many forms since the 1960s when the Battered Child Syndrome was identified. Policy makers, legislators, professionals, and concerned citizens have struggled to find effective ways to prevent violence against children. The goal of child abuse prevention is simple—to stop child abuse and neglect from happening in the first place, sparing children and family's emotional and physical trauma and decreasing

the need for costly intervention and treatment services. Many principles of prevention have been suggested from time to time.

Mrazek et al., [87] have categorized prevention programs along three dimensions: universal, selected and indicated. 'Universal' or population-based interventions are designed to prevent child maltreatment for the entire population of children and are mass distributed. Examples include child protection policies, mass media campaigns, and public service announcements to draw public attention to the importance of protecting children and promoting their optimal development. 'Selected' interventions are directed towards families who are at high risk for child neglect and abuse. Their goal is to reduce the incidence of child maltreatment, often by reducing the risk factors that are associated with child neglect and abuse. Thus selected interventions are often targeted to families experiencing social isolation, high stress, few resources, poverty and substance abuse. 'Indicated' interventions are directed towards families in which child neglect / abuse has already occurred and aim at minimizing the negative effects of neglect or abuse on the child and breaking the cycle by preventing further occurrence. In a review of prevention strategies for neglect, Holden et al., [88] recommended the need for theory-driven, longitudinal programs that extend beyond the prevention of neglect to the promotion of healthy care-giving practices.

The best way to prevent child abuse and neglect is to support families and provide parents with the skills and resources they need. It has been established that those in closest proximity to the child perpetrate a considerable amount of child maltreatment. While it is accepted that maltreatment can occur because of wider societal system, organizations and processes, it's also agreed that a major step in developing preventive strategies is to recognize those dangers, which are in the control of individuals closest to the child. These individuals are most frequently family and community members and in some cases strangers whom the child encounters within his / her day-to-day environment. Interventions can be at the primary, secondary or tertiary level. Primary prevention aims at proving education to the parents to be, about good parenting. At the secondary level those parents who were abused, as children should be given education regarding child rearing. Once the abuse has occurred tertiary prevention helps to avoid recurrence of the same.

# MULTIDISCIPLINARY TEAM APPROACH

Effective assessment of the alleged child victim and intervention in the field of child maltreatment is the result of collective insights of many professionals. The mechanism for achieving consensus and thus shared responsibility is the multidisciplinary team. Each discipline holds one piece of the puzzle, and until all of the interlocking pieces are joined, a full understanding of the child's experience and the needs of the victim and family will not be appreciated. The multidisciplinary team is the mechanism to ensure improved coordination, better information gathering, and thus a more comprehensive assessment of an allegation. A comprehensive assessment assists in providing direction for the law enforcement, mental health, and the child-protection system to respond, when coordinated intervention occurs in a coordinated manner, appropriate service needs can be identified early in delivered timely.

The legal system can respond without compromising its focus and taking into account the needs of the victims and their families. Although valid statistics are lacking, the

multidisciplinary team is likely to lead to more successful prosecution when appropriate. The artificial turf issues in each discipline and the boundaries that exist to form the turf diminish with time, and all disciplines begin to interact to reflect a more global understanding of the system and the victim's needs. The issue of confidentiality of the communication as well as discovery rules may vary from state to state and must be addressed, with particular attention to how the team's discussions and recommendations are recorded.

A multi-disciplinary protection team should include representation from the medical, mental health, education, social service and legal professions working with government agencies. Non Governmental Organizations can also play a lead role. An emotionally secure infancy is a fundamental requirement for a stable adulthood. Hence treatment services must focus on the child as well as the parent. They must help children develop feelings of self worth and learn interpersonal skills and coping mechanism. For this to happen the basic needs of the family and the community should be met. World Health Organization is active in the areas of improving mother and child interaction, in promoting life skills education for children and adolescents and in promoting child-friendly schools. Parents need to rediscover the art and skill of our yester year grandmothers, in preventing child sexual abuse in the families, in a subtle way by anticipating situations and acting appropriately in time. As professionals, the health care workers need to be committed to protect the rights of the children by motivating the parents to be more proactive and systematically sensitizing the other agencies like the police and judiciary about the various aspects of the problem.

*Providing ongoing support* to the victims of childhood abuse may include assisting young people to:

- Regain trust and establish supportive relationships;
- Understand the effects on their current thoughts and feelings;
- Validate the normality of their thoughts and feelings;
- Manage impacts on other relationships;
- Rediscover their strengths and value what is special about them;
- Help them to make meaning from their experiences; and
- Help them to make the changes they want to make, at their pace.

# WHY DOES PREVENTION MATTER?

Through prevention activities such as parent education, home visitation, and parent support groups, many families can find the support they need to stay together and care for their children in their homes and communities. Prevention efforts help parents develop their parenting skills, understand the benefits of nonviolent discipline techniques, and understand and meet their child's emotional, physical, and developmental needs. Prevention programs also can help parents identify other needs they may have and offer assistance in getting that additional support.

Prevention of socially undesirable and hazardous behaviors may not only save lives, but also resources. While it is impossible to entirely eradicate certain kinds of behavior that can have tragic human consequences, including the maltreatment of children, human service professionals have been buoyed by improvements over time across numerous major indices that measure the health and well-being of individuals and families.

# **INTERVENTION AND TREATMENT**

The primary concern of child welfare agencies has been child protection and safety. Increasingly, in Australia, the United States and the United Kingdom there have been movements to redress the imbalance, so that child- and family-support services achieve much greater importance. This may be welcomed, because maltreated children and their families are typically embedded within a range of family difficulties and social stresses, which need to be tackled to obtain positive outcomes for children. Intervention therefore is multifaceted and may involve several professionals working together [89]. The overall objectives of management are to help the child and his or her family by:

- Stopping abuse or reversing neglect and ensuring adequate care-taking;
- Improving the capacity for positive interpersonal relationships between family members;
- Treating symptoms of psychological disorders;
- Stopping sexually aggressive, violent, or exploitative behavior that is either directed towards child or which is likely to have an impact on him or her;
- Facilitating understanding by the child and his or her parents of the nature and dynamics of maltreatment.

The process of treatment may be conceptualized in three phases: first, an initial acknowledgement of the problem; second, improving parental competence and sensitivity to their child's needs and third, a resolution phase, either in the direction of alternative care or, if possible within the family [89].

Treatment of child sexual abuse requires a psycho-educational component, so that the nature of sexual exploitation and the way in which children become coerced into sexual activity can be understood by the child without a sense of personal blame. Cognitivebehavioral treatments have been particularly helpful especially when combined with a program of intervention for the non-abusive caretaker [33]. Abusers should be assessed for their potential to respond to psychological treatment [90]. In the field of physical abuse, cognitive-behavioral treatment and family therapy have been shown to be more effective than routine community services at reducing the levels of violence and parental distress, as well as family conflict [13]. Psychological treatments in the field of neglect have concentrated on improving parenting skills and sensitivity through direct encouragement of positive interactions in feeding, play and general care, combined with individual therapy for parents themselves, who have frequently experienced multiple deprived childhoods. Psychiatric treatment of parental mental health problems, such as depression or substance abuse, is critical. Mobilizing community-based supports and networks to overcome social isolation, and linking neglectful mothers with other parents who can provide role modeling and support, appears to be promising approach.

# FORENSIC ASPECTS OF INTERVIEWING ALLEGEDLY ABUSED CHILDREN

Since there is currently no valid syndrome for child sexual abuse and there is a strong likelihood of child sexual abuse symptoms having multiple explanations, short of a valid offender confession [91] and incontrovertible medical evidence, the crucial aspect of a proper forensic evaluation of suspected sexual abuse in children is usually the interview itself, which must be done in an 'ethically permissible manner' [92]. Mandatory state and provincial reporting laws require the health care professionals as well as others in positions of trust to report the suspected child sexual abuse to child protection services or the police. However, since the initial disclosure interview with the child is usually the most valid one, the clinical practitioner should be prepared to provide minimally optimum standards of forensic interviewing at this crucial juncture. Furthermore, there has been increasing recognition that, under certain circumstances, false memories of abuse may develop in children [93]. Preschoolers, in particular, have been found vulnerable to confusion and false memories because of faulty interview techniques. Recent studies inform clinical practice by suggesting that children are most likely to be misled during interviews if (a) they feel it is preferable to 'guess' or 'pretend' during interviews rather than to acknowledge 'not knowing' [94]; (b) their memories are 'weak' [95]; (c) they are confused about 'adult' language that does not take the child's development level into account; (d) the interviewer is seen as unfriendly, intimidating, or authoritarian [96]; or (e) the interviewer has a preconceived bias about the event.

Statement validity analysis is a serious attempt to construct a valid and systematic forensic interview schedule for children that emphasizes free recall as its key component. There also is a general consensus that the use of open-ended questions when interviewing children suspected to have been sexually abused, although associated with lowest risk of confusion, also is associated with the generation of least information. In contrast, leading questions, which pressure a child to agree with the interviewer, can lead to a permanent distortion of memory for what happened. Focused questions, which fall between free recall and leading questions, have been found to be necessary by many clinicians when faced with a child reluctant to discuss abuse [97]. To minimize suggestibility in forensic interviews of allegedly abused children, Reed [98] has recommended that interviewers systematically clarify to children what is expected of them by specifically teaching children to state when they are 'confused' or 'do not know' answers to the questions.

Several structured and semi-structured interview protocols have been introduced that were designed to maximize the amount of accurate information and to minimize mistaken or false information provided by children. These approaches include the 'cognitive interview', which encourages witnesses to search their memories in various ways, such as recalling events forward and then backward. The step-wise interview is a funnel approach that starts with open-ended questions and, if necessary, moves to more specific questions. The interview protocol developed at the National Institute of Child Health and Human Development includes a series of phases and makes use of detailed interview scripts [99].

Although these protocols may be particularly important in a forensic context, experienced clinicians endorse flexibility to size up the situation and use techniques that are likely to help the youngster become comfortable and communicative. One victim might need a favorite

object (e.g., a teddy bear or a toy truck); another might need to have a particular person included in the interview. Some children are comfortable talking; others prefer to draw pictures. An unrelated joke, a shared cookie, or a picture on the evaluator's wall may lead to a disclosure of abuse. Important comments may be made while chatting during the break time, instead of during the structured interviews. The interviewer should remember the following principles:

- Audiotape or videotape the interview, if possible;
- Use a minimum number of interviews, as multiple interviews may encourage confabulation;
- Avoid repetitive questions, either-or questions, leading and suggestive questions;
- Use restatement, that is, repeat the child's account back to the child (This allows the interviewer to see if the child is consistent and ensures that the interviewer understands the child's report);
- Conduct the interview without the parent being present, however, if the child is young, consider having a family member in the room;
- Use an examination technique that is appropriate to the child's age and developmental level;
- Determine the child's terms for body parts and sexual acts; do not educate or provide new terms.

However, the interview should not take the form an interrogation. The interviewer should note the child's affects while discussing the topic and be tactful in helping the child manage anxiety. The 'step-wise interview, which is primarily intended for forensic evaluation, consists of the following components: building rapport, asking the child to describe two specific past events to assess the child's memory, establishing the need to tell the truth, introducing the topic of concern, eliciting a free narrative, posing general questions, posing specific questions, using interview aids etc., while trying to explore:

- Whether the child was told to report or not to report anything;
- Who the alleged perpetrator was;
- What the alleged perpetrator did;
- Where did it happen;
- When it started and when it ended;
- Number of times the abuse occurred;
- How the child was initially engaged and how the abuse progressed over time;
- How the alleged perpetrator induced the child to maintain secrecy;
- Whether the child is aware of specific injuries or physical symptoms associated with the abuse;
- Whether any photography or videotaping took place.

# FORENSIC EVIDENCE COLLECTION IN CHILD SEXUAL ABUSE

It is first necessary to consider what the legal system views as scientifically useful evidence and how this applies to behavioral testimony, whether given by psychologists, psychiatrists, or social workers. Such psychological considerations are quite different from clinical ones in a variety of ways, which turn on the rigor of findings (i.e., probability), their admissibility in court (i.e., relevancy) and our avoidance of biasing the jury. Although there is a good deal of variation in how these 'rules of evidence' are interpreted and administered in courts at various states and provincial levels, these legal tools are nevertheless broad guidelines to good forensic practice in general. When science and the law work well together, there is a relatively smooth transition between the two as to the validity of data. Thus, if one's assessment is knowingly to be used in court, then it should be prepared as a scientific exhibit rendered interpretable according to common law rules of evidence.

Several general guidelines about evidence should be addressed before outlining a more specific approach to specimen collection. First, specific details of collection, labeling and packaging of specimens should be worked out with the laboratory processing the specimens. Second, a specimen collection protocol should be used to ensure that all appropriate specimens are collected for both routine and non-routine circumstances. Third, collection kits should be standardized, providing containers, collection devices, and checklists to assure proper collection of specimens. Fourth, the procedures for collecting specimen are best explained in advance to the child and the caretakers, because cooperation is the key to the collection of proper specimens. Fifth, proper consent must be obtained from the parent and child before performing the examination and collecting evidence. Finally, the handling of collected specimens should be documented to maintain the 'chain of evidence'.

Forensic evidence collection, usually done at the same time as the physical examination, has been recommended when the examination occurs within 72 hours of acute sexual abuse [100]. This recommendation is based at least in part on the fact that seminal fluid and other foreign substances are rarely recoverable after 72 hours of the sexual contact. The types of evidence sought include sperm, seminal fluid, and foreign materials on the victim's body surface or clothing. Semen, blood, saliva, body hair, bite marks, and other materials occasionally found on the body of the victim can also be used to identify the perpetrator with the help of genetic markers. DNA typing or profiling is the most specific development in the identification of the perpetrator of sexual abuse [101]. Christian et al., [102] recommended that guidelines for evidence collection for adult sexual assault victims may not be appropriate for pre-pubertal victims and the clothing and linens that yield the majority of evidence should be vigorously pursued for analysis.

# MEDICOLEGAL ASPECTS OF CHILD ABUSE AND NEGLECT

All states have reporting laws mandating that professionals in the areas of health care, social service, law enforcement, and education report suspected cases of child abuse or neglect. Conversely, the ethical principles of medicine, nursing and other professions require

professionals to safeguard confidential information. But in legal proceedings, the presence or absence of a privilege is important. In court, a professional may have to answer questions that require disclosure of information the professional is ethically bound to protect.

Confidentiality arises from three sources: (a) the broad ethical duty to protect confidential information, (b) laws that make certain records confidential, and (c) privileges that apply in legal proceedings. Communication between a patient and a professional is privileged when three requirements are fulfilled. First, the communication must be between a patient and a professional with whom privileged communication is possible. The fact that a third person is present when a patient discloses information may or may not eliminate the confidentiality required for a privilege. Second, the patient must consult the professional to obtain advice or therapy. Despite the fact that a patient may consult a physician who refers the patient to a second professional, communication between the patient and the referring physician is privileged even though the patient does not receive treatment from the referring physician. Third, only communications that the patient intends to be confidential are privileged. Furthermore, privileged communications remain privileged when the relationship with the patient ends as in case of death of the patient. It must be stressed that the privilege belongs to the patient, not the professional. In legal parlance, the patient is the 'holder' of privilege and as a privilege holder; he can prevent the professional from disclosing privileged information in legal proceedings. When the patient is a child, parents normally have authority to make decisions about confidential and privileged information. However, when a parent is accused of abusing or neglecting a child, it may be inappropriate for the parent to make decisions regarding the child's confidential information. In the event of a conflict between the interests of the child and the parents, the court may appoint someone else, such as a 'guardian ad litem', to make decisions about confidential and privileged information.

Child abuse reporting laws require professionals to report suspected child abuse and neglect to designate authorities [103]. The reporting laws override the ethical duty to protect confidential client information. Moreover, the reporting requirement overrides privileges for confidential communications between professionals and their patients. However, although reporting laws abrogate privileges, abrogation usually is not complete. In many states, professionals may limit the information by reporting the specific information required by law. Information that is not required to be reported remains privileged.

Statement of the abused child recorded during examination and interview has medical as well as legal significance [1]. If children's statements are properly documented they may be admissible in subsequent legal proceedings. In some cases, the child's statements to the professionals are the most compelling evidence of maltreatment. However, the rule in all states is that 'hearsay' is inadmissible in criminal and civil litigation unless the particular hearsay statement meets the requirements of an exception to the rule against hearsay. In these cases, a child's words are hearsay if three requirements are fulfilled: (a) the child's words were intended by the child to describe something that happened; (b) the child's words were spoken before the court proceedings at which the words are repeated by someone who heard the child speak and (c) the child's words are offered in court to prove that what the child said actually happened [104 - 106].

An 'excited utterance' is a hearsay statement that relates to a startling event. The statement must be made while the child is under the acute emotional stress caused by the startling event. Excited utterances can be used in court even though they are hearsay. Accordingly a professional can document the following important factors:

- Nature of the event;
- Amount of time elapsed between the startling event and the child's statement relating to the event (the more time that passes between a startling event and a child's statement describing it, the less likely a court is to conclude that the statement is an excited utterance, however, exceptions are there);
- Indications that the child was emotionally upset when the child spoke (whether crying, frightened or otherwise upset);
- Extent to which the child's statement was spontaneous (the more spontaneous the statement the more likely it meets the requirements of this exception);
- Number and type of questions used to elicit the child's statement (asking questions does not necessarily destroy the spontaneity required for the excited utterance exception, however, as the questions become leading, spontaneity may dissipate, undermining the applicability of this exception);

First safe opportunity to disclose what happened (in many cases, abused child remains under the control of the abuser for hours after the abusive incident and has the first safe opportunity to disclose when released. A child's statement at the first safe opportunity may qualify as an excited utterance even though considerable time has elapsed since the abuse occurred)

A child's initial disclosure of sexual abuse may be admissible in court under an ancient legal doctrine called 'fresh complaint of rape or sexual assault'. In most states, a child's fresh complaint is not, technically speaking, hearsay.

Most states have an exception to the hearsay rule for certain statements to professionals providing diagnostic or treatment services, commonly called the 'diagnostic or treatment exception'. The primary rationale for this exception is that hearsay statements to these professionals are reliable. Reliability exists because the patient has a strong incentive to be truthful with the professional. To increase the probability that a child's statements satisfy the diagnosis or treatment exception to the rule against hearsay, the professional can take the following steps:

- Discuss with the child the clinical importance of providing accurate information and of being completely forthcoming.
- The diagnosis or treatment exception requires that information supplied to the professional be pertinent to diagnosis or treatment. Thus it is important to document how information disclosed by the child is pertinent to diagnosis or treatment.
- If the child identifies the perpetrator, the professional should document why knowing the identity of the perpetrator is pertinent to diagnosis or treatment. For example, knowing the identity of the perpetrator may be important to determine whether it is safe to send the child home or the physician needs the perpetrator's identity if sexually transmitted disease is a possibility.
- Document the foregoing factors in the child's chart.

Many states have a hearsay exception known as 'residual or catch all exception', which allows use in court of a reliable hearsay statement that does not meet the requirements of one of the traditional exceptions. Some states also have a special hearsay exception for statements by children in child abuse cases. These 'child hearsay exceptions, and residual exceptions allow use in court of children's reliable hearsay statements that do not fit into another exception.

Proper documentation by the medical professionals is needed not only to preserve the child's words, but also to preserve a record of the factors indicating that the child's hearsay statements meet the requirements of an exception to the hearsay rule. Without careful documentation of exactly what questions are asked and exactly what the child says, the professional will not likely remember months or years later, when called as a witness and asked to repeat what the child said. Furthermore, uncorroborated clinical inference and contaminated interviewing have no place in preparing scientifically informed child abuse evidentiary material. When clinical descriptions of children are improperly elevated to the status of empirically contentious syndromes, disorders or even simple profiles, numerous psycho-legal errors and miscarriages of justice can result because they do not sufficiently reduce uncertainty below reasonable doubt and can then bias the trier of fact and invade the province of jury. It is also essential to know the scientific limitations of one's discipline to provide a balanced and useful forensic evaluation.

# CONCLUSION

Many efforts are directed towards identification of victims, and yet prevention is the ultimate goal in the management of child abuse and neglect. Multidisciplinary teams provide an excellent opportunity for clinicians to become involved in the child protection system and to take leadership role in their community. Medical professionals functioning as a part of team can educate team members concerning the scope of medicines' contribution in the field. At the same time, they can learn to appreciate the complexity of the system and then advocate for necessary changes that will positively affect children and their families. However, different professionals, particularly those of medicine and law sometimes seem like ships passing in the night, yet if children are to be protected, all must work together. Only genuine interdisciplinary cooperation holds realistic hope of reducing the tragic number of abused and neglected children. Furthermore, given the presumption of innocence in common law and the need to protect the rights of both the accused and the victims, it is especially incumbent upon health professionals and allied service providers to adhere to very high scientific standards when assessing child sexual abuse. Regrettably, however, most health-care professionals and lawyers are not specifically versed in terms of valid child sexual abuse interviewing. Presumably then, their knowledge of valid behavioral manifestations of child sexual abuse is also likely to be deficient. Accordingly, there is a pressing need for more user-friendly literature that integrates behavioral and legal issues.

#### REFERENCES

Sharma BR. Medicolegal aspects of child abuse. *Physician's Digest 2005;* 14 (1): 41 – 48

[2]	U.S. Department of Health and Human Services, Administration on Children, Youth and Families: Child Maltreatment 1998: reports from the States to the National Child Abuse and Neglect Data System. Washington, DC, US Government Printing Office, 2002
[3]	U.S. Department of Health and Human Services, Administration on Children, Youth and Families: <i>Child Maltreatment 1998: reports from the States to the National Child Abuse and Neglect Data System.</i> Washington, DC, US Government Printing Office, 2000
[4]	Sedlak A, Broadhurst MLA. <i>Third National Incidence Study of Child Abuse and Neglect Washington DC</i> , U.S. Department of Health and Human Services, 1996
[5] [6]	<i>Kids Help Line report 2003 – 2004</i> , available at <i>www.kidshelp.com.au</i> Black MM. The roots of child neglect. In Rees RM. Ed. <i>The treatment of child abuse</i> . Baltimore; Johns Hopkins University Press; 2000
[7]	Zuravin SJ. Child neglect: a review of definitions and measurement research. In Dubowitz H. ed. Neglected children; Thousand Oaks CA: <i>Sage Publication 1999</i> ; 24 – 46.
[8]	Helfer RE. The neglect of our children. Pediartic Clin North Am. 1990; 37: 923.
[9]	Berman BD, Winkleby M, Chesterman E. et al. After-school child care and self-esteem in school-age children. <i>Pediatrics 1992</i> ; 89: 654
[10]	Jaudes PK, Diamond LJ. Neglect of chronically ill children. <i>Am J Dis Child 1986</i> ; 140: 655
[11]	Nandy A. <i>Principles of Forensic Medicine</i> , 1 <sup>st</sup> Ed. 1995; New Central Book Agency (P) Ltd.: 352 - 357
[12]	Caffey J. Multiple fractures in the long bones of infants suffering from chronic subdural hematoma. <i>Am J Rontgenol 1946</i> ; 56: 163 – 173.
[13]	Skuse D, Bentovim A. Physical and emotional maltreatment. In M Rutter, E Taylor, L Hearsov eds. Child and Adolescent Psychiatry, <i>Blackwell Science, Oxford, 1994</i> , 209 – 229.
[14]	Westcott HL, Jones DPH. Annotation: the abuse of disabled children. J Child Psychology and Psychiatry 1999; 40: 497 - 506
[15]	Kolko D. Child physical abuse. In J Briere, L Berliner J Bulkley et al., eds. The APSAC hand book on child maltreatment, <i>Sage</i> , <i>London</i> , <i>1996</i> , 21 - 50
[16]	Feldman K. Evaluation of physical abuse. In R. Helfer, R Kempe and R Krugman eds. The Battered Child, <i>University of Chicago Press</i> , (5 <sup>th</sup> edn) 1997, 175 - 220
[17]	Cicchetti D, Toth S. A developmental psychopathology perspective on child abuse and neglect. <i>J Am Acad Child and Adolescent Psychiatry 1995</i> ; 34: 541 – 565
[18]	Kempe CH. Sexual abuse, another hidden pediatric problem. <i>Pediatrics 1978</i> ; 62: 382 – 389
[19]	Finkelhor DH, Hotaling GT. Sexual abuse in the National Incidence Study of Child Abuse and Neglect: an appraisal. <i>Child Abuse Negl</i> 1984; 8: 23 – 32
[20]	Wyatt GE, Peters SD. Methodological consideration in research on the prevalence of child sexual abuse. <i>Child Abuse Negl 1986</i> ; 10: 241 – 251
[21]	Sharma BR, Gupta M. Child abuse and its implications. <i>J Clinical Forensic Medicine</i> 2004; 11 (5): 258 – 266
[22]	Beriner L, Conte JR. The process of victimization: the victim's perspective. <i>Child Abuse Negl 1990</i> ; 14: 29 – 40

- [23] American Prosecutors Research Institute National Center for Prosecution of Child Abuse. *Investigation and prosecution of child abuse, Alexandria, VA*: American Prosecutor's Institute, 1987
- [24] Finkelhor D. The victimization of children: a developmental perspective. Am J Orthopsychiatry 1995; 65: 177 – 193
- [25] Summit RC. The child sexual abuse accommodation syndrome. *Child Abuse Negl* 1983; 7: 177 193
- [26] Hunter RS, kilstrom N, Loda F. Sexually abused children: identifying masked presentations in a medical setting. *Child Abuse Negl 1985*; 9: 17 25
- [27] Seidel JS, Elvik SL Berkowitz CD, et al. Presentation and evaluation of sexual misuse in the emergency department. *Pediatr Emerg Care 1986*; 2: 157
- [28] DeJong AR, Emans SJ, Goldfarb A. Sexual abuse: what you must know? Patient Care 1989; 23: 145
- [29] Klevan JL, deJong AR. Urinary tract symptoms and urinary tract infection following sexual abuse. *Am J Dis Child 1990*; 144: 242 244
- [30] Meadow R. ABC of child abuse, BMJ Publishing London, 3<sup>rd</sup> Ed. 1997
- [31] Jones DPH, Lynch MA. Diagnosing and responding to serious child abuse. *BMJ 1998*; 317: 484 - 485
- [32] Jones DPH. Management of the sexually abused child. Advances in Psychiatric Treatment 1996; 2: 39 – 45
- [33] Jones DPH, Ramchandani P. *Child sexual abuse informing practice from research,* Radcliffe Medical Press, Oxford, 1998.
- [34] Rind B, Tromovitch P, Bauserman R. A meta-analytic examination of assumed properties of child sexual abuse using college samples. *Psychological Bulletin 1998*; 124: 22 – 53
- [35] Brassard M, Hardy D. Psychological maltreatment. In R. Helfer, R Kempe and R Krugman eds. *The Battered Child*, University of Chicago Press, (5<sup>th</sup> edn) 1997, 392 – 412
- [36] Department of Health, Health and personal social services statistics for England, HMSO London, 2000. Website *http://www.Doh.gov.uk/HPSSS/INDEX.HTM*
- [37] Sharma BR, Harish D, Gupta M. Innocence abused: an overview of the nature, causes and prevention of child abuse. *J Indian Academy of Forensic Medicine 2003*; 25 (3): 87 92
- [38] Zuravin S. Child abuse, child neglect and maternal depression: is there a connection? National Center on Child Abuse and Neglect: child neglect monograph: proceedings from a symposium. Washington DC: Clearinghouse on child abuse and neglect information, 1988.
- [39] Polansky N. Damaged parents: an anatomy of child neglect. Chicago: University of Chicago. 1981
- [40] Wolock I, Horowitz H. Child maltreatment and maternal deprivation among AFDC recipient families. Soc Serv Res 1979; 53: 175.
- [41] Friedrich WN, Tyler JA, Clark JA. Personality and psycho-physiological variables: its abusive, neglectful and low- income control mothers. J Nerv Ment Disord 1985; 173: 449
- [42] Martin M, Walters S. Familial correlates of selected types of child abuse and neglect. J Marriage Family 1982; 44: 267.

498	B. R. Sharma
[43]	Ory N, Earp J. Child maltreatment and the use of social services. <i>Public Health Rep</i> 1981; 96: 238.
[44]	Shapiro J. Family reactions and coping strategies in response to the physically ill or handicapped child. <i>Soc Sci Med 1983</i> ; 17: 913
[45]	Benedict M, White K. Selected perinatal factors and child abuse. <i>Am J Public Health</i> 1985; 75: 780.
[46]	Herrenkohl EC, Herrenkohl RC. Some antecedents and developmental consequences of <i>child maltreatment</i> . In Rizley R, Cicchetti D, eds Developmental perspectives on child maltreatment. San Francisco: Jossey-Bass, 1981.
[47]	Glaser D, Bentovim A. Abuse and risk to handicapped and chronically ill children. <i>Child Abuse negl 1979;</i> 3: 565
[48]	Dietrich KN, Starr RH, Weisfeld OE. Infant maltreatment: caretaker-infant interaction and developmental consequences at different levels of parenting failure. <i>Pediatrics 1983</i> ; 72: 332.
[49]	Crittenden PM. Maltreated infants: vulnerability and resilience J Child Psychol Psychiatry 1985; 26: 85.
[50]	Egeland B, Brunquell D. An at-risk approach to the studies of child abuse and neglect. <i>J Am Acad Child Adolesc Psychiatry 1979</i> ; 18: 219.
[51]	Azar S, Robinson DR, Hekimien E. Unrealistic expectations and problem solving ability in maltreating and comparison mothers. <i>J Consult Clin Psychol 1984</i> ; 52: 687
[52]	Twentyman C, Plotkin R. Unrealistic expectations of parents who maltreat their children: an educational deficit that pertains to child development. <i>J Clin Psychol 1982</i> ; 38: 497.
[53]	Herrenkohl RC, Herrenkohl EC, Egolf BP. Circumstances surrounding the occurrence of child maltreatment. <i>J Consult Clin Psychol 1983</i> ; 51: 424.
[54]	Kadushin A. Neglect in families. In Nunnally EW, Chilman CS, Cox FM. Eds Mental illness, delinquency: addictions and neglect. Newbury Park: sage, 1988.
[55]	Dubowitz H, Black MM, Kerr M, et al. Fathers and child neglect. Arch Pediatr Adolesc Med 2000; 154: 135
[56]	Polansky NA, Ammons PW, Gaudin JM. Loneliness and isolation in child neglect. Social Casework 1985; 66: 38
[57]	Polansky NA, Gaudin JM, Ammons PW, et al. The psychological ecology of the neglectful mother. <i>Child Abuse Neglect 1985</i> ; 9: 265
[58]	Crittenden PM. Child neglect: causes and contributors. In Dubowitz H eds Neglected Children. Thousand Oaks, CA: <i>Sage publications 1999</i> ; 47 – 68.
[59]	Garbarino J, Crouter A. Defining the community context of parent-child relations. <i>Child Dev 1978</i> ; 49: 604
[60]	Garbarino J, Sherman D. High-risk neighborhoods and high-risk families: the human ecology of child maltreatment. <i>Child Dev 1980</i> ; 51: 188
[61]	Rimsza ME, Berg RA. Sexual abuse: somatic and emotional reactions. <i>Child Abuse Negl. 1988</i> ; 12: 201 - 208.
[62]	Walker E, Katon W, Harrop J, et al. Relationship of chronic pelvic pain to psychiatric diagnosis and childhood sexual abuse. <i>Am J Psychiatry 1988</i> ; 145: 75 - 80.
[63]	Arnold RP, Rogers D, Cook DA. Medical problems of adults who were sexually abused in childhood. <i>BMJ 1990</i> ; 300: 705 - 708.

- [64] Koss MP, Koss PG, Woodruf WJ. Deleterious effects of criminal victimization on women's health. Arch Intern Med. 1991; 151: 342 - 347.
- [65] Walker EA, Katon WJ, Roy-Byrne PP. Histories of sexual victimization in patients with irritable bowel syndrome or inflammatory bowel disease. *Am J Psychiatry 1993*; 150: 1502 - 1506.
- [66] Moeller TP, Bachmann GA, Moeller JR. The combined effects of physical, sexual and emotional abuse during childhood: long term health consequences for women. Child *Abuse Negl. 1993*; 17: 623 640.
- [67] Lodico MA, DiClemente RJ. The association between child hood sexual abuse and prevalence of HIV related risk behaviors. *Clin. Pediatr. 1994*; 498 502.
- [68] Golding JM. Sexual assault history and physical health in randomly selected Los Angeles women. *Health Psychol. 1994*; 13: 130 138.
- [69] Taylor ML, Trotter DR, Csuka ME. The prevalence of sexual abuse in women with fibromyalgia. *Arthritis Rheum. 1995*; 38: 229 234.
- [70] Golding JM. Sexual assault history and women's reproductive and sexual health. Psychol. *Women Quart. 1996;* 20: 101 - 121.
- [71] Golding JM. Sexual assault history and limitations in physical functioning in two general population samples. *Res. Nurs. Health 1996;* 19: 33 44.
- [72] McCauley J, Kern DE, Kolodner K, et al. Clinical charasteristics of women with a history of childhood abuse: unhealed wounds. *JAMA 1997;* 277: 1362 1368.
- [73] Sharma BR, Gupta M. Child rape save the innocents. Social Welfare 2004; 51 (9): 3 –
   8
- [74] Walker EA, Gelfand A, Katon WI, et al.. Adult health status of women with histories of childhood abuse and neglect. *Am J Med 1999*; 107: 332 – 339
- [75] Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. Am J Prevent Med 1998; 14: 245 – 258
- [76] McEwen BS. Protective and damaging effects of stress mediators. *NEJM 1998*; 338: 171 - 179
- [77] Gunnar MR. Quality of early care and buffering of neuroendocrine stress reactions: potential effects on the developing human brain. *Prevent Med 1998*; 27: 208 211
- [78] Taylor SE. The lifelong legacy of childhood abuse. *Am J Med 1999*; 107: 399 400
- [79] Green A. self-destructive behavior in battered children. Am J Psychiatry 1978; 135: 579 – 582.
- [80] Deykin E, Alpert J, McNamarra J. A pilot study of the effect of exposure to child abuse or neglect on suicidal behavior. *Am J Psychiatry 1985*; 142: 1299 – 1309
- [81] Farber E, Kinast C, McCord W, et al. Violence in families of adolescent runaways. *Child Abuse Negl 1984*; 8: 295 – 299
- [82] Kaplan S, Pelcovitz D, Salzinger S et al. Adolescent physical abuse: risk for adolescent psychiatric disorders. Am J Psychiatry 1998; 155: 954 – 959
- [83] Hawton K, Roberts J, Goodwin G. The risk of child abuse among mothers who attempt suicide. Br. J Psychiatry 1985; 146: 486 – 489.
- [84] Kaplan S, Pelcovitz D, Salzinger S et al. Adolescent physical abuse and suicide attempts. J Am Acad of Child Psychiatry 1997; 36: 799 – 807.
- [85] Sharma BR, Sharma V, Harish D. Suicides in Northern India causes, Methods used and Prevention thereof. J Med. Sci. Law 2003; 43 (3): 221 - 229.

[86]	Schmidt P, Muller R, Dettmeyer R, Madea B. Suicide inchildren, adolescents and
F0 = 3	young adults. Forensic Science International 2002; 127: 161 - 167
[87]	Mrazek PJ, Haggerty RJ. Reducing risks for mental disorders: frontiers for preventive intervention research. Washington DC: <i>National Academy Press, 1994</i>
[88]	Holden EW, Nabors L. the prevention of child neglect. In Dubowitz H. ed Neglected
[]	children. Thousand Oaks CA: sage Publications, 1999; 174 – 190
[89]	Jones DPH. Treatment of the child and the family where child abuse or neglect has
	occurred. In R. Helfer, R Kempe and R Krugman eds. <i>The Battered Child</i> , University
[00]	of Chicago Press, (5 <sup>th</sup> edn) 1997, 521 – 542 Becker J. Offenders: characteristics and treatment. Future of Children 1994; 4: 176 –
[90]	197
[91]	Kassin SM. The psychology of confession evidence. Am Psychol 1997; 52: 221 – 233
[92]	Bruck M, Ceci SJ, Hembrook H. Reliability and credibility of young children's report:
	from research to policy and practice. <i>Am Psychol 1998</i> ; 53: 136 – 151
[93]	Ceci S, Bruck M. Suggestibility of the child witness; a historical review and synthesis. <i>Psychol Bull 1993</i> ; 113: 403 – 439
[94]	Saywitz KJ, Moan-Hardie S. Reducing the potential for distortion of childhood
	memories. Conscious Cogn 1994; 3: 408 - 425
[95]	Saywitz KJ, Goodman GS, Nicholas E, et al. Children's memories of physical
	examinations involving genital touch: implications for reports of child sexual abuse. $J$
50 (3	<i>Consult Clin Psychol 1991</i> ; 59: 682 – 691
[96]	Goodman GS, Batterman-Faunce JM, Kenney R. <i>Optimizing children's testimony:</i>
	research and social policy issues concerning allegations of child sexual abuse. In Cicchetti D, Toth S eds Child Abuse, Child Development and Social Policy. Norwood,
	NJ, Ablex, 1992
[97]	Saywitz KJ, Goodman GS. Interviewing children in and out of court: current research
L ]	and practical implications. In Briere J, Berliner L, Bulkley J. et al., eds The APSAC
	hand Book on Child maltreatment. Thousand Oaks, CA, Sage, 1996, 297-318
[98]	Reed LD. Findings from research on children's suggestibility and implications for
	conducting child interviews. Child Maltreatment 1996; 1: 105-120
[99]	American Professional Society on the Abuse of Children, Guidelines for psychosocial
	evaluation of suspected sexual abuse in young children, Chicago, American
[100]	Professional Society on the Abuse of Children, 1997
[100]	American Academy of Pediatrics Committee on Child Abuse and Neglect, Guidelines for the evaluation of sexual abuse of children: subject review. <i>Pediatrics 1999</i> ; 103:
	186 - 191.
[101]	Gill P, Jeffreys AJ, Werrette DJ. Forensic application of DNA fingerprints. <i>Nature</i>
L ' J	<i>1985</i> ; 318: 577 - 579
[102]	Christian CW, Lavelle JM, DeJong AR, et al. Forensic evidence collection in
	prepubertal victims of sexual assault. Pediatrics 2000; 106: 100 - 104
[103]	Kalichman SC. Mandated reporting of suspected child abuse: ethics, law and policy.
51043	Washington, DC: American Psychological Association, 1993
[104]	Mueller CB, Kirkpatrick LC. Federal evidence, 2 <sup>nd</sup> ed. Rochester, NY: Lawyers
[105]	Cooperative Publishing, 1994 Myers JEB. <i>Evidence in child abuse and neglect cases, 2<sup>nd</sup> ed.</i> New York: Aspen Law
[103]	and Business, 1997

[106] Myers JEB. Legal issues in child abuse and neglect practice, 2<sup>nd</sup> ed. Thousand Oaks, CA: Sage, 1998

Chapter 19

# ADOLESCENT SEXUAL RISK BEHAVIOR: ASSOCIATIONS WITH FAMILY, PEERS AND SCHOOL ENVIRONMENT\*

Aleksandra Jovic-Vranes<sup> $1,\dagger$ </sup> and Boris Vranes<sup>2</sup>

<sup>1</sup> Institute of Social Medicine, Medical School University of Belgrade, Dr Subotica 15, 11000 Belgrade Serbia
<sup>2</sup> Institute of Gynecology and Obstetrics, Clinical Center of Serbia

# ABSTRACT

Objectives

During the dynamic period of adolescence when the passage from childhood to maturity takes place, sexuality takes on new dimensions; feelings become more intense, relationships become more complex, and the consequences of sexual behavior are radically altered. In general, earlier puberty, later marriage, a decline in the family leading to less control and more autonomy, and intense exposure to sexual stimuli via the mass media and travel across cultural boundaries have made pre-marital adolescent sexual activity more common. Adolescent sexual risk behavior places them at risk of unwanted pregnancy and childbirth, induced abortion in hazardous circumstances, HIV infection, and other sexual transmitted diseases. Objective of the study was to achieve a picture of adolescent sexual risk behavior. It also provides better material for the planning of specific preventive activities.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> e mail: aljvranes@yahoo.co.uk

#### Methods

Random cluster sample consisted of 1540 15-year old adolescents, 822 female and 718 male. The multiple choice questionnaire included 119 items, covering topics such as: demographic characteristics, psychosocial determinants of health, leisure time behavior, family and peer context, risk behavior (tobacco use, alcohol and drugs consumption, sexual activity), and perceptions of school and the school's influence. Descriptive statistics, Chi-square test and logistic regression were applied for statistical analyzes.

#### Results

Thirteen percent of adolescent reported that they had sexual intercourse before age 15 years. Average age of first sexual intercourse was 13.88. Average number of sexual partners they had was 3.10. One out of three never used condom. Eight percent reported multiple sex partners, and fourteen percent alcohol use before intercourse. Factors associated with adolescent sexual risk behavior include other form of risk behaviors, family and peer context, and school environment.

#### Conclusion

Adolescent relationships are complex and future research should consider not only causality of adolescent sexual risk behavior, but also the etiology of the satellite behaviors.

Keywords: adolescence, sexual risk behavior, family, peers, school

#### INTRODUCTION

In bringing medicine to the community the provision of health services for children of school age should have a high priority. The requirements of children are basically similar everywhere, and a study of the historical growth and geographical pattern of school health services reveals that they have developed on very similar lines in most parts of the world.

#### Historical

It was not until the nineteenth century that serious study was given to the health of children at school.

In 1812 James Ware reported on the eyesight of school children in London and of students at Oxford University. In 1840 several doctors were appointed in a number of training colleges in Sweden. In 1866 Herman Cohn investigated the eyesight of over 10 000 children in Breslau; by 1883 he was urging the appointment of school doctors and had the eminent support of Virchow. The first school doctor in Germany was, in fact, appointed that year in Frankfurt-am-Main. Two years later one was appointed in Lausanne. In 1888 the Swedish Government inquired into the physical condition of over 11 000 Swedish children. At about

the same time school medical inspection was started in all the Departments of France. In 1895 six school physicians were appointed to supervise the elementary schools in Moscow [1]. Adolescence

Adolescents are important! Throughout the world, they constitute a large proportion of any population, and their absolute numbers are increasing. Financially, they are the most expensive group of any age with regard to the investment of the community in their training and education, and socially they are the most commercially powerful, exploited and important since patterns of behavior established in adolescence inevitably continue into adulthood [2]. Adolescents are also politically important, but despite their enormous power to influence the community of the present day and the future, society has not yet learned to provide its adolescents with a structure of normal care during the period of their transition from child to adult [2].

Indeed, adolescence is a critical age for the development of coping behaviors and responses. This period is characterized by rapid physical, psychological, sociocultural and cognitive changes, and is unfortunately fraught with many threats to health. Much of the adverse health consequences experienced by adolescents are, to a large extent, the result of "risk behaviors". The initiation of risky behavior is occurring at a progressively younger age [3,4]. As a result of these trends, many adolescents may be vulnerable to experiment and initiate risk behaviors that have deleterious consequences during adolescence and later.

Risk behaviors and their associated adverse health outcomes represent a serious threat to life time health [5,6].

#### Reproductive Health in Adolescence

The health and well-being of adolescents is closely intertwined with their physical, psychological and social development, but this is put at risk by sexual and reproductive health hazards which are increasing a much of the world. Changes in population growth and distribution, the rise of telecommunications, the increase in travel and a decline in the family, as well as a generally earlier start of menarche and later age of marriage are contributing to an increase in unprotected sexual relations before marriage. This combined with risks from early marriage, result in early or unwanted pregnancy and childbirth, induced abortion in hazardous circumstances, and sexually transmitted diseases, including HIV infection leading to AIDS [7].

Sexuality is a fundamental quality of human life, important for health, happiness, individual development, and indeed for the preservation of the human race. During the dynamic period of adolescence in which the passage from childhood to maturity takes place, sexuality takes on new dimensions; feelings become more intense, relations become more complex, and the consequences of sexual behavior are radically altered. This not only affects the behavior of young people but also of those who interact with them, their families and peers, and those who work in health, education, youth, social welfare, and other sectors [8].

Adolescent sex has become a complex phenomenon among the researches, due to high sensitivity of this area. During the last two to three decades, the importance of this issue has increased because, its association with teenage pregnancies, sexual abuse such as rape and molestation, illegal birth, illegal and unsafe abortion and risk of STD and HIV-AIDS among adolescent. The basic problems of such studies are gathering of reliable information [9].

Unfortunately, there are still large numbers of misconceptions among adolescents towards sexual behavior. Some of them include that all teens are having sex, having sex makes you an adult, something is wrong with an older teen (17-19) who is not having sex, a girl can't get pregnant if she's menstruating, or a girl can't get pregnant if it's her first time [10].

#### Which Factors Influence Adolescents' Decisions about Sex?

Relevant factors include both risk factors and protective factors, which may be equally important in terms of their relevance. Risk factors are those that encourage behavior that could result in a pregnancy or sexually transmitted disease (STD) or, conversely, that discourage behavior that could prevent them. Protective factors are those that discourage behavior that could lead to a pregnancy or STD or that encourage behavior that can help prevent them. Put another way, as the number of risk factors in a adolescent's life increases and/or the number of protective factors decreases, the likelihood that he/she will have sex, become pregnant/cause a pregnancy, or contract an STD increases [11].

More than 400 factors are identified that affect one or more sexual behaviors (the initiation of sex, frequency of sex, number of sexual partners, use of condoms, and use of other contraceptives) or consequences of those behaviors (pregnancy, childbearing or STD).

Important risk and protective factors include characteristics of the adolescents' states, communities, families, friends and peers, romantic partners, and the adolescents themselves. Factors also involve adolescents' relationships with these important individuals or organizations in their environment. Some factors directly involve sexuality, while others do not [11,12].

These risk and protective factors may be grouped into four key themes:

- 1. Individual biological factors (e.g. age, physical maturity and gender)
- 2. Disadvantage, disorganization and dysfunction in the lives of the adolescents themselves and their environments (e.g. rates of substance abuse, violence, and divorce; also levels of education)
- 3. Sexual values, attitudes, and modeled behavior (e.g. adolescents' own values about sexual behavior as well as those expressed by parents, peers, and romantic partners)
- 4. Connection to adults and organizations that discourage sex, unprotected sex, or early childbearing. (e.g. attachment to parents and other adults in their schools and places of worship) [11].

## METHODS

A survey was conducted applying the WHO research protocol for a cross-national survey "Health behavior in school aged children" (HBSC), as a cross-sectional study among Belgrade adolescents. It was a research project that aimed to gain new insight into and

increased understanding of health and risk behavior, lifestyles and their context regarding young people.

#### Sample

A total of 64 public schools, and school classes at the appropriate grade levels were randomly selected. Thus, cluster sample design was used; and once the first level of sampling occurred at the school or school class level, then all adolescents in an appropriate age group were surveyed. Minimum sample size derived from HBSC recommended protocol was 1536 adolescents. This sample size assumes 95% confidence interval and a design effect of 1.44 [5]. A total of 1540 adolescents, aged 15 years, were included in this study.

#### Instrument

The multiple choice questionnaire for adolescents includes 119 items, covering topics such as: demographic characteristics, psychosocial determinants of health, leisure time behavior, family and peer context, risk behavior (tobacco use, alcohol and drugs consumption, sexual activity), and perceptions of school and the school's influence.

#### Measures

The data collection was performed by specialized staff, trained to address issues related to risk behavior, and to approach children in schools. These research assistants documented the process of data collection in a classroom report. The completion of the questionnaire by the participants took ca. 45 minutes. Teachers were absent from the classroom during this process. The project team guarantees anonymity of the answers provided by students.

Risk behavior was estimated through questions regarding sexual activity, smoking, alcohol and drug consumption. Adolescents were asked whether they had ever had sexual intercourse. Sexually active adolescents were asked how many times in lifetime they had engaged in sexual intercourse, how many partners they had, consistency of contraceptive use, and practice of other risk behavior. To determinate the prevalence of current tobacco use adolescents were asked, "how often do you smoke at present?" Only those who reported smoking at least once a week were referred to as current smokers. In addition the adolescents participating in the survey were asked about their use of alcohol. Specifically, they were asked how frequently they took alcoholic drinks, and whether they had ever been drunk. Current consumption of alcohol was defined as drinking some kind of alcoholic beverage at least every week. The term "illicit drugs" includes substances like: marijuana, a combination of alcohol and pills, inhalants and cocaine. Any frequency of the use of drugs above mentioned in a life time were taken into account and analyzed. Questions regarding family context included family structure, communications with parents, and family support. Questions addressed peer context included number of close friends, communication and

spending time with friends, and attitudes towards classmates. School environment was estimated through questions about attitudes towards school and teachers.

## Statistical Analysis

Data were analyzed by methods of descriptive statistics, chi-square test regarding differences by sex, and multivariate logistic regression. Associations between variables were initially tested using Chi-square tests. Logistic regression was used to identify the odds ratios for adolescent early onset of sexual behavior. Independent variables were related to other form of risk behavior, family and peer context and school environment. The odds ratio (OR) and 95% confidence intervals were obtained, and p value <0,05 was taken as the minimum level of significance.

#### Results

A study included a total of 1540 15-year old adolescents, 822 female and 718 male. Overall questionnaire was successfully conducted in 89.16% of the study population.

#### School as a Setting

Girls reported better school performance compared to their boys classmates. Asked how they feel about school at present seventy per cent of adolescents answered that they like school a bit, with no statistical difference by gender. Almost half of adolescents agree that their school is nice place to be, and that they belong at that school. Asked about their teachers, 10.3% of adolescents strongly agree that they are encouraged to express their own views in classes by teachers, and 10.4% that teachers treated them fairly. If they need extra help at school, 52.4% of respondents believe that they can get it by teachers. One out of three adolescent feel that their teachers expect too much of them at school, and 19.2% of them feel a lot pressured by the schoolwork they have to do. Almost half of adolescent neither agree nor disagree that their teachers are interested in them as a person. Less than 4% strongly agree with this statement. One out of five of respondents never or rarely feel tired when go to school in the morning, and 15.9% feel tired every day. Great number of boys had opinion (33.6%) that going to school is boring. Eleven per cent of girls had opposite opinion that going to school is never boring. Differences between boys and girls were also found in skipping classes or school. Twelve per cent of boys did it at least once, and 5% four or more times. Most of adolescent (78.7%) feel safe at school (Table 1).

## Parent Relationships

Most of adolescent (84.2%) live with both biological parents. Boys reported better communication with father compare with girls who reported better communication with mother.

One out of three girls has difficult communication with father about things that really bother her, and 11.4% of boys recognized the same problem in communication with mother. Most of respondents (84.4%) think if they have problems at school, their parents are always ready to help, 71.2% have parents who are willing to come to school to talk with teachers, and 85.2% are encouraged to do well at school by their parents. That their parents expect too much of them at school believe 29.8% of adolescents (Table 2).

variable	gender		total	$X^2$	р
	boys n (%)	girls n (%)			
Good school performance	315 (43.9)	512 (62.4)	827 (53.8)	89.09	0.000
Like school a bit	510 (71.9)	563 (68.9)	1073 (70.3)	2.41	0.492
School is a nice place to be	337 (48.0)	420 (51.3)	757 (49.8)	1.68	0.194
I feel I belong at this school	404 (57.2)	464 (56.9)	868 (57.1)	0.01	0.909
I am encouraged to express my	84 (11.9)	73 (9.0)	157 (10.3)	16.99	0.002
own views in my class(es) Teachers treat us fairly	96 (13.5)	62 (7.6)	158 (10.4)	31.60	0.000
When I need extra help, I can get it	347 (49.2)	450 (55.3)	797 (52.4)	5.70	0.000
My teachers are interested in me as a person	33 (4.7)	24 (3.0)	57 (3.8)	12.52	0.014
Rarely or never feel tired when they go to school in the morning	169 (23.6)	162 (19.7)	331 (21.5)	7.96	0.047
Feel safe at school	559 (78.3)	650 (79.1)	1209 (78.7)	4.08	0.395
Going to school is often boring	240 (33.6)	171 (20.9)	411 (26.7)	40.79	0.000
Skipping classes at least once	85 (11.9)	79 (9.7)	164 (10.7)	22.12	0.000

## Table 1. School environment - gender differences

## Table 2. Family context – gender differences

variable	gen	der	total	X²	р
	boys n (%)	girls n (%)	-		
I live with both parents	602 (84.1)	691 (84.4)	1293 (84.2)	5.97	0.202
Easy communication with father	487 (70.0)	410 (50.6)	897 (59.6)	58.22	0.000
Easy communication with mother	569 (81.3)	694 (85.4)	1263 (83.5)	4.53	0.033
If I have problems at school, my parents are ready to help	605 (85.2)	685 (83.6)	1290 (84.4)	5.18	0.159
Parents are willing to come to school to talk to teachers	519 (72.9)	572 (69.7)	1091 (71.2)	3.32	0.345
Parents encourage me to do well at school	599 (84.0)	707 (86.2)	1306 (85.2)	1.55	0.671
Parents expect too much of me at school	245 (34.6)	210 (25.6)	455 (29.8)	14.56	0.000

variable	gen	gender		$X^2$	р
	boys n (%)	girls n (%)	-		
Three or more close friends	580 (81.2)	522 (63.6)	1102 (71.8)	61.39	0.000
Easy to make new friendship	656 (91.9)	731 (89.3)	1387 (90.5)	3.04	0.081
Easy communication with friends	254 (37.4)	385 (47.5)	639 (42.9)	45.52	0.000
of the same sex					
Easy communication with friends	183 (27.0)	182 (22.8)	365 (24.7)	11.49	0.022
of the opposite sex					
Spending time with friends after	185 (26.3)	111 (14.0)	296 (19.8)	57.31	0.000
school 4-5 days a week					
Students enjoy always being	300 (42.6)	252 (31.1)	552 (36.5)	27.19	0.000
together					
Most of the students are kind and	569 (80.5)	687 (84.4)	1256 (82.6)	4.04	0.045
helpful					
Students accept me as I am	659 (93.3)	760 (93.3)	1419 (93.3)	0.01	0.943
Take part in bullying other	56 (7.9)	25 (3.0)	81 (5.3)	49.28	0.000
students even once					

## Table 3. Peer context – gender differences

#### Peer Relationships

Seventy two per cent of adolescents have three or more close friends. Being withdrawn by classmates was present in 1.6% of adolescents. Making new friendships was easy for 90% of both girls and boys. Girls reported better communication with friends of the same sex. Almost half of them (47.5%), can very easy to talk about things that really bother them with friends of the same sex. One out of four of respondents have very easy communication with friends of the opposite sex with no statistical difference according to gender. There was a significant difference according to sex noted in the way students spent time with friends right after school. 26.3% of boys and 14% of girls went out with friends four to five times a week. Forty three per cent of boys believe that the students in their class(es) always enjoy being together compare to thirty per cent of girls. Twenty five per cent of girls think that is possible only sometimes. Good opinion about their classmates was noticed among adolescents, with no difference according to sex. That most of the students in their class(es) are kind and helpful reported 82.6% of adolescents, and 93.3% that other students accept them as they are. Eight per cent of boys have taken part in bullying other adolescents in school at least once in that term in comparison to 3.0% of girls (Table 3).

#### Substance Use

The current prevalence of cigarette smoking was 14,8% with non significant difference according to sex observed (129 female and 99 male smokers).

5,6% of students consumed alcohol regularly. The percentage of students reported to having "been drunk" were 34,9% of males and 13,2% of females, with significant difference by sex. Having five or more drinks in a row was found commonly among students who use alcohol with 25% of males, regularly drinking alcohol, in particular on weekends.

variable	gen	der	total	$X^2$	р
	boys n (%)	girls n (%)			-
Substance use					
cigarette smoking	99 (13.8)	129 (15.7)	228 (14.8)	1.10	0.294
regular alcohol consumption	41 (8.1)	15 (3.1)	56 (5.6)	11.71	0.001
been drunk	189 (34.9)	70 (13.2)	259 (24.2)	68.63	0.000
have five or more drinks in row	132 (24.2)	44 (8.2)	176 (16.3)	84.51	0.000
drug use	98 (13.6)	69 (8.5)	167 (10.8)	10.95	0.001
Sexual behavior					
sexual intercourse before age 15	160 (23.8)	19 (2.5)	179 (12.6)	146.21	0.000
years	× ,				
first sexual intercourse was	117 (84.2)	12 (92.3)	129 (84.9)	1.66	0.647
wished by both partners	× ,				
had a relationship with their	122 (81.3)	8(61.5)	130 (79.8)	2.90	0.088
partners in the moment of first	× ,				
sexual intercourse					
main reason for the first sexual	61 (40.9)	5 (38.5)	66 (40.7)	1.28	0.973
intercourse - curiosity		× ,			
main outcome that bothers	91 (62.3)	10 (76.9)	101 (63.5)	4.54	0.338
adolescents- pregnancy			· · · · ·		
condom use	95 (64.2)	12 (100)	107 (66.9)	6.43	0.01
multiple sex partners	13 (8.8)	0 (0)	13 (8.1)	1.24	0.265
alcohol use before intercourse	20 (13.4)	2 (15.4)	22 (13.6)	0.14	0.932

## Table 4. Adolescent substance use and sexual behavior - gender differences

Lifetime experience of illicit drug use was 14% for boys and 8,5% for girls, with significant difference according to sex. The most prevalent illicit drug was marijuana. Half of the students who had experience with drugs used marijuana, 11,3% used a combination of drugs (alcohol and pills), 9,4% used inhalants and 4% used cocaine (Table 4).

#### Sexual Behavior

Thirteen percent of adolescent reported that they had sexual intercourse before age 15 years, more boys (23.8%) than girls (2.5%). Average age of first sexual intercourse was 13.88. Average number of sexual partners they ever had was 3.20 for boys, and 1.50 for girls. Most of adolescents (84.9%) reported that their first sexual intercourse was wished by both partners, and 79.8% that they had a relationship with their partners in the moment of first sexual intercourse. For the majority of respondents main reasons for the first sexual intercourse were curiosity (40.7%), passion (19.1%), and to stay in relationship with partner (19.8%). Five percent did it because it was peer expectation, and ten per cent of adolescents out of love. Other reasons were physical attraction (3.7%) and under alcohol or other drugs (1.9%). Pregnancy was the main concern for most of adolescents (63.5%). One third was afraid of AIDS, 1.9% of parents' disapproval, 2.5 reported other reasons, and less than one per cent was concerning other sexually transmitted diseases. Despite these facts, use of contraceptive methods during each sexual intercourse was inconsistent and sporadic. Condom was the most used contraceptive method by adolescents. Girls always use condoms, and one

out of three boys did not use it at all. Other methods were contraceptive pills (5.6%), and interrupted coitus (5.6%). Rest of the adolescent reported some other methods. Eight percent of boys and none of the girls reported multiple sex partners. Fourteen percent of respondents admitted alcohol use before intercourse, more girls than boys (Table 5).

Table 5. Odds ratios of adolescent early sexual intercourse, according to factors from
family, peer, school setting and other form of risky behavior

variable	OR*	95% CI**	р
I live without both parents	1.54	1.04-2.26	0.033
Difficult communication with mother	1.75	1.20-2.55	0.004
Parents don't encourage me	2.16	1.23-3.80	0.008
Difficult communication with friends of the opposite sex	0.49	0.34-0.71	0.000
Three or more close friends	1.56	1.07-2.29	0.022
Difficult to make new friendship	0.36	0.16-0.78	0.010
Spending time with friends after school	1.39	1.28-1.54	0.000
School is not a nice place to be	1.41	1.02-1.94	0.037
I don't feel that I belong at this school	1.45	1.06-1.99	0.022
Cigarette smoking	2.74	1.91-3.94	0.000
Regular alcohol consumption	6.43	3.61-11.45	0.000
Been drunk	7.45	5.15-10.79	0.000
Drug use	3.46	2.34-5.13	0.000

\* Odds ratio from multiple logistic regression models.

\*\* Confidence Interval.

#### Associations with Adolescent Sexual Risk Behavior

Using logistic regression analysis we found that factors associated with early onset of adolescent sexual behavior include other form of risk behaviors, family and peer context, and school environment. Boys and girls who live in families without both biological parents [odds ratio (OR)=1.53, 95% confidence interval (CI)=1.04-2.26] and in families without support (OR=2.16, 95%CI=1.23-3.80) often have firs sexual intercourse before age 15 years compare with others. Difficult communication with mother (OR=1.75, 95%CI=1.20-2.55) and easy communication with friends of the opposite sex (OR=0.49, 95%CI=0.34-0.71) were found to be associated with early sexual activity. Other factors regarding peer context were large number of close friends (OR=1.65, 95%CI=1.07-2.29), easy way of making new friendships (OR=0.36, 95%CI=0.16-0.78), and spending a lot of time with friends after school (OR=1.39, 95%CI=1.28-1.54). Attitudes towards school also influence sexual activity among adolescents. Those who did not think that school is nice place to be (OR=1.41, 95%CI=1.02-1.94), and that they don't belong at that school (OR=1.45, 95%CI=1.06-1.99) experienced with sexual behavior earlier. Other forms of risk behavior, such as smoking cigarettes (OR=2.74, 95%CI=1.91-3.94), alcohol (6.43, 95%CI=3.61-11.45) and drug consumption (OR=3.46, 95%CI=2.34-5.13) are related to a younger age at the onset of sexual activity (Table 5).

## CONCLUSION

Sexual and reproductive ill-health mostly affects women and adolescents. Women are disempowered in much of the developing world and adolescents, arguably, are disempowered everywhere. Sexual and reproductive health services are absent or very often of poor quality and underused in many countries because discussion of issues such as sexual intercourse and sexuality make people feel uncomfortable. The increasing influence of conservative political, religious, and cultural forces around the world threatens to undermine progress made since 1994, and arguably provides the best example of the detrimental intrusion of politics into public health [13].

According to a growing body of research, most young people begin having sexual intercourse during their teenage years. Current data suggest that slightly more than half of females and nearly two-thirds of males have had intercourse by their 18th birthday. In the last several decades there have been substantial increases in the proportion of adolescents who report sexual activity at each year of age. Increases have been greatest among females, especially among young females. Thus, more than twice as many females ages 14, 15, and 16 are sexually active now, compared with young women of the same ages just 15 years ago. Moreover, on average, there are seven years for women and ten years for men between first intercourse and first marriage [14]. Our findings show that thirteen percent of adolescent had sexual intercourse before age 15 years, more boys (23.8%) than girls (2.5%), and average age of first sexual intercourse was 13.88. Average number of sexual partners among adolescents in our study was 3.20 for boys, and 1.50 for girls. Data from Center for disease control and prevention suggest that in 2005, 47% of high school students had ever had sexual intercourse, and 14% of high school students had had four or more sex partners during their life [15]. In addition, young people use alcohol and other drugs at high rates. Adolescents are more likely to engage in high-risk behaviors, such as unprotected sex, when they are under the influence of drugs or alcohol [16]. In 2005, 23% of high school students who had sexual intercourse during the past three months drank alcohol or used drugs before last sexual intercourse. 34% of currently sexually active high school students did not use a condom during last sexual intercourse [15]. In our study fourteen percent of adolescents admitted alcohol use before sexual intercourse, more girls than boys. One out of three boys did not use condom at all.

Many variables are related to the timing of first sexual intercourse. On average, males begin having sex at younger ages than females. In addition families and schools, provide an environment within which adolescents make decisions related to sexual activity. Unconventional psychosocial attitudes and behaviors--as reflected by early use of alcohol, tobacco and other drugs, school problems, delinquency, and physical aggression--are associated with earlier onset of adolescent sexual intercourse. Parents' marital disruption and living with a single parent have been found to be associated with earlier onset of adolescent sexual behavior. Similarly, having sexually active siblings and friends is strongly related to a younger age at the onset of sexual activity [14, 17].

Our findings are consistent with other studies [14,17,18]. We found that factors associated with early onset of adolescent sexual behavior include other form of risk behaviors, family and peer context, and school environment. Adolescents who live in families without both biological parents, in families without support, and who have difficult communication with mother often have firs sexual intercourse before age 15 years compare

with others. Other factors regarding peer context were easy communication with friends of the opposite sex, large number of close friends, easy way of making new friendships, and spending a lot of time with friends after school. Attitudes towards school also influence sexual activity among adolescents. Those who did not think that school is nice place to be, and that they don't belong at that school experienced with sexual behavior earlier. Other form of risk behavior, such as smoking cigarettes, alcohol, and drug consumption are related to a younger age at the onset of sexual activity.

When interpreting the findings of this study, a number of limitations should be considered. When responding to questions about risk behaviors, some adolescents might have felt uncomfortable and untrusfull and had the need to give responses that were socially acceptable. Moreover, the cross-sectional study design does not allow us to establish causal relationships among variables. Despite these limitations, we believe that there are several important trends in the data. Our data are consistent with other studies, and our results may stimulate further research, and provide valuable reference point for future studies of this issue.

To prevent problems before they happen, we must identify the factors that put adolescents at risk and either eliminate, reduce, or somehow buffer the effect of exposure to the risk. Risk for adolescent problem behavior exists in many domains of life. However, it is important to remember that exposure to a single risk factor does not condemn an individual to problem behavior [19]. The findings summarized above have important implications for parents, teachers, clinicians and others responsible for the well-being of adolescents. There is an urgency to develop and implement behavioral interventions that motivate adolescents to adopt and/or maintain risk taking behavior prevention practices. Behavioral change interventions have attempted to reduce or eliminate the consequences of adolescents risk behavior. Interventions have been attempted to delay the initiation of sexual activity, to improve contraceptive use among sexually active adolescents, to (in some cases) influence pregnancy resolution decisions among those who become pregnant, and to eliminate or reduce sexual transmitted diseases. Adolescents who begin having sexual intercourse need to understand the importance of using an effective contraceptive every time they have sex. This requires convincing sexually active adolescents who have never used contraception to do so. In addition, sexually active adolescents who sometimes use contraceptives need to use them more consistently (every time they have sex) and use them correctly. Finally, sexually active adolescents need to take actions to prevent sexually transmitted infections, as well as unintended pregnancy.

## REFERENCES

- [1] Henderson P. The health of the school child. *Public health* 1971; 85 (2): 58-66.
- [2] Bennett LD. Worldwide problems in the delivery of adolescent health care. *Publ Hlth Lond* 1982; 96: 334-340.
- [3] Office of Disease Prevention and Health Promotion. Intervention for adolescents at risk. *Prevention Report*, 1993 February/March (pp. 1-2). Washington, DC: United States Department of Health and Human Services, U.S. Public health Service.

- [4] DiClemente R, Hansen BW, Ponton EL. Adolescent at Risk, a Generation in Jeopardy. In: *Handbook of Adolescent Health Risk Behaviour*, New York; Planum Press; 1996: 1-52.
- [5] Currie C. Health Behavior in School-Aged Children: Research Protocol for the 1997-98 Survey. Edinburgh: Edinburgh University; 1998.
- [6] Jovic Vranes A, Vranes B, Marinkovic J, Cucic V. Adolescent substance abuse, the importance of family, school and peers: data from the health behavior in school children. *Soz Praventivmed*. 2005; 50(2):119-24.
- [7] Friedman HL. Reproductive health in adolescence. *World Health Stat Q.* 1994; 47 (1): 31-5.
- [8] Friedman HL. Changing patterns of adolescent sexual behavior: consequences for health and development. J Adolesc Health 1992; 13(5): 345-50.
- [9] Nainakwal R, Nainakwal M. Adolescent sexual behavior, contraception, and unsafe abortion (A study on the college student in Mumbai, India). University of Mumbai 2001. Available from: URL: http://www.iussp.org/Brazil2001/s80/s85 p06 nainakwal.
- [10] Focus Adolescent Services. *Teen Sexual Behaviors: Issues and Concerns.* 1999. Available from: URL: http://www.focusas.com/SexualBehavior.html.
- [11] Kirby D, Lepore G, Rayan J. Sexual Risk and Protective Factors. Factors affecting teen sexual behavior, pregnancy, childbearing and sexual transmitted disease: Which are important? Which can you change. ETR Associates, Washington 2005.
- [12] Sallis JF, Nader PR. Family determinants of health behavior. In: D.S. Gochman, *Health Behaviour: Emerging Research Perspectives*, Plenum Press, New York ;1988.
- [13] Glasier A, Gulmezoglu AM, Schmid GP, Moreno CG, Van Look PF. Sexual and reproductive health: a matter of life and death. *Lancet*. 2006 Nov 4; 368(9547):1595-607.
- [14] Moore AK, Miller CB, Sugland WB, Morrison RD, Glei AD, Blumenthal C. Adolescent sexual behavior, pregnancy and parenthood. U.S. Department of Health and Human Services. 2006. Available from: URL: http://www. aspe.os.dhhs.gov/hsp/cyp/xsteesex. htm.
- [15] CDC. Youth Risk Behavior Surveillance—United States, 2005. Morbidity & Mortality Weekly Report 2006; 55(SS-5):1–108.
- [16] Leigh B, Stall R. Substance use and risky sexual behavior for exposure to HIV: issues in methodology, interpretation, and prevention. *American Psychologist* 1993; 48:1035– 1043.
- [17] Klingon YS, O Sullivan AL. The family as a protective asset in adolescent development. *J holist Nurs* 2001; 19(2): 102-21.
- [18] Jacar J, Dittus JP, Gordon VV. Maternal Correlates of Adolescent Sexual and Contraceptive Behavior. *Family Planning Perspectives* 1996; 28: 159-165.
- [19] Hawkins JD. *Risk for Health and Behavior Problems*. Fifteenth Annual Gsela Lectureship, University of Minnesota 1993.

Chapter 20

## SEXUAL AGGRESSIONS AMONGST UNDERAGED: VULNERABILITIES, RISK FACTORS, SIGNS OF CALLING FOR HELP\*

**Philippe Bessoles** 

"Victimology and Criminology" University Pierre Mendés France, Grenoble II, France

## In collaboration with Myriam Linnell-Boutaud

Keywords: Victimology, traumatism, criminology, sexuality, resilience.

## **VULNERABILITY AND TRAUMA**

Post-trauma notion of vulnerability, following a sexual aggression, appears too general in infantile clinical psychopathology to be sufficiently operating. Differences regarding criteria for symptomatological assessments, differential diagnosis with post-trauma disorders, personality changes inherent to child development, specificity (or not) of a post-traumatic semiology inherent to under fifteen, etc. add to the complexity of the question. According to research, vulnerability can be, at once, synonym of psycho-social factors of risk, signs of victimarius call for help or badly circumscribed psychological vulnerabilities. International studies often diverge as investigation methodologies are so contradictory and epidemiological data not consensual enough... Notions of abuse, touching, incest, remain blur and badly defined despite a number of publications in which categorizations are suggested by searchers like M. Montés de Oca, C. Yohant and A. Markowitz (1990), who differentiate:

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

- *Sexual abuse without body contact*: public or private exhibitionism, pornographic pictures, verbal incitation towards sexual activity, and erotic talk on sex in front of children.
- Sexual abuse with body contact: erotic caresses or kisses, touching, calls for masturbation.
- Sexual abuse with penetration: rape attempt or rape with anal, vaginal or oral penetration (the French penal code completes this definition of rape "with the use of any object".

Legal, clinical or medical definitions overlap or oppose themselves. Inside a research field, contradictions add to the complexity of psychopathological and therapeutic aims such as "*compliant hymen*" regarding medico-legal expertise or the statute of limitation after 10 years in the case of rape. Legal expertise is paradigmatic of these difficulties.

In this relatively complex context and based on our clinical victimology consultations in the ward of clinical forensic (CHU Montpellier, France – University Hospital Center), we have taken an interest (over a period of three years; 1998-2000) in two major questions on ill-treatments inflicted to underage:

First question: by differentiating sexual and non-sexual ill-treatments, what is the proportion of girls or boys who are affected?

Second question: *amongst observed disorders, do vulnerabilities, specific to under 15, exist that could lead to the hypothesis of a post-sexual-abuse syndrome?* 

## **SYNTHETIC REVIEW OF THE QUESTION (CLINICAL ASPECTS)**

The main epidemiological problem comes from the fact that some works are carried out based on retrospective sociological data gathered from adults, surveys from significant adult samples or studies limited to notional territory. These simple references demonstrate a disparity of assessment and methodological referential. These referential have no epistemological consensus and lead to bias in analysis.

We limit out research to a clinical approach of the phenomenon and the main publications on the subject.

The decade 70 – 80 can be called phenomenological to the extent that epidemiological data are statistically not very numerous. They are essentially, testimonies from female adults victim of incest during their childhood. The interest resides in the fact that research community has becoming aware of the need for granting this phenomenon with fundamental and applied research it deserves. Y.H.L. Haesevoets in 1997 and in 2004 emphasized that an important step has been taken in 1976 in K. Meiselman's study. This study, carried out between 1973 and 1976, on a sample of 58 adults abused in their childhood, observed chronic traumatic pathologies i.e. persisting 15 years after the trauma occurred. From this observation, analyses diverge. This divergence is partly due to the absence of comparative studies using either a control group or a clinical group of reference, the non-existence of standardized instruments in order to diagnose, the size of samples or the scarce anamnesistic studies prior

to sexual trauma. This group of methodological approximations generates impassable bias for a rigorous research.

From the years 1980, researchers attempted to harmonize their investigation protocols even if the undertaken notions appear more phenomenological than clinical or psychopathological. H. Van Gijseghem demonstrates the irreversibility of psychological wounds; S. Groi, of symptoms of anxiety and exacerbated fright (insomnia, nightmares, sleeping disorders, somatic disorders,...). Other research report after-effects such as food disorders, depression or phobia,...

In the years 1990, J. Kieser et al. show, with a sample of 10 children between 2 and 6, the development of symptoms in connection with Post Traumatic Stress Disorder criteria with reference to the DSM III R. Other studies balance these works [26]. We can retain:

- Loss of self-confidence (Mac Leer, 1998).
- Loss of self-esteem (shame and culpability) (Herman, 1981; S. Groi, 1982).
- Symptoms of depression, sometimes severe.
- Significant correlations between sexual brutalities and delinquent or criminal behaviors (David and Earls, 1987).
- Intellectual development disorders and sociability disorders (Gomes-Swartz; Horowitz, 1985).
- Depressions and suicidal thoughts (Wozencraft, 1991).
- Alteration in identity structure (Hunter; Childers, 1988).
- Typically psychiatric severe disorders (Cole; Putman, 1992).

In 1997, Y.H. L. Haesevoets suggests a grid of the principle psychopathological disorders connected with incest. This grid analyses somatic, physical and physiological, psychosomatic, cognitive, behavioral and relational, familial and social factors. The author yet specifies that the consequences of an abusive situation appear to depend on uncontrollable variables such as precocity of detection, the beginning of therapeutical treatment, designation of the sexual trauma (touching, rape, incest, sexual brutality, etc...) as well as the individual experience of the trauma by the underage.

The unpredictability of symptoms is not synonymous of their inexistence. From an empirical point of view, we observe symptomatic "*silences*" becoming loud, months – if not decades – later (as in peri- menopausic psychological changes, for example), symptomatic "*over-determinations*" revealed during subsequent minor trauma, for example. Some clinical constants appear and most theoretical models take them into account:

- Some symptoms seem inseparable from pre-existing symptoms. Post-trauma disorders intervene as an accentuation of the previous psychopathology.
- The clinical picture tends to put into relation some symptoms, with the real coercions of traumatic break-in, particularly coercions of humiliation as in collective rapes or torture.
- The recurrence of sexual violence during childhood seems to correlate with the severity of ulterior disorders, from a psychiatric point of view as much as a somatic decompensation one.

• With the exception of some sexual conduct disorders in the underage, there does not seem to be any evidence of causality between sexual abuse and symptom production.

Methodological "*weaknesses*" observed in different studies temporarily lead, not to postulate for a post sexual abuse answering to precise diagnostic assessment, specific diagnosis and sufficient anamnesistic backgrounds (research involving control groups).

Although, some of these bias appear in our works, considering the complexity of interactions whether psychological, pathological or inherent to child or teenage development, they seem to aim towards an opposite conclusion, which is the existence of such pathology.

## STUDY 1

## Method

The repartition according to sex of the victim and nature of sexual/non sexual illtreatment has been studied in a first longitudinal research over three years (1999, 2000 and 2001). The data collection is from consultations in victimology in forensic ward (University Hospital Centre – Montpellier – France) during the first appointment.

Non-sexual ill-treatment: assaults, injuries, scratches, tortures, moral harassment, brutalities, punishments, deprivations.

Sexual ill-treatments: rape, incest, exhibitionism, voyeurism, sexual touching, sexual brutality, sexual tortures.

#### Results

Type of aggression	% girls	% boys	
Non-sexual ill-treatments	17.6	29.4	
Sexual ill-treatments	44.1	14.7	
All ill-treatments overcome	61.7	38.2	

#### Table 1. Study 1999- number of consultations: 61

#### Table 2. Study 2000- number of consultations: 111

Type of aggression	% girls	% boys	
Non-sexual ill-treatments	3.6	25.2	
Sexual ill-treatments	52.2	18.9	
All ill-treatments overcome	55.8	44.1	

Type of aggression	% girls	% boys	
Non-sexual ill-treatments	6.5	23.7	
Sexual ill-treatments	64.75	5.3	
All ill-treatments overcome	71	29	

#### Table 3. Study 2001- number of consultations: 160

#### Table 4. Average of the studies 1999, 2000,2001- number of consultations: 332

Type of aggression	% girls	% boys	
Non-sexual ill-treatments	9.15	26.1	
Sexual ill-treatments	53.6	12.9	
All ill-treatments overcome	62.8	37.1	

## STUDY 2

#### Method

The data collection is for the year 2000 from 160 consultations. The semiology is identical to the one observed during the first interview. The interview's technique is semidirective and aims towards a diagnosis. We do not use the repartition by age slice.

Data analysis was done in two parts. The first part was a generic draft of types of syndromes, of disorders encountered amongst underage, in order to, in a second part, specifies items susceptible to help us towards the creation of a symptomatic analysis grid. The aim of our research was the creation of a clinical assessment tool for sexually abused underage or underage suspected of having been sexually abused. This study was part of a HPCR (Hospital Project in Clinical Research) in relation to the French legal context – June 1998 law on audio-visual recording of interviewed of underage who are suspected to have been sexually abused, and the effects of "over victimization".

## Results

# Table 5. Frequency of the clinical signs expressed in % - year 2001- Number of people:160- Clinical picture after the first consultation- Sex and age overcome.

Oral sphere disorders	26.5	
Sphincterian functioning disorders	18.9	
Disorders of the cognitive sphere	94.9	
Behavior and conduct disorders	60.7	
Sleeping and falling asleep disorders	22.7	
Genital sphere disorders	20.25	
Social link disorders	20.25	

Disorder of the oral sphere	Percentage	
Anorexia	5.06	
Bulimia	11.3	
Weight loss	6.3	
Weight gain	3.7	

## Table 6. Repartition per item. Oral sphere disorders.

## Table 7. Repartition per item. Sphincterian functioning disorders

Sphincterian functioning disorders	Percentage		
Enuresis	2.53		
Encopresis	2.53		
Diarrhea	0.01		
Constipation	5.06		
Abdominal pains, stomach aches	8.96		

## Table 8. Repartition per item. Disorders of the cognitive sphere

Disorders of the cognitive sphere	Percentage	
Sudden school failure	7.5	
Scholastic drop	17.7	
Scholastic non-implication	20.25	
Mnemic disorders	10.12	
Learning disorders	8.86	
Concentration disorders	7.59	
Language and communication disorders	11.3	
Graphics and writing disorders	11.3	

## Table 9. Repartition per item. Behavior and conduct disorders

Behavior and conduct disorders	Percentage		
Self-aggressive conducts	5.06		
Aggressive behaviors	7.55		
Passivity-depression- submission	13.8		
Repetitive activities	5.6		
Psycho-motor activities agitation	6.86		
Psychosocial instability, within the group	8.75		
Behavioral impulsivity	11.3		

## Table 10. Repartition per item. Sleeping and falling asleep disorders

Sleeping and falling asleep disorders	Percentage		
Nightmares	7.6		
Falling asleep disorders	11.5		
weariness	3.8		

Genital sphere disorders	Percentage	
Masturbation – touching	5.06	
Exaggerated modesty	14.65	
Frigidity	9.82	
Anorgasmia	11.5	
Definitive interruption of sexual intercourse	20.56	
Gynecological problems	8.87	
Compulsive washing disorders	19.86	

#### Table 11. Repartition per item. Genital sphere disorders.

#### Table 12. Repartition per item. Social link disorders.

Social link disorders	Percentage	
Aggressiveness	6.32	
Isolation-withdraw	12.08	
Inhibition	7.34	

#### Results' Synthesis

General data show a strong incidence of sexual trauma on the cognitive sphere for the quasi majority (94.9%) of the population examined. The study of individual items tends to confirm that troubles are not of an instrumental nature. It is the relation with didactic content and implication towards learning, which are being disturbed. The post-traumatic effect seems to interfere with concentration, representations and memorization. Scholastic drop (17.7%) and non-implication (20.25%) underline this "*leech off*" in the attitude towards learning like tools for appropriation of knowledge (language in particular).

Symptoms in relation to oral sphere disorders are feeding disorders of a bulimic nature (11.3%). Bulimic behavior emphasizes (based on the number of people we observed) the alterations of the body image especially the upheaval of body topography (inside, outside, bottom, top front, back...).

Sphychterian function disorders are not directly linked with mastering of sphincter -anal or urethral, for example- intestinal transit causing colonopathy or abdominal pain (8.96%). Sleeping disorders occur mainly in the form of difficulties to go to sleep (11.5%).

Behavior and conduct disorders (BCD) appear contradictory, when only reading on quantitative level. The item "passivity-depression- submission" (13.8%) is apparently in opposition with the item "behavioral impulsiveness" (11.3%). It seems that this opposition shows a double aspect of the BCD meaning predominance of passivity from which facial expressions–often anxious- of behavioral impulsiveness emerge. This aspect can be corroborated by social links disorders in which the item "isolation-withdrawal" is at 12.08%. The behavioral propensity would be a propensity towards withdrawal and isolation more than a post-traumatic pathology loud manifestation.

Finally, disorders of the genital sphere can appear surprising in particular the item related to the definitive interruption of sexual intercourse, considering our study is on under 15. This interruption regards the eldest underage we observed (20.56%). The item "compulsive

disorders" (19.86%) qualifies compulsive washing behaviors (of an obsessional type), frequent change of underwear with cleaning control, etc.

#### Comments and Analyses

The qualitative analyses of the above quantitative data permit to better assess psychological and psychical mechanisms striving, even though our preliminary study calls for other complementary and deepened researches. It still corroborates the 20 indicators of sexual exploitation defined by S. Groi as early as 1986 (cited by Y.H. Haesevoets).

On the level of cognitive process analysis, it seems that post-traumatic effects do not invalidate the efficiency of the cognition mechanisms (language, memory, perception, etc...). We, essentially, observe phenomenon of "interference" and inhibition more of less important, which, indeed, have some serious and durable consequences. This seriousness of sexual trauma tends to be correlated with the sensitive phases in the development (for example, the pubertal phase and its identity changes) or their duration (recurrent incest over many years).

Although, longitudinal studies are essential, when studying sexually abused underage, in order to avoid anamnesistic reconstructions or the after effect, it appears that phenomenon of anxiety consecutive to trauma pervert (literally) learning processes. H. Van Gijseghem (1975, 1985) underlines intellectual development deficiencies particularly of symbolic and Piagetian thought. Scholastic symptomatology shows a disharmony of learning processes generated by *"fright, inhibition of thought and relational passivity"*. For school children, sexual trauma disturbs learning process even if that learning can help towards a better control of the sustained event as well as helping towards it verbalization.

A research consensus agrees that the earlier the abuse occurs in life, the more risks that wounds become irreversible at all levels; "*particularly on identity level*" adds H. Van Gijseghem. At more advanced ages, observed phenomenon on cognition level are identical to those observed amongst younger children. On the other hand, behavioral and conduct disorders (depression, dysthymia, running away, suicidal ideation, etc...) generate, for the essential, these cognitive disturbances.

Pathogenic components that touch the oral sphere, such as sphyncterian or functional (stomach ache) malfunctioning appear as somatic symptoms for want of sufficient psychical elaboration (alteration of representational mechanisms, in general). Bulimic behaviors, for example, emphasize alterations of processes of identity construction, in particular, the progressive construction of the body image. Bulimia appears during clinical examination like a reactional formation when facing the breaking of primary psychic sheaths. It acts like "*a lipidic protection*" for want of "*protective shield*" efficiency.

The "*psychosomatic*" aspect of psychical investments tends to accreditate the hypotheses of impossibilities in the representation of sexual trauma, such as pain affects, mentioned by M. Bertrand (cited by L. Crocq and P. Bessoles). This aspect of a symptomatology of the psychic body and the skin-Ego (D. Anzieu, 1987) attests of the victimarius destruction in the underage. All recent works confirm this double aspect of psychosomatic symptoms and disturbance in the image of the body. These same research corroborate others, more ancients, such as the model of D. Finkelhor and D. Browne in 1985, published in the American Journal of Orthopsychiatry.

Disturbances in falling asleep reveal the failures of the work of thought and dream. Abused underage appear to be fighting against the decline of the level of vigilante consciousness to which, falling asleep, lead, being too scared to be invaded with anxieties and traumatic reviviscence. This attitude, highly phobic, that can take the clinical shape of diurnal and nocturnal terrors is typical of post-traumatic syndromes amongst children as well as adults. We wish to underline this phobic aspect, which is at the source, it seems, of dysmorphophobia which particularly pathogenic amongst female teenagers.

The visible contradiction in our study, due to passivity ...... Withdrawal, isolation or inhibitions coexist through aggressive behaviors –sometimes violent-. Hyper-vigilance and hypersensitivity translate psycho-affective insecurity of the underage. It is expressed by constant submission and passivity. Violence appears like the manifest expression of anxious faces, facing paroxystic anxieties of an anaclitic type. This behavioral "*dyschrony*" attests of the underage internal disturbances, of which, acting out behaviors –auto or hetero aggressive – are the mirror of post-traumatic suffering.

Genital sphere disorders are, for the essential, psychosomatic (dermatosis, cystitis, prurit, inflammatory process, vaginal herpes attacks...). These phenomena, well known by medical corps, can be treated by using psychotherapeutic processes and without any medicine apart from comfort. On the other hand, it is true, particularly for female teenagers that, rituals of washing rituals of the uro-genital zone, particularly their frequency and intensity, have a tendency to modify the self-immune balance of the pelvic belt. Sexuality and sexual conducts disorders, in its different components according to the age of the underage, are constant. They lead to sexual prematureness, compulsive conducts of masturbation or important inhibitions (notably nudity) often observed by pediatricians or doctors (in the context of school medical).

With the exception of very few particularly suggestive mimes of coitus using a doll (which does not necessarily express the reality of the sexual trauma), a majority of underage patients did not show any seductive or provocative attitude, as it is often stated in literature. On the contrary, underage children show important inhibitions which require some form of talk prior to any somatic examination.

In an attempt to summarize clinical and psychopathological forms of vulnerabilities amongst sexually abused underage children, we can divide them into four main domains:

- A symptomatology of school learning. It is about learning process which become disturbed, inhibited and leeched off by post-traumatic effects. It is, therefore, the relationship with knowledge and knowledge which is at cause and not the cognitive efficiency.
- A symptomatology on behavior and conduct disorders centered on a dominating inhibition of the psychical link and social link. This domination sometimes turns into anxious acmes in the form of auto or hetero-aggressive raptus.
- A more or less invalidating phobic symptomatology, with obsessional behaviors mainly centered on rituals of washing.
- A somatoform symptomatology which indicates the alterations of representational processes and the effects of psychical staggering due to sexual trauma.

## **DISCUSSION**

The evaluation of post-traumatic vulnerabilities amongst sexually abused underage children emphasizes, independently from clinical characteristics, the main – and durable – traumatogenic impact on the process of identity construction.

Ancient Freudian's works (1920) and Férenczian's (1927-1933) have broadly described post-traumatic psychopathologies for the same reasons as those more recent from L. Crocq, F. Lebigot, L. Daligand, etc... Our research confirms the presence of a psychopathological entity of trauma neurosis independent from psychopathological backgrounds, amongst underage children (as also observed amongst adults).

This traumatic neurosis adds to pre-existing pathologies but appears specific, even if the interactions with post-traumatic pathologies complexify the differential analyses of the clinical picture.

Our arguments rest on the above study as well as the therapeutic treatment of the underage children. We would like to specify that the hospital conditions for admitting patients (clinical victimology ward CHU Montpellier), were purposely limited to post-immediate sharp states. Underage children were addressed to partners of the network we collaborated with (specialized institution, CMPP -Medico Psycho Pedagogic Center-, child psychiatry...).

- A direct incidence exists between the event sustained and the emergence of a specific psychopathology described above (even if this incidence, still today, calls for more in depth research in order to specify their characteristics).
- Post-traumatic psychopathology observed cannot be limited to a reading of comorbidity or accentuation factors in previous disorders. Even if it can be the case, for some of the examined underage children (like boarders in an institution for mentally deficient children), it is wrong to assert (from an empirical point of view to this date) that sexual trauma only produces a quantitative and qualitative surcharge of previous disorders.
- The specificity of a post-traumatic syndrome is argued in the differential diagnosis, i.e. significant correlations between sexual trauma characteristics and generated disorders. Thus, the "unfavorable" diagnosis has been confirmed for those, victim of sexual aggression with penetration, humiliation, repetition and torture.
- At an adult age, traumatic pathologies are even more persistent when the person has not been able to undertake any treatment, independently from therapeutical fields of reference.
- Sexual trauma generates real "*cataclysms*" in the underage child's psychic organization, not only in his psychical system of economy but above all in his structural organization. This aspect seems more pathogenic at time of phases of psychical organizations such as Oedipus, puberty or adolescence phase.
- The alteration of identity process is apparent in the clinical picture's somatoform expressions. This expression does not just reveal the difficulties in psychically elaborating the trauma but also the breaking of primary psychical sheaths (*"secure base"*).
- Finally, "we have to take into account" the fact that most sexual abuses are committed by a member of family (or a relative) in 85 to 90% of cases (according to

our panel). These epidemiological data confirm the idea of pervertion of affective, familial, social,... bond, even if the degree of relativity between the underage child and his abuser remains to be specified.

## CONCLUSION

We wish to emphasize the phobic dimension of post-traumatic states amongst the underage. On the semiological level, phobic symptoms observed in our study are not only due to classical description of child and teenager psychopathology. The intensity of anxiety does not only answer to phobic situations, it is often close to states of panic, notably in the difficulties in going to sleep. Likewise, dysmorphobia, observed in our research, do not appear to be phobia in the "strict" meaning, but disorders in the construction of body image, independently from corporal mutations inherent to the development of the young subject.

Anxiety is not only related to a potential traumatic situation as in traumatic reviviscences (notion of "anxiety warning"). It generates a constant and chronic feeling of anxiety, as well as an extremely invalidating subjacent depression of a neglected or anaclitic type, for most observed underage children.

The phobic aspect appear to evocate some border-line psychic functioning, as these phobia are atypical because of the context in which they appear, their content and the fact that they have no impact on anxiety. On the contrary, it soaks-in the entire underage mental function, and compromises relationships with others. The whole of the subject is threatened as anxiety is so massive and intrusive.

If, as stated by D. Marcelli, on the subject of phobic psychopathologies amongst the under aged and teenagers, "*it is a phobic neurosis from a clinical point of view*, [...] there is a reason to look for anxiety of a paranoid type". Empirically, we agree with the term of paranoid anxiety, for some of the patients, during the sexual trauma post-immediate phase. It is characterized by the following aspects:

- The formulation of the trauma sustained, sometimes, appears incoherent, badly systemized, if not incomprehensible, independently from the syntactic or semantic level of the under-aged. This formulation appears terrifying and incommunicable, sometimes completed with invasions of interfering coenesthesia (twitches, shivers, spasms, interfering actions, mimes,...).
- The under-aged sometimes feel, spied on, observed or threatened and uses strategies such as para-verbal or psychomotor allusions (scribbles in dolls crotch done by young children, for example) or whispers in the doctor or psychologist ear "*but, you know*...".
- The threat of psychic dismantling can often be spotted in the form of persecutions, notably in graphic productions realized during the clinical interview. The people drawn are often being attacked, aggressed; sometimes, a whole has been made in the paper with the felted-pen. The intensity of anxiety is such that in some cases, dolls are literally destroyed, as if the esthetic experience was close to that feeling of having one's body cut into pieces.

• The way the body is apprehended, amongst the most dramatic cases (collective rapes perpetrated on female teenagers), is close to real processes of depersonalization with, a feeling of body dispossession, ideas of body transformation and for two clinical cases conducts of self-mutilations and autolysis.

These productive moments can be transitory, but we cannot, due to the absence of longitudinal follow-up, specify the evolution of chronicity. On the other side, retrospective studies on adults tend to confirm long lasting psychopathological after-effects as well as pathogenic reviviscences decades after the infantile trauma. It seems, in an empirical at this point in time, that specific phases of vulnerability, which are not related to life psychical hazard, exist. These phases are the one when identity changes are directly "requested" like peri-partum or pre-menopause phases.

On the other side, we feel that it is important to emphasize, in the same way we have done it elsewhere, on the resilient capacities, sexually abused children and teenagers are capable of. Psychical and psychological "fragilities" underlined hereby appear as, as much plasticity in psychical changes and elaborations of the sexual trauma. Recent research by M. Lemay (1998, 1999) or B. Cyrulnick (199, 2001) show these under-aged can elaborate particularly terrible and destroying sexual traumatisms. Nevertheless, it would be a mistake to think that these psychical elaborations are definite.

## REFERENCES

Anzieu D. Les enveloppes psychiques. Paris : Ed. Dunod ; 1987

- Baccino E, Bessoles P. Le traumatisme sexuel et ses devenirs. Lecques : Ed. Le Champ Social ; 2001.
- Baccino E, Bessoles P. *Traumatisme et résilience : lien psychique, lien social*. Lecques : Ed. Le Champ Social, 2003.
- Bessoles P. Torture et temporalité : contribution à une sémiologie de psychose post traumatique. In : Les cahiers de psychologie clinique. Bruxelles : Ed. De Boëck University ; 2004.
- Bessoles P. 2000, Le *meurtre au féminin, clinique du viol*. Saint Maximin : Théétète Editions, 1997.

Bessoles P. La psychose post-traumatique. In : L'encéphale. Paris : Elsevier Editions ; 2005.

- Brusset B. *Etats anxieux et névroses* : critique des concepts. In : Marcelli D, Braconnier A, editors. Adolescence et psychopathologie, 4. Paris : Ed. Masson ; 1999. p. 203-38.
- Coutenceau R. Vivre après l'inceste. Paris : Ed. Desclée de Brouver ; 2004.
- Crocq L. *Névrose traumatique et élaboration résiliente*. In : *Victime-Agresseur*, Tome 4. Lecques : Ed. Le Champ Social ; 2003. p. 27-41.
- Crocq L. Bessoles P. *Récidive, réitération, répétition : lien d'emprise et loi des séries.* Lecques : Ed. Le Champ Social ; 2004.
- Cyrulnik B. Un merveilleux malheur. Paris : Ed. Odile Jacob ; 1999.
- Cyrulnik B. Les vilains petits canards. Paris : Ed. Odile Jacob ; 2001.
- Dalignand L, Gonin D. Violence et victimes. Lyon : Ed. Méditions ; 2004.
- Damiani C. Les Victimes : violences publiques et crimes privés. Paris : Ed. Bayard ; 1997.

- Defrancis V. *Protecting the child victim of sex crimes committed by adults*. Denver, CO : American Human Association, Children's division ; 1976.
- Fattah EA. Le rôle de la victime dans le passage à l'acte : vers une approche dynamique du comportement délictuel. Rev Int Criminol Police Tech 1973 ; XXVI(2) : 173-88.
- Ferenczi S. 1927-1933, Confusion de la langue entre l'adulte et l'enfant, in : Œuvres complètes, Psychanalyse IV, Paris, Payot, 1968.
- Finkelhor D. Sexuality victimized children. New-York : Free Press ; 1979.
- Finkelhor D. Browne A. *The traumatic impact of child sexual abuse : a conceptualization*. Am J Orthopsychiatry 1989 ; 55 :530-41.
- Gabel M. Les enfants victimes d'abus sexuels. Paris : PUF ; 1994.
- Gabel M. Le traumatisme de l'inceste. Psychiatrie de l'enfant. Paris : PUF ; 2000.
- Gutton P. Adolescens. Paris : PUF ; 1996.
- Haesevoets YH. L'enfant victime d'inceste. Bruxelles : Ed. De Boëck Université ; 1997.
- Haesevoets YH. L'enfant en question. De la parole à l'épreuve du doute dans les allégations d'abus sexuels. Bruxelles : Ed. De Boëck Université ; 2003.
- Lebigot F. Bessoles P. Situation d'urgence, situation de crise : clinique du psychotraumatisme immédiat. Lecques : Ed. Le Champ Social ; 2004.
- Lemay M. Résilience devant la violence. J Psychologues 1999 ; 164 :36-47.
- Lighezzolo J, de Tyché C. *La résilience : se (re)construire après le traumatisme*. Paris : Ed. In Press ; 2004.
- Marcelli D, Braconnier A. Adolescence et psychopathologie. Paris : Masson ; 1995.
- Marcelli D. Enfance et psychopathologie. Paris : Ed. Masson, 1996.
- Mazet P, Stoléru S. *Psychopathologie du nourrisson et du jeune enfant*. Paris : Masson ; 1993.
- Montes de Oca M, Yohant C, Markowitz A. *Revue critique des travaux épidémiologiques* anglo saxons récents sur l'abus sexuel à l'égard des enfants. In : Les abus sexuels à l'égard des enfants, CTNERHI. Paris : PUF ; 1990. p. 7-70.
- Van Gijseghem H. L'enfant mis à nu. L'allégation d'abus sexuel : la recherche de vérité. Québec : Méridien ; 1992.
- Van Gijseghem H. *Autre regard sur les conséquences de l'inceste père-fille*. Revue canadienne de psycho-éducation ; 1985 ;14(2) :138-45.
- Van Gijseghem H. Facteurs interférents avec la qualité du témoignage de l'enfant au tribunal dans les causes d'abus sexuels. Québec : Revue canadienne de Psycho-éducation 19.1.11.22 ;1975.

Chapter 21

## PREDICTORS OF MIDDLE SCHOOL YOUTH EDUCATIONAL ASPIRATIONS: HEALTH RISK ATTITUDES, PARENTAL INTERACTIONS, AND PARENTAL DISAPPROVAL OF RISK<sup>\*</sup>

Regina P Lederman<sup> $\dagger$ ,1</sup>, Wenyaw Chan<sup>2</sup> and Cynthia Roberts-Gray<sup>3</sup>

<sup>1</sup>University of Texas Medical Branch, University of Texas Medical Branch, Galveston, Texas 77555-1029, USA <sup>2</sup>University of Texas School of Public Health, Galveston, Texas 77555-1029, USA <sup>3</sup>Resource Network, Galveston, TX 77550, USA

## ABSTRACT

School-wide surveys in five middle schools were used to measure educational aspirations, attitudes toward sexual health risk behaviors and drug use, and perceptions of parental interactions and disapproval of risk behavior at baseline and one year later. Participants were male and female students of Black (n = 222), Hispanic (n = 317), White (n = 216), and Asian or other heritage (n = 85), ages 11 to 14. Analyses were performed for three factors with Cronbach's alpha coefficients  $\geq 0.65$  (youth's attitudes, discourse with parents, and parents' disapproval of risk behavior), and three single items inquiring about use of alcohol, use of marijuana, and sexual behavior. Generalized Linear Model (GLM) with logit link was used to evaluate the contribution of these measures at baseline as predictors of educational aspirations at the one-year follow-up. Results showed race/heritage (p < .001), attitudes toward health risk behavior (p < .01), extent to which youth talked with parents' disapproval of risk behavior (p < .05) each made significant contributions in predicting educational aspirations. Gender did not contribute to the

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Contact Author: Regina P Lederman, PhD, University of Texas Medical Branch, 301 University Blvd. Galveston, TX 77555-1029. Phone: 713 666 0172. 409 772 6570. email: rlederma@utmb.edu

prediction of educational aspiration nor did self-report of actual risk behavior. These results indicate that youth interactions with parents regarding health risk behaviors is worthy of further exploration to develop interventions to reduce adolescent health risks and increase educational aspirations.

## INTRODUCTION

Parent-child relations occupy important positions in pathways linking education and health. Parental encouragement for educational attainment is, for example, closely linked to youth educational aspirations (Looker and Thiessen, 2004; Goyette and Xie, 1999). Educational aspirations of youth (Trusty and Harris, 1999) and parental involvement in their children's education (Trusty, 1999) are linked to educational attainment. Lack of educational attainment and low educational aspirations in turn are associated with a variety of health and health behavior problems, including substance abuse (Caetano, 2002) and rates of teenage birth (Pamuk, Makuc, Heck, Reuben, and Lochner, 1998). The association of poor education and poor health is a consistent finding from domestic and international research in economics, epidemiology, and sociology (Blane, 2003; Morowsky and Ross, 2003).

In some formulations of the pathways linking education and health, education is identified as a health-protecting factor (see Hannum and Buchman, 2003). Because of the abundant empirical evidence that better educated people lead longer, healthier lives, those who espouse the health protection formulation recommend investment in programs and practices that promote educational attainment as a means for improving health and reducing health inequalities. At a local program level, such an initiative might include parenting education and supportive services to assist families in encouraging their children to aspire to and achieve higher educational attainments, anticipating that positive health outcomes will accrue.

Other analyses demonstrate that good health in adolescence predicts both higher educational attainment and better adult health (Chandola, Clarke, Blane and Morris, 2003). Recommendations for action from this perspective focus on interventions to help families encourage healthy behaviors in childhood and adolescence with the expectation that better educational and adult health outcomes will follow.

In this chapter we examine a secondary analysis of data from a prevention research project exploring relationships among the educational aspiration of middle school youth, their health risk attitudes and behaviors, and interactions with their parents regarding risk behavior. The data, originally obtained to test the effectiveness of an innovative parent-child program for the prevention of sexual health risk behaviors (see Lederman, Chan, and Roberts-Gray, 2004; Lederman and Mian, 2003), were collected in school-wide surveys at five middle schools during a baseline period and again the next school year. The purpose of this secondary analysis of the data is to explore the possibility that, in addition to putting the health of youth at risk, negative health attitudes and behaviors distract youth from the pursuit of educational goals and thereby constrain educational aspirations, while appropriate parental involvement in discouraging negative health behaviors in youth can contribute to higher educational aspirations. We hypothesized that middle school youth involvement in and attitudes toward health risk behaviors, and their perceptions of interactions with their

parents/guardians regarding health risk behaviors would be predictive of educational aspiration measured during the subsequent school year.

## **Research Design and Methods**

This study analyzed data collected in school-wide surveys over two successive school years from students in five middle schools in two school districts in coastal Texas. The written survey questionnaire provided self-report measures of: (1) sexual involvement and the use of alcohol and other drugs; (2) attitudes toward sexual health risk behavior; (3) the frequency and breadth of discourse with parents about sexual health topics; (4) perceptions of parents' disapproval of the youth's involvement in sexual and other health risk behaviors; and (5) educational aspiration. The protocol was approved by the Institutional Review Board of the University of Texas Medical Branch at Galveston. Active assent of students and consent of parents was obtained and documented.

#### Participant Recruitment

Participants were recruited in four steps. First, we sent invitations through the mail and through presentations at conferences to all urban school districts located in and near Galveston, Texas. Leaders in two school districts indicated willingness to have their schools participate in the prevention intervention research project. The next step was meeting with principals and counselors at middle schools in the consenting districts. Two out of the three middle schools in one district and three of the four middle schools in the second district agreed to participate. The third step was partnership with the consenting schools to invite families to participate. A cover letter, consent form, and informational flyer were sent to parents via the students and through the mail informing them about the program and providing details regarding the research protocol. The school principals at each of the participating schools endorsed the program and signed the letters to parents, which were written in both English and Spanish.

The participating schools represented diversity in school size, ethnic heritage of the student body, and performance characteristics. Two of the schools had approximately 500 students, two had between 700 and 1,000 students, and one had more than 1,200 students. Across the five schools, the percent of students of different heritage ranged from 0.0 to 3.4 Asian, 7.0 to 54.3 African American, 32.8 to 85.8 Hispanic, and 5.2 to 52.1 White. Three of the schools had more than two-thirds of the student body identified as economically disadvantaged. The percent of students with special circumstances ranged from 3.7 to 23.6 with limited English proficiency, 4.5 to 13.6 participating in gifted and talented programs, and 2.2 to 4.3 percent with disciplinary placement. Total enrollment across the five schools was 3,881. Informed consent and actual participation in the survey was obtained for 848 youth.

### Demographic Characteristics of the Participating Youth

More than half of the students were in sixth grade at the time they completed the baseline survey. There were approximately equal numbers of females and males. The majority was youth of color with nearly 40% indicating Hispanic/Latino/Mexican ethnicity. These data are displayed in Table 1.

Youth characteristics	Number of Youth	Percent
GENDER		
Male	383	45%
Female	465	55%
RACE/ETHNICITY		
Hispanic/Latino/Mexican	317	38%
Black/African American	222	26%
White – not Hispanic	216	26%
Asian or Mixed or Other	85	10%
GRADE LEVEL		
Sixth	480	60%
Seventh	249	31%
Eighth	76	9%

## **Table 1. Demographic Characteristics of Survey Participants**

#### Survey Administration

Students completed the survey questionnaire during a special assembly at the school. The survey items originally were selected to be relevant in evaluating outcomes of prevention education interventions to reduce sexual health risk behaviors and prevent teen pregnancy and the spread of HIV and other sexually transmitted infections (STI). Adapted from the National Youth Survey (Elliot, Huizinga, and Ageton, 1985) and a middle school survey developed by ETR Associates (Kirby et al, 1997), the 9-page, 94-item survey questionnaire was divided into sections asking: "What do your parents think?," "What do you think?," "What do you do?," and "Do you talk with your parents?" Students typically completed the survey in 30 to 45 minutes. Unique identifying codes were marked on the survey forms so that records could be matched from year to year without disclosing individual identifying information.

#### Measures

Survey items addressing a common content domain (e.g., youth attitudes toward risk behaviors) and having similar response options (e.g., strongly agree to strongly disagree) were combined to produce domain scales which then were tested for internal consistency. Those obtaining a Cronbach alpha of 0.65 or greater were retained as scales. The scales and survey items used in the current analyses are described as follows.

• *Educational Aspiration* was measured with a single item asking, "How far do you think you will go in school?" Response options were assigned ordinal codes of 1 = "Won't graduate from high school," 2 = "Will finish high school," 3 = "Will go to trade, vocational, or business school after high school," 4 = "Will attend college."

Youth involvement in and attitudes toward risk behaviors were measured with two sets of survey items, one that counted the number of self-reported risk behaviors and the other a scale measuring perceived acceptability of "someone my age" having sex.

- *Health Risk Behaviors* was measured with three single items: (1) "In the last 30 days, did you drink alcohol such as beer, wine, wine coolers or hard liquor?," (2) "In the last 30 days, did you use marijuana?," and (3) "In the last 6 months, did you try to get someone to have sex with you?" Response options were coded 1 = No and 0 = Yes.
- Attitude Toward Sexual Health Risk Behavior was measured with a 14-item domain scale asking about conditions under which it is acceptable to have sex. Sample items are: "I believe it's OK for someone my age to have sex with someone they like, but don't know well," "I believe people my age should always use a condom if they have sex," and "If my boyfriend or girlfriend wanted to have sex and I didn't, it would be OK to say 'no'." Response alternatives were "Strongly agree," "Agree," "Disagree," and "Strongly Disagree," coded 1 through 4 with more desirable responses receiving a higher numerical score. Cronbach's alpha for the baseline sample was 0.76.

*Interactions with parents* regarding risk behaviors was assessed with two domain scales, one measuring the frequency and breadth of youth discourse with parents about sexual health topics and the other measuring youths' perceptions of whether or not parents or guardians would disapprove of youth engaging in risk behaviors.

- *Discourse with Parents about Sexual Risk Behavior* measured how often in the last three months youth had talked with their parents about five topical items: menstruation, the risk of getting pregnancy or getting someone pregnant, being a teenage parent, different kinds of birth control, and sexually transmitted diseases (STDs) or AIDS. Response options were coded 0 = "Never," 1 = "1-3 times;" and 2 = "More than 3 times." Cronbach's alpha for this domain scale at baseline was 0.73.
- *Parental Disapproval of Risk Behavior* was measured by asking if the parent or guardian would "approve" or "disapprove" if the youth engaged in six types of risk behavior: used alcohol; sniffed paint or glue or used marijuana; used hard drugs such as heroin, cocaine or crack; had sex (made love, went 'all the way'); got pregnant or got someone pregnant; and failed a grade in school. Responses were coded 1 and 0. The Cronbach alpha for this scale at baseline was 0.65.

A study reported elsewhere (Lederman, Chan, and Roberts-Gray, 2004) showed the domain scales selected for the predictive model tested in the current study were independent of one another. The obtained inter-domain correlations were low with Pearson correlation

coefficients of r = 0.10 for Discourse and youth Attitude, r = 0.04 for Discourse and perceived Parental Disapproval, and r = 0.29 for youth Attitude and perceived Parental Disapproval.

#### Data Analyses

Generalized Linear Model with logit link was applied to evaluate the contribution of the baseline measures of interactions with parents and youth involvement in and attitudes toward risk behaviors as predictors of youths' educational aspirations measured in the subsequent school year. Additional variables included in the statistical models were students' age (11 and 12 versus 13 and 14), self-reported ethnic heritage, and gender. A limitation of this study is that all analyses assume the data are taken from random samples. Like many studies reported in the research literature, this study cannot verify this assumption. Generalization of the results should be cautious.

## RESULTS

*Educational Aspirations*. Virtually all (99%) of the middle school students that participated in the survey indicated they plan to finish high school. Ten percent indicated completion of high school is the highest level of education they expect to attain. The vast majority of the students (86%), however, indicated they plan to attend college.

## Table 2. Generalized Linear Model with PREDICTORS Measured at Baseline and the Criterion Variable EDUCATIONAL ASPIRATION Measured the Subsequent School Year

Predictors	Beta Weight	SE	Chi Square	<i>p</i> -value
Gender	0.1835	0.2376	0.60	NS
Age (11-12 versus 13-14)	0.4917	0.3779	1.69	NS
Race/Ethnicity				
Hispanic / Latino, including Mexican	-1.3867	0.4196	10.92	<.01
Black / African American	0.1924	0.4792	0.16	NS
White – not Hispanic	-0.0617	0.4722	0.02	NS
Asian or Mixed or Other	0			
Health Risk Behaviors				
Use alcohol	-0.1672	0.2894	0.33	NS
Use marijuana	0.1890	0.4317	0.19	NS
Try to get someone to have sex	-0.3618	0.4139	0.76	NS
Parental disapproval of risk behaviors	2.0630	1.0186	4.10	<.05
Discourse with parents about risk behaviors	-0.7926	0.3454	5.27	<.05
Attitudes about acceptability of having sex	2.0928	0.6970	9.02	<.05

Predictors of Educational Aspirations. Race/heritage (p < .001), youth attitudes toward sexual health risk behaviors (p < .01), discourse with parents about health risk behavior (p < .05), and perception of parental disapproval of youth involvement in risk behavior (p < .05) each made significant contributions in predicting educational aspirations. The full model is displayed in Table 2. Neither age nor gender made significant contributions in predicting

Lower educational aspiration was predicted by higher levels of discourse with parents about sexual health risk behaviors and with being Hispanic/Latino/Mexican. Higher educational aspiration was predicted by youth holding attitudes less accepting of sexual activity and perceiving that their parents/guardians disapprove of their being involved in risk behaviors.

## CONCLUSION

This analysis demonstrates that the attitudes of middle school youth toward health risk behaviors and their perceptions of interactions with their parents/guardians regarding health risk behavior are useful in predicting educational aspiration measured during the subsequent school year. Youth who are at baseline less accepting of involvement in sexual activity and who perceive that their parents disapprove of their involvement in risky behaviors are more likely than those with more tolerant attitudes and perceptions to continue into the next school year with high educational aspirations. The finding that perceived parental disapproval of health risk behavior predicts youths' educational aspiration is consistent with other research indicating that maternal connectedness may facilitate the development of a positive future time perspective by young adolescents (Aronowitz and Morrison-Breedy, 2004).

An unexpected finding in the current analyses was that frequency and breadth of youths' discourse with their parents about sexual health topics showed a negative relationship with youths' educational aspirations. Youth who reported more discourse with parents about sexual health topics were significantly more likely than their peers to report lower educational aspirations. This finding underscores conclusions presented in Healthy People 2010 (U S Department of Health and Human Services, 2000, Chapter 25) regarding the need to give further attention toward helping parents impart information about sexual health risks. The research literature suggests there is no simple, robust relationship between parent-adolescent communication about sexuality and subsequent adolescent sexual health behaviors (Miller, 1998; Dittus and Jaccard, 2000). A hypothesis for exploration in future research is the possibility that young adolescents talk with their parents about risk behavior only when their behavior has created a predicament that forces discussion between parent and child. Another possibility is that more discussion about risks and protection may have a normalizing effect on the youth's perceptions of risk behavior. Yet another possibility is that more discourse about risk behavior occurs in families where parental control experienced as excessive or coercive, and impels the youth toward acting out and/or toward a more negative future time perspective (Tiongson, 1997).

The current study indicates that the interactions of youth with parents regarding health risk behaviors is worthy of further exploration and the development of interventions to improve adolescent health, and educational and adult health outcomes. There is a clear need for additional research to identify ways to marshal, nurture, and work in concert with parental efforts to reduce adolescent health risks and encourage educational attainment.

## REFERENCES

- Aronowitz, T., and Morrison-Breedy, D. (2004). Comparison of the maternal role in resilience among impoverished and non-impoverished early adolescent African American girls. *Adolescent and Family Health*, 3(4), 155-163.
- Blane, D. (2003). Commentary: Explanations of the difference in mortality risk between different educational groups. *International Journal of Epidemiology*, 32, 355-356.
- Caetano, R. (2002). *Education, psychiatric diseases and substance use/abuse.* Presentation made at conference on Education and Health: Building a Research Agenda. National Institutes of Health. Available to read online: www.wws.princeton.edu/chw conferences/conf1002/session/Caetanofiles/frame.html
- Chandola, T., Clarke, P., Blane, D., and Morris, J. (2003). *Pathways between education and health: A causal modeling approach*. University College London (UCL). Published online: www.ucl.ac.uk/epidemiology/Chandola/healtheducation.htm
- Dittus, P., and Jaccard, J. (2000). The relationship of adolescent perceptions of maternal disapproval of sex and of the mother-adolescent relationship in sexual outcomes. *Journal of Adolescent Health*, 26, 268-278.
- Elliot, D., Huizinga, D., and Ageton, S. (1985). *Explaining delinquency and drug use*. Beverly Hills, CA: Sage.
- Goyette, K., and Xie, Y. (1999). Educational expectations of Asian-American youth: Determinants and ethnic differences. *Sociology of Education*, 71, 24-38.
- Hannum, E., and Buchman, C. (2003). The consequences of global educational expansion: Social science perspectives. Cambridge, MA: American Academy of Arts and Sciences. Available to read online: www.amacad.org
- Kirby, D., Korpi, M., Adivi, C., and Weissman, J. (1997). An impact evaluation of SNAPP, a pregnancy- and AIDS-prevention middle school curriculum. *AIDS Education Prevention*, 9(Suppl A), 44-67.
- Lederman, R., Chan, W., and Roberts-Gray, C. (2004). Sexual risk attitudes and intentions of youth age 12-14 years: Survey comparison of parent-teen prevention and control groups. *Behavioral Medicine*, 29(4), 155-163. Available to read online: www.findarticles.com/p/articles/mi m0GDQ/is 4 29/ai n6192603
- Lederman, R., and Mian, T., (2003). The Parent-Adolescent Relationship Education (PARE) Program: A curriculum for prevention of STDs and pregnancy in middle school youth. *Behavioral Medicine*, 29, 33-41.
- Looker, D., and Thiessen, V. (2004). Aspirations of Canadian youth for higher education. Learning Policy Directorate, Strategic Policy and Planning, Canada. Available to read online: www.pisa.gc.ca/SP-600-05-04E.pdf
- Miller, B. (1998). Families matter: A research synthesis of family influences on adolescent pregnancy. Washington, DC: National Campaign to Prevent Teen Pregnancy.
- Morowsky, J., and Ross, C. (2003). *Education, Social status, and Health.* New York: Aldine de Gruyter.
- Pamuk, E., Makuc, D., Heck, K., Reuben, C., and Lochner, K. (1998). Socioeconomic Status and Health Chartbook. Health, United States, 1998. Hyattsville, MD: National Center for Health Statistics, Table 9, page 185 of 464. Available to read online: www.cdc/nchs/data/hus/hus98.pdf

- Tiongson, A. (1997). Throwing the baby out with the bathwater: Situating young Filipino mothers and fathers beyond the dominant discourse on adolescent pregnancy. In: M. Root (Ed), *Filipino Americans: Transformation and Identity*, Thousand Oaks, CA: Sage.
- Trusty, J. (1999). Effects of eighth-grade parental involvement on late adolescents' educational expectations. *Journal of Research and Development in Education*, 32, 224-233.
- Trusty, J., and Harris, M. (1999). Lost talent: Predictors of the stability of educational aspirations across adolescence. *Journal of Adolescent Research*, 14, 359-382.
- U S Department of Health and Human Services. (2000). *Healthy People 2010: Understanding and Improving Health* (2<sup>nd</sup> ed). Washington DC: US Government Printing Office.

Chapter 22

## YOUNG CHILDREN'S DECISIONS ABOUT THE POWER OF INFERENCE<sup>\*</sup>

## Norman H. Freeman<sup>1</sup>, Angeliki Varouxaki<sup>1</sup> and Katerina Maridaki-Kassotaki<sup>2</sup>

<sup>1</sup>University of Bristol, UK <sup>2</sup>Harokopio University, Athens, Greece

## ABSTRACT

Seeing in a box is a direct way of knowing what is in the box. There is evidence that 4-year-olds engage with the idea that equally reliable knowledge can also be gained indirectly, via inference. For example, if you see that there is one cup to each saucer, the cups can be put away out of sight, and just by counting the saucers you can infer the number of the uncounted cups. But sometimes it is possible to make an inferential mistake. We review recent evidence of young children's decisions on whether people know things via inference, and add new evidence to test a claim about a false inference test. That test involved asking children to judge another's knowledge through inference when the other was misled to input the wrong premises into her calculation. The finding that children performed better in the false inference task compared to the true inference task, together with their explicit verbal justifications, attests to children's theoretical understudying of inference. We suggest that that is the growth point for the next round of research.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

## INTRODUCTION

The human mind makes much use of inference. Three-year-olds use inference in planning action in very simple tasks (Halliday, 1977). Four-year-olds, as reported by Pears and Bryant (1990), are prepared to use transitive inference about spatial position without any prior training when working out how to organise their actions in building a tower (e.g. if the blue block has to go higher than the red block, and the red block has to go higher than the green block, then the inference is that the red block will be lower than the blue block). That is, in deciding what to do in simple situations, preschool children come to rely on inference: but to what extent do the children understand what it is that underpins their decisions? It has been argued by Karmiloff-Smith (1992) that once children achieve practical behavioural mastery of something, they normally progress towards a theory of what they are doing, eventually becoming aware of their theory. So the question of how explicitly children can explain their inferential judgements bears closely on assessment of the speed of representational change in normal childhood.

One way of thinking about childhood representational change is to posit that children start by (a) grasping that seeing manifestly causes knowing, and then (b) progress to conceptualising unobservable inference as a generator of knowledge. Sodian and Wimmer (1987, p.432) suggested that a young child might come to make 'an amplification of his or her empiricist theory of knowledge'. Another testable hypothesis is that children's minds teem with ideas, and they have problems in discriminating (a) when they know something, from (b) when merely hopeful guesses arise in their minds. Children have to learn that there are permission conditions for claiming knowledge; and a sign of that would be a readiness to acknowledge their ignorance even when it would be easy to make a hopeful guess. In that light, inference can be regarded as a constraint on knowledge: if you follow a chain of inference, your mind is channelled towards one piece of knowledge. It is possible that children do not so much need to realise that inference *generates* knowledge as that inference constrains knowledge.

In this chapter, we consider some new and old evidence on

- a) children's use of inference under different conditions,
- b) children's adeptness in spotting conditions under which inference is safe and apposite,
- c) children's readiness to give satisfactory explanations of how their own and other people's inferences operate, and
- d) children's readiness to extend their expertise into more than just solving problems, into the situation where they are asked to predict when an inference will or will not lead to knowledge.

## **STUDIES ON YOUNG CHILDREN'S GRASP OF INFERENCE**

One line of work on young children's efficiency at making inferences in action planning (Halliday, 1977; Pears and Bryant, 1990) was noted at the outset. Such work is invaluable in identifying the earliest manifestations of inferential decision-making, but is not so suitable for

investigating further development towards reflective awareness. Here, another line of work seems to be proving its worth. The work stems from that of Sophian, Wood and Vong (1995) on numerical decisions. Briefly, children watched as a few Hawaiian frogs went to a party, each frog in its own boat. The frogs hopped into the party (actually a box that was then closed up, each frog leaving its boat moored outside. From counting the number of boats one can infer how many frogs were at the party; and one cannot infer how many dolphins had gone to the party because the dolphins swam in a group, leaving no visible traces outside the party. Sophian et al reported that with encouragement, something like half the preschoolers inferred the number of frogs, and appropriately refrained from using the boats to infer the number of dolphins who were at the party. The dolphin control condition gives one confidence that the successes with frogs were not a mechanical procedure, e.g. due to the children merely repeating the last number-word that was in their heads after counting boats: the last-word repetition tendency is always a worry in number research (Freeman, Antonucci & Lewis, 2000). In short, the 'frogboats paradigm' seems suitable for gaining evidence on children's adeptness in spotting conditions under which inference is safe and apposite. For follow-up research on 4-year-olds' numerical inference see Muldoon, Lewis and Freeman (2003); Muldoon, Lewis and Towse (2005). For confirmation of a contribution of logical reasoning to the learning of mathematics in primary school see Nunes et al (2007). We shall return to numerical inference below. For the present, we note one final line of study on inference. The work stems from an interest in children's dawning understanding of the power of the mind.

An early mentalistic insight is understanding that seeing can lead to knowing. Thus, many 3-year-olds grasp that a person who looks in a box knows what is in the box, and that someone who has not looked in a box does not know what is in it (Pillow, 1989; Pratt and Bryant, 1991). A discrimination between 'seeing is knowing' and 'not-seeing is not-knowing' is immensely useful as a step in grasping that there are causes of knowledge. But it can become counterproductive if held to too firmly; because under some conditions, inference can safely be used as a *substitute* for direct observation. Imagine that instead of one box, there are two boxes: you and your partner see a red pen and a green pen on the table, and do not see which pen goes into which box. Your partner need only look into one box to infer what is in the *second* box. You in turn infer that she knows because you apply a 'seeing is knowing' rule to the box she looks into, and then set aside the rule (a) that she has to see into the second box in order to know what is in it, in favour of (b) its inferential *functional equivalent* for the second box with its unseen contents. That is a complex procedure where a seeing rule has to be switched in and then inhibited; especially if you yourself had not looked into any box and knew that you did not know what your partner knew, so not-seeing is still not-knowing for you. The test formed part of the design invented by Keenan, Ruffman and Olson (1994). The work provoked interest, see Friedman, Griffin, Brownell & Winner (2003); Rai and Mitchell (2006). Yet there is one limitation on what the work showed. There was little evidence bearing on the question of children's readiness to give satisfactory explanations of how their own and other people's inferences operate. That was studied by Varouxaki and Freeman (1998);

Varouxaki et al (1999). Briefly, many of the many preschoolers studied managed to give satisfactory explanations of how their own and other people's inferences operated in the twobox game. Further, there was evidence of children's readiness to extend their expertise into the situation where they are asked to predict when an inference will or will not lead to knowledge. We shall return to that shortly. Before doing so, it is useful to take a close look at the method used to study children's inferential knowledge.

We shall adopt the notation used in Sodian and Wimmer (1987, Table 1) who first tested for a concept of inference as a source of knowledge, and label the basic task as '-+'. That symbol order means that the child does not know (-) the colour of the item that the second person does know (+). Suppose a child gives correct answers to two inference questions in this test: 'Do you know the colour in the other person's box? and 'Does the other person know the colour in your box?' by answering 'no' and 'yes' respectively. There is a danger of a false positive, in that the child might credit the other person with omniscience, so one needs the converse +- condition to elicit 'yes' for self and 'no' for other. In Keenan et al's (1994) Experiment 2, only a meagre 4/16 of their 4-year-olds seemed to show complete success. The majority error was to neglect the other's inference. In the authors' Experiment 3, efforts were made to remind children that the other remembered the items seen at the start, and success rose to 9/16. Before that point in the paper the authors had given up asking the children to explain their answers, because they had encountered very uninformative replies, so one cannot chart the progress of those children towards fully explicit understanding. But it seems that at least some 4-year-olds have enough of a concept of inference to make correct judgements in practice. Development seems to be a slow process: Keenan et al (1994) found that success on their inference test showed little and nonsignificant improvement from 4-yearolds to 5-year-olds, and lagged behind success on a false belief test. There are, though, three things to find out in order to explain what achievement some children had shown.

The first worry relates to what Keenan et al (1994) did which just about doubled the success level between their second and third experiments. Just before asking the critical questions the Experimenter intervened with words and pictures to stress the facts that the other had seen two colours (shapes) initially and knew what they were. We are concerned that some children might then have solved the task by contrasting seeing with remembering in the following way. In the -+ condition, the child might think something along the lines of 'I am being asked if the other knows what he cannot see; I have just been told that the other knows the two things, so he must be remembering the other thing he knows'. In the +- condition, the child might be able to work out that the other knows two things and is not looking into any box so cannot give a one-box answer. Certainly, inference is involved, but there is more to inference than realising that others might have good reiterative memory for some knowledge that had just been explicitly mentioned. The minimal step one can think of is to make the inferential demand more generative: the task should involve coming up with an answer that had not been explicitly stated up to then. Thus, one can have tests where items varied in number, and the child and other count them together. If there are, say, four items, to be divided between two boxes, any one box could contain one, two, or three items, so an inference involves coming up with a new cardinal number. In short, placing either a red or a green item inside a box does not alter the items' colour-identity, whereas dividing a set of items between two boxes generates subsets of new numerosity. There is an additional benefit from using simple arithmetic: counting out loud might draw to children's attention that something needs to be worked out. It was reported by Varouxaki and Freeman (1998); Varouxaki et al (1999) used tests that dispensed with the questionable intervention noted above, and reported that there was no difference between colour-inference and number inference results. So an understanding of inference as a source of knowledge has some degree of generality across tests.

The second thing one wants to know concerns the extent to which 4-year-olds who succeed on the test have a theoretical understanding of what they are doing. One source of evidence was mentioned at the outset: to ask the child to explain her judgements in order to tap whether she uses her ideas in explanatory manner (see Wellman, 1990). Another source of evidence for the existence of a theory is needed, namely that a child should be ready to use it to make predictions (see Perner, 1991). One can incorporate a test to find out if the child can *predict* whether she and another person would infer what is in a particular closed box under hypothetical viewing conditions. In such a situation, it was reported by Varouxaki et al (1999, Experiment 1) that the preschoolers made some 89% correct predictions, mostly with lucid explanations. Here is an excerpt from a child in a later replication we ran: 5-year-old Ella in a – condition, volunteered that if she were to look inside the other person's box she would know how many pens were in her own box: 'There were four pens and he'll have some and I'll have some, and I'll know how many I've got 'cause how many he's got. If he's got two I'll have two'. There can be not a shadow of a doubt that a child such as that feels safe in

basing her decisions upon an inferential process that runs from observable to unobserved

displays. The third thing that one wants to check concerns whether the children realise that other people can think in precisely the same way that they themselves do. Riggs and Robinson (1995) recently found evidence of young children having a greater awareness of the workings of their own minds than that of others. At first sight, the phenomenon of a child in the +condition saying that she knows what is in the second box without looking into it and denying that someone else in the -+ condition can gain knowledge by inference ('inference neglect') is congruent with the notion that children solve advanced mentalistic problems for self before other. The strongest evidence for a representational theory of mind is if, whenever a new mentalistic insight occurs, it occurs simultaneously for first-person and third-person application (Gopnik, 1993). Yet here we have a type of a judgement which sometimes reliably occurs for self before being extended to others (Sodian and Wimmer, 1987; Keenan et al, 1994), and sometimes does not (Varouxaki and Freeman, 1998; Varouxaki et al, 1999). Again, there was one phenomenon that was noted in the very first experiment of Sodian and Wimmer (1987) where 88% of all wrong answers to the question of whether the child herself knew the colour came from the child claiming to know when she actually could not know. So it seems that many 4-year-olds both overestimate their own knowledge and underestimate another's knowledge, thus treating other minds as very much 'stupider' than their own in two respects. The tendency of children up to the age of 6 years or so to fail to acknowledge their own ignorance has long been documented (e.g. Beal and Flavell, 1982; Markman, 1977; Mitchell and Robinson, 1992; Sophian and Somerville, 1988). Acknowledging when oneself is ignorant is a key step in coming to grips with the problem of knowledge, since it involves the realisation that there are constraints on knowledge. From that perspective, realising that other people can attain knowledge via inference might merely be part of the general problem of extending a grasp of constraints from one's own mind to that of others. Accordingly, one needs a design that will make it easy to track whether those particular children who are ready to acknowledge their own ignorance are also ready to credit others with knowledge via inference. That can be done by studying the full Sodian-Wimmer matrix of --, +-, -+ and ++ conditions where Keenan et al (1994) used only the +- and -+ conditions. It was from using the full matrix that enabled Varouxaki and Freeman (1998); Varouxaki et al (1999) to report that the child's understanding of inferences for self kept pace with inferences for others.

One of the long-term aims of research is to model the emergence of new ideas in competition with old ideas. Kuhn (1992, p.256) produced the general argument that "The most formidable problem for subjects appears to be not the acquisition and consolidation of new strategies but rather the ability to abandon old less adequate strategies -- a reversal of the way in which we typically think about development". It might indeed be the case that 4-year-olds understand that inference leads to knowing but overvalue the immensely useful rule that seeing is knowing, so their competence is masked. It is there that the -+ and -- conditions are so useful: the child cannot see into any box so she cannot know what is in either box, and any false claim to knowledge must be in despite of 'seeing is knowing'. In that light, one looks to whether a readiness to acknowledge own ignorance is a predictor of crediting others with inferential knowledge, as part of an emerging theory of constraints on knowledge.

# **EVALUATION OF THE RESEARCH SO FAR**

Keenan *et al* (1994) were correct to assert that the fifth year of life is about right for a test of a fledgling understanding that inference leads to knowledge. The similarity in results between Colour and Number found by Varouxaki et al (1999) enhances confidence in generalizing from the test. Some children had reached the highest level of articulating reflective awareness; other children were still showing inference neglect. We concentrate next on them.

The test involves converse inference: the child is asked whether she can infer the contents of the other's box and whether the other can infer the contents of the child's box. So success cannot come from egocentric reiteration since different colours/numbers are asked about even in the ++ condition where both protagonists have knowledge. That ++ condition should be the easiest in the matrix of conditions, since the child can see that it is possible for her to make an inference and that the other has a comparable experience. Yet inference neglect occurred in a third of those children who correctly assessed own knowledge. A crucial new finding by Varouxaki et al (1999) wass that when attention was focused on those children who avoided inference neglect, there was no significant drop in the frequency with which inferential explanations were given of how the other knew compared with how self knew. So another's mind was treated the same as own mind: in making a judgment of whether another knew and in reflecting on *how* the other knew. That is not congruent with the argument of Karmiloff-Smith (1992) that children mentally represent what they know more than once as they develop, with conscious awareness coming at the end of a process of what she termed 'explicitation'. Note that a two-step process has been suggested before for a lower-level problem in theory of mind: another's false belief is computed at an implicit level before becoming available for judgment at the verbal level (Clements and Perner, 1994; Freeman, Lewis and Doherty, 1991). It would be prudent to leave open the possibility that advance in technique might eventually show that the child's encounter with the problem of inferential knowledge similarly involves a step of collating own and other minds more than once. It is possible to be more precise about the developmental issue via the following consideration of the problem of knowledge.

Sodian and Wimmer (1987) argued that children need to 'enrich' their theory of mind to give inference equivalent weight to seeing as a source of knowledge. Keenan *et al* (1994), in

their Discussion section emphasized the importance of not over-valuing seeing is knowing. The emphasis is on the acquisition of a generative theory of mind to encompass the notion that other people generate knowledge from inference. We suggest that the data on acknowledgment of own ignorance helps make an account less one-sided by focusing attention on the functional point that both seeing and inferring are ways of constraining belief. Seeing provides one constraint which can be directly experienced for oneself and observed in the behaviour of another person, inference provides another constraint which cannot be directly observed in others' behaviour.

# **THE NEXT RESEARCH STEP**

The coding of experimental conditions as +-/++/--/-+ was possible because there are two possible states of mind in the experiment: either (a) someone cannot make an inference on that trial and so they are left uninformative about the hidden item(s), or (b) someone can make an inference and the result will be correct and informative about the hidden item(s). Yet it is commonly easy to be able to make an inference yet it turn out to be wrong. It would arise if the child were to watch someone who miscounted the frogboats, so that person would tell you a wrong number for how many frogs she inferred to be at the party. So that is an interesting case to study, to find out whether the child's conception of inference encompasses only true inference and does not yet extend to false inference. A person who is making a false inference is doing so in all honesty because she has a false belief about the number of frogboats. Accordingly, we need briefly to consider the acquisition of an understanding of false belief.

In experiments concerned with the developmental contingency between having a concept of inference and understanding false belief, Keenan et al (1994) found that performance on a false belief test preceded understanding of inference, whereas Varouxaki et al (1999) reported no pattern of association between children's performance on the two tests. Yet one would expect an association, if the following brief argument is on the right lines. In an inference test, to be correct children need to grasp that the other person has no direct access to the critical fact but that she does have access to information that allows her to make an inference. Similarly to pass the false belief test, children have to grasp that another person (a) has no direct access to an actual state, but that she (b) has access to relevant yet misleading information (Wimmer et al, 1988). If this 'informational access' approach accounts for difficulties in understanding both inference and false belief, then children should perform consistently across tests. There is an alternative approach: young children have difficulty with the representational complexity involved in false belief (Wimmer and Perner, 1983). The difficulty for the child arises because of the necessity of representing that the other holds a certain proposition as a true representation whilst this very same proposition represents misrepresentation. The mental truth of the other's belief needs to be represented rather than the child's own truth. The representational complexity of representing conflicting truth values does not apply to the Sodian and Wimmer and Keenan et al types of inference, because the same premise information holds true for the child and the story protagonist. So it might be comprehensible why different researchers have obtained different findings on the relations between false-belief and inference understanding: wherever there is a task-demand mismatch,

error may enter into the data. Accordingly, we add a third test to those two: a false inference test. Here, we adopt the frogboats design and the child sees the other person miscount the boats. In total, two studies were conducted, and children's performance on the false inference and standard false belief test was cross-compared. Before reporting what thappened, it is well worth briefly scanning the existing literature on true and false inference tests.

# **TRUE AND FALSE INFERENCES**

Research into the child's theory of mind began with the problem of false belief. That is, preschool children were often adept at spotting when someone held a true belief about a state of affairs vet could not seem to understand that someone could hold a false belief. So the children's' problem was held to be that of mastering a particular type of mentalistic 'representational complexity'. Children's performance would naturally vary under particular conditions, but the order of acquisition of true belief first then false belief was reliable (a 'true-belief advantage'). A problem arose when Riggs and Simpson (2005) found a situation in which both true and false beliefs were equally easy, or equally hard, for preschoolers to ascribe. That finding generated puzzlement; and in respect of false belief it was suggested that 'we should think harder about it' (Russell, 2005). Russell, Hill and Franco (2001) had independently also reported no difference between true belief and false belief success in a novel intention-reporting test (and for more standard null results under various conditions see Garnham & Ruffman, 2001; Roth and Leslie, 1998; Surian and Leslie, 1999). Lohmann, Carpenter and Call (2002) reported conditions under which the true-belief advantage only appeared under low processing demands; that is, the particular tasks effectively modulated the true-belief advantage. Now let us return to the inference test. What would one predict for an inference test that involved the participant forming a false belief? Would the child be as ready to ascribe inference to that participant as in the usual situation where the participant formed a true belief? Some insight into that may be gained from Ruffman (1996) who reported that children performed reliably better on the false inference test: a reversal of the true-belief advantage that had started the whole theory-of-mind tradition. Yet note that Ruffman (1996) preceded the Think question (what colour the protagonist thinks the sweet in the box is) with a Know question (whether the protagonist knows what colour the sweet is) as in the true belief inference condition. That pegged question-order was by design, to suit Ruffman's ingenious prediction of a false-belief advantage. Once children had judged the protagonist not to know what colour sweet was in the box, they should then attribute a false belief about the colour in order to be consistent( knowing is linked to true belief, not knowing therefore cannot be: the concept of a 'lucky guess' might be beyond preschoolers). But the procedure then needs amending to drop the Know question, cutting out any possibility of carry-over from that question to the Belief question. That is, we need to make the procedure, asking simply what the protagonist thinks without preamble so as to make the inference test comparable to the standard false belief test. And there is a need to avoid ambiguity possibly attendant on the lucky guess problem, and on worries about a protagonist who might not know if something was true compared with a protagonist who did know about something, only they knew the wrong thing resulting in an inferential falsity, see Bradmetz and

Bonnefoy-Claudet (2003). We can ensure that the protagonist knows the wrong thing as follows.

We pick up on numerical one-to-one correspondence as an inferential cue. It was noted earlier on that there was an interesting instance of numerical inference pioneered by Sophian, Wood and Vong (1995) whereby children watched as frogs went to a party, each frog in its own boat, with the boats being left outside. From counting the number of boats one can infer how many frogs were at the party. Roughly half the preschoolers succeeded (for extensions of the work see Muldoon, Lewis and Freeman, 2003); Muldoon, Lewis and Towse, 2005). The 'frogboats paradigm' gives one evidence on children's adeptness in spotting conditions under which inference is safe and apposite.. The paradigm can readily be adapted for judgement of self and other's inference. And one can easily ask for explanations to track the emergence of verbalisable explicit awareness. One can add a condition in which the protagonist happens to end upwith a miscount, so making a false numerical inference

We worked with 47 four-year-old children (22 girls and 25 boys) attending a primary school in a middle class residential area in Avon. They ranged in age between 49 to 60 mo (mean age M = 54 mo and SD = 4 mo).

The procedure always began with a version of the standard false belief task before any of the inference conditions. The procedure for the standard false belief test was a version of Gopnik and Astington's (1998) deceptive container test. Children were shown an egg-box with its true contents replaced with eggs and were asked what they thought was inside. Then, the box was opened revealing onions. The box was closed up and children were asked 'When somebody else from your classroom sees the box, all closed up like this, what will she think is inside?', then 'What really is inside?'

For the inference tests, the material was as follows. A total of 18, hand-drawn four page picture booklets were for the inference ++, 6 for the inference +-, and 6 for the false inference conditions. For each condition there were three sets of stories (a cooking story, cake story, and a picnic story). For each story there were two versions varying the number of target items. For the standard false belief test, there was an egg box with the normal contents replaced with onions. All testing took place in a quiet room outside the classroom.

Children received three conditions: inference ++, inference +-, and false inference. Three orders of conditions were created according to a balanced Latin square design. For the purpose of simplicity the following description of methods will be based on the cooking story; for the cake and picnic stories, the questions and prompts were phrased appropriately.

Each booklet showed two protagonists about to cook some fish. There were five experimental groups. One group were asked to count the number of fishes on the plates, while the other groups counted the plates on the table. There was always one extra plate next to the ones containing the fishes and it was pointed out to the children that this plate was left empty because it was cracked. The children who counted plates were either asked to count forwards or backwards (counting backwards means that counting ends up with the number 'one', so that the half of the children there cannot rightly simply reiterate their last-mentioned number in answer to the inference question). Of those children who counted the plates half were asked to include the empty, cracked plate, while the other half were instructed to ignore it (again that forces half the children to verbalise a last number that is different from the number they need because of the over-run onto that irrelevant plate). Then one of the protagonists left the room, whilst the other put some fishes into a saucepan and closed the lid. The children recounted the fishes/plates on the table. They were the asked the Know question about their

*own* inference e.g. "Do you know how many fishes are in the pan? [If the child had not spontaneously given a number, she was then asked "How many"? to ensure that she was ascribing a true inference to the story protagonist], followed by a request for explanation "How do you (why don't you) know that?" Then the absent protagonist returned, re-counted and the target child was asked the Know question about the *other*'s inference: "Does s/he know how many fishes are inside the saucepan"? (again if necessary asking how many, followed by a request for explanation).

The +- and false belief inference trials followed the same procedure, except that (a) in the +- condition the absent protagonist never returned to the kitchen and (b) in the false inference condition a cat ate some of the fishes left on the plates before the protagonist's return. Because the fishes that were left on the plates and those now in the pan do not add up to the initial number of fishes on the table, and because the absent protagonist had not seen the fishes being transferred into the pan, he could not know how many were really in it. Participants were asked the false belief Think question: 'How many fishes does he think are inside the pot?' and the Reality control question 'How many are really in the pan?'. It may be recalled that Ruffman (1996) in his study of true inference preceded the Think question with a Know question as in the true belief inference condition. We chose not to do so here because traditionally tests that are concerned with false beliefs in self and others rely on the Think question, children's answers to the Think question were sufficient to gain an insight into their understanding of inference as set out by our research questions, and we were concerned about children mechanically going in for response switching.

We begin the analysis with the basic data of yes/no answers to the question of whether self or other knows the number of items transferred into an opaque container over five questions (two for self and three for other). Inspection of the data revealed that there was no effect of the counting manipulations (counting targets versus plates, forward or backward, with or without the inclusion of the extra plate) on total correct responses, confirmed by one way Analysis of Variance (F(4, 42) < 1). Hereafter these are not considered as variables.

Individual response patterns. The correct response pattern was to say (a) 'yes' for one's own and the other person's knowledge in the ++ condition, and (a) 'yes' for self but 'no' for the other in the +- and false belief inference condition. There were 17/47 (36%) children who were completely correct. Two incorrect response patterns were identified. Inference neglect involved neglecting the other's inference in the ++ trial whilst being correct for self --17/47 children (36.2%) showed this pattern. The remaining 13 children (27.6%) were either yea/no-saying, response switching, overestimating the other's knowledge, or giving uninterpretable responses.

The standard false belief test which was passed by 23/47 (48.9%) children was not a significant predictor of performance on the inference task. An analysis of covariance where standard false belief is the independent variable, inference is the dependent variable and age is the covariate showed a significant main effect of standard false belief understanding on the total number of correct responses on the inference task, F(1, 44) = 8.83, p < .01. Age did not have a significant effect as a covariant F(1, 44) < 1). The suggestion is that children who pass the false belief test are likely to give more correct answers on the inference trials.

Yet more correct answers do not necessarily translate into a correct response pattern. Table 1 is a cross-tabulation of children's false belief performance and response patterns for Inference. There was a significant difference between the 17 children who were correct for Inference and the remaining 30 children in whether they passed or failed the false belief test,  $\chi^2(1, n = 47) = 8.08, p < .01$ . Passing the false belief test is not a sufficient condition for understanding inference as 9 out of the 17 children who showed neglect on the inference task had passed the standard false belief test. In fact there was no significant difference between the 17 children who were correct in the inference trials and those 17 who showed neglect in whether or not they passed the false belief test,  $\chi^2(1, n = 34) = 2.06, p > .05$ . That (a) 23 children passed the false belief test but only 13 were completely correct for inference together with (b) the absence of a pattern of association in children's performance on the two tests supports Wimmer at al's (1988) claim that understanding inference is a distinct and later attainment than understanding false belief.

False Belief	Infe	erence Response Pattern	
	Correct	Neglect	Other
Pass	13	9	1
Fail	4	8	12

 Table 1. Cross-tabulation of performance on the standard false belief test and inference response patterns

There was no significant difference in the number of children (31/46 = 66.0%) who passed the false inference test and the 49% of children who had passed the standard false belief test (binomial p > .05). Table 2 shows the contingency between performance on the new false inference test and the traditional deceptive container test. In Table 2 the association between the false belief inference and traditional false belief tests was high,  $\chi^2$  (1, N = 47) = 17.7, p < .001. There were 37/47 children (78.7%) who either passed both or failed both the deceptive container and false inference task. Of the remaining 10 children, 9 passed the traditional false belief test and failed the new false inference task whilst only one child showed the reverse pattern. The results are congruent with the hypothesis that both false belief and inference require an appreciation of the contingency between informational access and the resulting knowledge.

Table 2. Cross-comparison of performance on the faise inference and standard faise							
belief tasks							

False Inference	Standard Fa	lse Belief
	Pass	Fail
Pass	22	1
Fail	9	15

An analysis of covariance where false belief inference is the independent variable, inference is the dependent and age is the covariate showed a significant main effect of false belief inference understanding on the total number of correct responses on the inference task, F(1, 44) = 6.56, p = .01. Age was not a significant covariant F(1, 44) < 1. Next, we examined whether success on the false inference task predicts the patterns of children's responses on the inference conditions, Table 3.

There was no significant difference between the 17 children who were correct on the inference trials and the remaining 30 children in whether they passed or failed the false inference test,  $\chi^2(1, n = 47) = 1.84$ , p = .17. Importantly, 17 children passed the false inference task but not the inference task and only 3 showed the reverse pattern, Binomial p = .03. Children actually found it easier to ascribe an inferential false belief than to ascribe an inferential true belief.

 Table 3. Cross-tabulation of performance for false inference and true inference response patterns

False Inference		True Inference	
	Correct	Neglect	Other
Pass	14	14	3
Fail	3	3	10

Children's explanations were considered inferential when sufficient information was given of how reasoning had been involved. We classified children's explanations across conditions for self and other as: (a) equal proportions of inferential explanations for self and other, (b) proportionately more inferential explanations for self, (c) proportionately more inferential explanations, and (e) insufficient evidence (it was possible to have insufficient evidence because children who showed inference neglect had no opportunity to give inferential explanations for other). On inspection of the data there were 19/47 children (40.4%) for whom we had insufficient evidence to classify the pattern of their explanations.

Contrary to Keenan et al's (1994) study, all the remaining children were able to give sufficient information to see how reasoning had been implicated in solving the inference. Of the 28 children for whom we had sufficient evidence 24 children (88.9%) gave equal proportions of inferential explanations for self and other, 2 children (7.4%) gave proportionately more inferential explanations for self, and 2 children (7.4%) gave proportionately more inferential explanations for other than for self. These data are consistent with Ruffman (1996) and Varouxaki et al (1999).

On cross-comparison of the pattern of children's responses with the pattern of their explanations, 17/19 children for which we had insufficient evidence had shown neglect whilst the remaining 2 children were 'other'. There is no discernible difference between the remaining 17 children who were correct and those 11 who were 'other', in whether they gave equal proportions of inferential explanations for self and other or were biased (Fisher's Exact Probability Test p = .57).

There was no discernible difference between the 24 children who gave equal proportions of inferential explanations for self and other and the remaining 23 children in whether they had passed or failed (a) the standard false belief test, and (b) the false inference test. Twelve and 13 children had passed false belief respectively. The respective numbers for the false inference test were 14 and 17. The results of the true inference test are in line with past research (Sodian and Wimmer, 1984; Varouxaki et al, 1999): over a third of the children were completely correct, an equal proportion neglected another's knowledge through inference, and with the remaining children showing no understanding that the task was inferential. As

with Keenan et al (1994) children's performance for inference lagged behind understanding false belief. A new finding was that there was a strong association between children's Think judgements on the false inference test and the standard false belief test. An important finding was that children found the new false inference task easier than the true inference task. The view that children at this age hold a theoretical understanding of inference was further supported by their explanations.

## CONCLUSION

The main finding to take forward into the future was the children's readiness to spot when someone was basing an inference on a false belief compared with when they were basing inference on a true belief. That is an interesting reversal of the usual because the theory-of-mind accounts in the literature agree that false belief understanding is a later accomplishment than true belief understanding. Could it be that children are ready primed to spot when someone is, as it were, riding for a fall, inferentially speaking?

## REFERENCES

- Beal, C. R. & Flavell, J. H. (1982). The effects of increasing the salience of message ambiguities on kindergartners' evaluations of communicative success and message adequacy. *Developmental Psychology*, 18, 43-48.
- Bradmetz, J., & Bonnefoy-Claudet, C. (2003). Do young children acquire the meaning of to know and to believe simultaneously or not? International Journal of Behavioral Development, 27, 109-115.
- Clements, W. A., and Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9, 377-395.
- Freeman, N.H., Antonucci, C., & Lewis, C. (2000). Representation of the cardinality principle: early conception of error in a counterfactual test. *Cognition*, 74, 71-89.
- Freeman, N. H., Lewis, C., & Doherty, M. (1991). Preschoolers' grasp of a desire for knowledge in false-belief prediction. *British Journal of Developmental Psychology*, 9, 139-157.
- Friedman, O., Griffin, R., Brownell, H., & Winner, E., (2003). Problems with the seeing = knowing rule. *Developmental Science*, 6, 505-513.
- Garnham, W. A., & Ruffman, T. (2001). Doesn't see, doesn't know: is anticipatory looking really related to understanding of belief? *Developmental Science*, *4*, 94-100.
- Gopnik, A. (1993). How we know our own minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences*, 16, 1-14.
- Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 59, 26-37.
- Halliday, M. S. (1977). Behavioural inference in young children. Journal of Experimental Child Psychology, 23, 378-390.

- Karmiloff-Smith, A. (1992). Beyond modularity: A developmental perspective on cognitive science. MIT Press, London.
- Keenan, T., Ruffman, T., & Olson, D. R. (1994). When do children understand logical inference as a source of knowledge? *Cognitive Development*, 9, 331-353.
- Kuhn, D. (1992). Cognitive development. In M. H. Bornstein and M. E. Lamb (Eds.), Developmental psychology: An advanced textbook (pp. 211-272). Hillsdale NJ: Erlbaum.
- Lohmann, H., Carpenter, M., & Call, J. (2002). Guessing versus believing and seeing versus believing – in false belief tasks. *British Journal of Developmental Psychology*,23, 451-469.
- Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development*, 48, 986-992.
- Marsh, R. W. (1967). Tables for testing the significance of difference between proportions. *Australian Journal of Psychology, 19,* 223-229.
- Mitchell, P. & Robinson, E. J. (1992). Children's understanding of the evidential connotation of 'know' in relation to overestimation of their own knowledge. *Journal of Child Language*, 19, 167-182.
- Muldoon, K., Lewis, C., & Freeman, N. H. (2003). Putting counting to work: preschoolers' understanding of cardinal extension. *International Journal of Educational Research*, 39, 695-718.
- Muldoon, K., Lewis, C., & Towse, J. (2005). Because it's there! Why some children count, rather than infer numerical relationships. *Cognitive Development*, 20, 472-491.
- Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A., & Carraher, J. (2007). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Developmental Psychology*, 25, 147-166.
- Pears, R., and Bryant, P. (1990). Transitive inferences by young children about spatial position. *British Journal of Psychology*, 81, 497-510.
- Perner, J. (1991). Understanding the Representational Mind. Cambridge, MA: MIT Press.
- Pillow, B. H. (1989). Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology*, 47, 116-129.
- Pratt, C., & Bryant, P. E. (1991). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development, 61,* 973-982.
- Rai, R., & Mitchell, P. (2006). Children's ability to impute inferentially-based knowledge. *Child Development*, 77, 1091-1093.
- Riggs, K. J. & Robinson, E. J. (1995). What people say and what they think: Children's judgments of false belief in relation to their recall of false messages. *British Journal of Developmental Psychology*, 13, 271-284.
- Riggs, K. J., & Simpson, A. (2005). Young children have difficulty ascribing true beliefs. *Developmental Science*, 8, F27-F30.
- Roth D., & Leslie A. M. (1998). Solving belief problems: Toward a task analysis. *Cognition*. 66, 1-31. 1998.
- Ruffman, T. K. (1996). Do children understand the mind by means of simulation or a theory? Evidence from their understanding of inference. *Mind and Language*, *11*, 388-414.
- Russell, J. (2005). Justifying all the fuss about false belief. *Trends in Cognitive Sciences*, 9, 307-308.

- Rusell, J., Hill, E.L., & Franco, F. (2001). The role of belief veracity in understanding intentions-in-action: Preschool children's performance on the transparent intentions task. *Cognitive Development*, 16, 775-792.
- Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development*, 59, 703-718.
- Sophian, C., Wood, A. M., & Vong, K. I. (1995). Making numbers count: The early development of numerical inferences. *Developmental Psychology*, 31, 263-273.
- Sophian, C. & Somerville, S. C. (1988). Early developments in logical reasoning: Considering alternative possibilities. *Cognitive Development*, *3*, 183-222.
- Surian L, & Leslie A. M. (1999). Competence and performance in false belief understanding: A comparison of autistic and normal 3-year-old children. *British Journal of Developmental Psychology*. 17, 141-155..
- Varouxaki, A., & Freeman, N. H. (1998). Young children's understanding of other people's inferences. *Psychologia: Journal of the Hellenic Psychology Society*, 5, 20-30.
- Varouxaki, A., Freeman, N. H., Peters, D., & Lewis, C. (1999). Inference neglect and ignorance denial. *British Journal of Developmental Psychology*, 17, 483-499.
- Wellman, H. M. (1990). The child's theory of mind. Cambridge, MA: MIT Press.
- Wimmer, H., Hogrefe, G. J., & Sodian, B (1988). A second stage in children's conception of mental life: Understanding informational accesses as origins of knowledge and belief. In J. W. Astington, P. L. Harris & D. R. Olson (Eds.), *Developing theories of mind* (pp. 173-192). Cambridge University Press.
- Wimmer, H, & Perner, J (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.

Chapter 23

# THE ROLE OF FILIAL PIETY IN THE CAREER DECISION PROCESSES OF POSTGRADUATE STUDENTS IN CHINA<sup>\*</sup>

# Leili Jin, Mantak Yuen, and David Watkins

The University of Hong Kong, Pok Fu Lam Road, Hong Kong

The research attention paid to career decision making among Mainland Chinese adolescents and college students is of recent origin (e.g., Creed & Wong, 2006; Hampton, 2005). Such attention is mainly due to major changes in two salient policy areas by the government of the People's Republic of China (PRC). Firstly, the labor distribution system has been changed from a planned to a market economic system in which employers and graduates are free to choose each other. Before the middle of the 20th century in the PRC, college graduates had to accept the jobs assigned by the government. The freedom of selection by both employers and fresh employees has stimulated Chinese graduates to consider how to choose a suitable career for their own sake. Secondly, higher education in the PRC is undergoing a radical transformation from an elite to a mass system. The increasing intake of the youth into college had led to the sharp competition in the labor market, which in turn has produced the employment pressure among college graduates. The number of college graduates pouring into the labor market each year has increased from 1.15 million in 2001 to 4.2 million in 2006 (China Ministry of Education, 2005). Meanwhile, graduate school admission has become a top priority for the majority of undergraduates: 0.714 million senior undergraduates sat for the national entrance examination for a Master's Degree in January,  $2007^{1}$ . Although the need for career and other counseling services in the Mainland is continually increasing (Zhang, Hu, & Pope, 2002), it is obvious that the antecedents of adaptive career decision making among Chinese postgraduates still remain unclear.

<sup>&</sup>lt;sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>1</sup> http://www.moe.edu.cn/edoas/website18/info23746.htm , 12-12-2006

#### Career Commitment and Decision Making Self-Efficacy

Theoretically, the process of career decision making covers sequential stages starting from a pre-awareness and moving through awareness, planning (exploration and crystallization), commitment, and implementation (Gottfredson, 2002; Harren, 1979; Jordaan, 1963; Super, 1957). For late adolescents and young adults entering into the transition from school to work, a salient achievement of career decision making is a strong commitment to a career choice through sufficient exploration of self and the world of work. Those who endorse higher career commitment are likely to have a clear sense of occupational preference, be awareness of the potential barriers and the willingness to overcome them; a confidence in abilities to achieve objectives; and actively preparing to implement a set of particular vocational goals (Blustein, Ellis, & Devenis, 1989; Harren, 1979; Jordaan & Heyde, 1979; Marcia, 1966; Super, 1957). Blustein et al. (1989) defined vocational exploration and commitment (VEC) as one's progress in attaining commitment to a career choice along with a process ranging from an uncommitted or exploratory phase to a highly committed and confident phase. They further asserted that there were individual differences in the means used to achieve career commitment. Some people may accept various experiences encountered in the career commitment process, while others would hastily make a career choice in order to avoid the ambiguous, uncertainty and anxiety. Such phenomena is defined as the tendency to foreclose reflecting "the desire to commit to important educational and career decisions as soon as possible and an analogous attempt to adhere to these choices even in the face of disconfirming evidence" (Blustein et al., 1989, p. 347). Presumably, such a tendency would restrict one's career exploratory space and yield a premature decision, which in turn hinder one's career satisfaction and success.

The construct of *career decision self-efficacy*<sup>2</sup> has been identified in the literature as a salient social cognitive factor in the career decision-making process. This construct is based on Bandura's (1997) self-efficacy theory and reflects beliefs in personal abilities to successfully complete tasks necessary in making career decisions (Taylor & Betz, 1983). Numerous studies have consistently showed that career decision self-efficacy is positively related to indices of adaptive career decision making, including career maturity, career exploratory activities, vocational identity, career certainty, and negatively to career indecision (see Betz, Hammond, & Multon, 2005; and Betz & Luzzo, 1996 for reviews). Therefore, career commitment, tendency to foreclose and career decision self-efficacy were used as outcome variables related to adaptive career decision making in the current study.

### Familial Factors and the Career Commitment Process

Researchers have called for more attention to the relationships between contextual / cultural factors and the process and content of career decision making (Blustein, 1997; Lent, Brown, & Hackett, 1994, 2000; Vondracek, Lerner, & Schulenberg, 1986a). This represents a shift of research focus from individual differences and self-implementation to individuals

<sup>&</sup>lt;sup>2</sup> The original construct was named Career Decision-Making Self-Efficacy. As stated by Nancy Betz and Karen Taylor, Career Decision-Making Self-Efficacy has been changed into Career Decision Self-Efficacy because another author has copyright of the term career decision making already. In the current study, we adopt the term career decision self-efficacy in order to be consistent with what Betz and Taylor propose.

embedded in a real social or relational context (Blustein, 2004). Among the contextual factors, family relationships have been indentified as salient in the personal career decision process, including attachment to and conflictual independence from parents, multiple indices of family environment (see Vondracek, Lerner, & Schulenberg, 1986b; and Whiston & Keller, 2004 for reviews), and perceived parental support and family barriers (Leal-Muniz & Constantine, 2005). Although past studies are quite impressive, the nature and magnitude of family relationship on career decision making still remain unclear (Blustein et al., 1991). In particular, Leong and Brown (1995) argued that some cultural-specific variables might miss when the existing career theories developed from the Western culture and value system are applied to culturally different populations.

For adolescents and young adults who endorsed collectivist values, family relationships presumably play a more prominent role in career decision making(Lent et al., 2003). A few of past studies reported that family involvement directly exerted a significant impact on career choices, inferring that Asian Americans (41.3% were Chinese Americans in their sample) who tended to follow their parent's wishes were more likely to choose careers in traditional areas. Although career self-efficacy was found to significantly predict career choices, it was not predicted by either family SES or family involvement as expected (Tang, Fouad, & Smith, 1999). In addition, higher levels of intergenerational or acculturation conflicts within families would likely lead to higher career indecision and lower career aspiration among Asian American youths. However, relational-interdependent self-construal was significantly predictive of career certainty, suggesting that individuals who had a closer relationship with their parents and an inclination to consult parents' opinions for the future were more likely to make a definite career decision (Constantine & Flores, 2006; Ma & Yeh, 2005).

Theorists generally assume that the intention of collectivists to consider or even follow parents' advice on career choices (Leong & Serafica, 1995) can be partially attributed to the high values placed on respect for seniors or obedience to authority (Ma & Yeh, 2005; Moy, 1992). Such values partially due to the obligation or responsibility to financially support the family and care for aged parents (Leong & Chou, 1994; Leong & Leung, 1994), and partially to the aspiration to bring honor to the family (Ying, Coombs, & Lee, 1999). The abovementioned devotion of children to the parents is typically reflected by the filial piety construct. To date, little research has examined the relations of filial piety and adaptive career decision making among Chinese students who are from a representative collectivist culture (Oyserman, Coon, & Kemmelmeier, 2002).

#### Filial Piety

*Filial piety* addresses how children should treat toward their parents, living or dead, as well as attitudes toward their ancestors, which is a prominent ethical rule guiding intergenerational or interpersonal relationships in Chinese families and society (Ho, 1996). Recently, Yeh and associates have proposed a dual filial piety model to reflect the developmental changes of filial piety in the contemporary era (Yeh, 1997a, 2003). The dual model involves four common dimensions. "Respecting and Loving Parents" and "Supporting and Memorializing Parents" were combined into one higher order factor termed reciprocal filial piety, whereas "Oppressing Oneself" and "Glorifying Parents" formed another superfactor labeled authoritarian filial piety.

*Reciprocal filial piety* is defined as "emotionally and spiritually attending to one's parents out of gratitude for their efforts in having raising one, and physical and financial care for one's parents as they age and when they die for the same reason" (Yeh & Bedford, 2003, p. 216). In other words, the harmonious relationships with and the responsibility to care about the seniors out of affection and gratitude are underlined by reciprocal filial piety based on the principle of reciprocity and favoring the intimate(Yeh, 2003; Yeh & Bedford, 2004). *Authoritarian filial piety* is characterized as "suppressing one's own wishes and complying with one's parents' wishes because of their seniority in physical, financial or social terms, as well as continuing the family lineage and maintaining one's parents' reputation because of the force of role requirements" (Yeh & Bedford, 2003, p. 216). It accentuates hierarchy and submission, as well as oppresses independence and self-autonomy based on the principle of absolute observance, obedience and subjugation of individual needs are not emphasized as much as before, but affective consideration and care of parents are still affirmed among Chinese people (Yeh & Bedford, 2003; Yue & Ng, 1999).

The dual model has only been applied to date to examine the relationships of filial piety to indices of intrapersonal differences and the frequencies and treatment of family conflict (Yeh & Bedford, 2003, 2004). This study will first probe its potential effectiveness in the realm of career decision making. More specific, we will examine the relationships between two types of filial piety and career commitment process and career decision self-efficacy.

Researchers (Yeh, 1997a; 2003; Yeh & Bedford, 2003) propose that the effects of reciprocal filial piety on family relationships and individual development would be consistent with the positive evidence documented by the past research. For instance, higher filial piety yields better intergenerational relationships, family cohesion and solidarity (Cheung, Lee, & Chan, 1994; Lawrence, Bennett, & Markides, 1992; Roberts & Bengtson, 1990), low frequencies of parent-child conflict (Yeh, 1997b). Higher filial piety makes children feel more obligated to offer financial, physical and emotional support for parents and more responsible for caring for them (Ishii-Kuntz, 1997). Family obligation facilitates adolescents to discuss personal issues with family members, actively involves in positive social interaction and seeks help with peers, and promotes their educational aspirations and expectations (Fuligni, Tseng, & Lam, 1999). Specifically, both relational factors (i.e. respect for and a sense of indebtedness to parents) and parental factors (i.e., expectation, pressure, and social support) were significantly predictive of Korean adolescents' self-efficacy, achievement motivation, and studying time, which in turn promoted their academic achievement (Kim & Park, 2006). More direct evidence showed that reciprocal filial piety was positively related to openness and negatively to conventionalism (Yeh & Bedford, 2003). Therefore, we expect that reciprocal filial piety is positively related to career decision self-efficacy and progress in career commitment but negatively to the tendency to foreclose.

One the other hand, authoritarian filial piety would be likely to play a harmful role in interpersonal relationships and individual development as also partially evidenced by previous research. For example, parents' strong filial piety attitudes would produce children's *rigidity* and impede their *cognitive complexity* (Boey, 1976). Children who trained according to traditional filial attitudes were inclined to show cognitive conservatism in performing various problem-solving tasks, that is, they were more likely to adopt a passive uncritical and uncreative orientation toward learning, to hold fatalistic, superstitious, and stereotypic beliefs, and to be authoritarian, dogmatic, and conformist (Ho, 1994). They were also more likely to

engage in superstitious practices, such as consulting an almanac or fortunetellers in making decisions (Ho, 1990). A sense of obedience to one's parents would force children to give up self-fulfillment, suppress self-expression in terms of impulse control. In a recent qualitative research to explore how child-parent relationships affect career and educational choices with a sample of senior undergraduates from the PRC, Deutsch (2004) found that parents' education backgrounds, individual history, family financial situation and social network (guanxi) strongly influenced children's career and further education options and geographic location. Especially, when parents disagreed with children's career or educational decisions, the majority of participants tended to follow their parents' wishes in terms of parents' attitudes and needs rather than fulfill personal occupation preferences. Deutsch argued that less obedience to parents' expectations would facilitate Mainland Chinese college students to more commit to their career goals. More direct evidence has further shown that authoritarian filial piety was positively associated with Neuroticism, Authoritarian Aggression and Conventionalism, and Particularism, negatively with Openness and Extroversion (Yeh & Bedford, 2003). Therefore, we hypothesized that authoritarian filial piety is negatively related to career decision self-efficacy and progress in career commitment but positively to the tendency to foreclose.

# Method

#### Participants

Convenience sampling was used to recruit 374 postgraduate students were recruited from a large comprehensive university in Beijing. Of these, 158 (42.2%) were male and 216 (57.8%) female, with a mean age of 23.9 (SD=1.16).

#### Measures

*Filial Piety Scale* (FPS; Yeh & Bedford, 2003). This 16-item scale assesses both reciprocal and authoritarian filial piety. The instructions and item statements were slightly modified to be suitable for Mainland postgraduates. Participants are asked to indicate their agreement with the items on a 6-point scale from Extremely Unimportant (0) to Extremely Important (5). Examples of items measuring reciprocal filial piety include, "Hurry home upon the death of a parent, regardless of how far away you live" and "Be grateful to parents for raising you". Authoritarian items include, "Live with parents even after marriage" and "Compliment your parents when needed to save their face". Yeh and Bedford (2003) reported the Cronbach alpha's coefficient of .90 for the reciprocal FPS (8 items) and .79 for the authoritarian scale (8 items).

*Commitment to Career Choices Scale* (CCCS; Blustein et al., 1989). The Vocational Exploration and Commitment (VEC) subscale of the Commitment Career Choices Scale taps a dimension that ranges from an uncommitted, exploratory phase to a highly committed phase of career exploration, and uses 19 items to assess one's progress in attaining commitment to career choices. Lower scores on the VEC represent a greater commitment in one's career

exploration and choice process. Reliability of the VEC scale has been reported with Cronbach's alpha coefficients ranging from .92 (for a sample of 565 US college students) to .91 (for a sample of 571 US college students). In terms of validity, Blustein et al. (1989) reported a correlation of -.72 between VEC and the Decision Making Task: Occupations scale of the Assessment of Career Decision Making (Harren, 1984) in a sample of 117 college students. Betz and Serling (1993) reported that VEC correlated .62 with the Fear of Commitment Scale (Serling, 1988) and .83 with the Career Indecision from Career Decision Status (Osipow, Carney, Winer, Yanico, & Koschier, 1976) in a sample of 138 US undergraduates.

The Tendency to Foreclose subscale (TTFS) is a 9-item scale to assess how one commits to career choices. Participants are asked to indicate their agreement with the items on a 6-point scale from Never true about me (1) to Always true about me (6) on both VEC and TTFS subscales. Low scores on the TTFS indicate being less foreclosed in commitment to a career choice. The TTFS has been found to have Crobach's alpha coefficients of .82 (for a sample of 565 college students) and .78 (for a sample of 571 college students). In terms of validity, Blustein et al. reported that the TTFS was related to a measure of foreclosure (Bennion & Adams, 1986) as expected.

*Career Decision Self-efficacy –Short Form* (CDSE-SF; Betz, Klein, & Taylor, 1996). This 25-item scale measures the degree of one's belief concerning his/ her ability to successfully complete tasks necessary to make career decisions on five hypothesized dimensions, including self-appraisal, gathering occupational information, goal selection, making plans to implement the decision, and problem solving. Betz et al. (1996) reported subscale reliabilities ranging from .73 (Self- Appraisal) to .83 (Goal Selection) and a Cronbach alpha coefficient of .94 for the total score. In terms of validity, Betz et al. (1996) reported that CDSE-SF correlated significantly with the Certainty and Indecision subscales of the Career Decision Scale (Osipow, Carney, Winer, Yanico, & Koschier, 1976), and with the Vocational Identity subscale of the My Vocational Situation Scale (Holland, Daiger, & Power, 1980) in a sample of 184 undergraduates. Betz, Hammond, and Multon (2005) recently provided extensive evidence to support the reliabilities and validities of the CDSE-SF.

Two items were excluded in the current study because they were not suitable for postgraduates in our sample. These two items are "Find information about graduate or professional schools" and "Select one major from a list of potential majors you are considering". The modified version of CDSE-SF consists of 23 items. The participants indicated how confident they feel about each statement along a 5-point Likert scale from No Confidence At All (1) to Complete Confidence (5).

**Demographic questionnaire.** A demographic questionnaire was used to collect background information regarding the participants' gender, age, academic major and year in study.

#### Procedure

In order to guarantee the literal accuracy of instruments, Commitment to Career Choices and Career Decision Self-Efficacy-Short Form were translated from English to Chinese by the researcher and back-translated from Chinese into English by a Chinese PhD. student with an English major. A native English PhD. student compared the back-translated version with the original English version, leading by a minor revision of the Chinese version. The Filial Piety Scale was not included in this procedure as it was developed in Chinese.

The data were obtained in first semester of the 2006-2007 academic year. Participants were solicited in public classes and by dormitory fliers and personal contact. The questionnaires were distributed and administered in classrooms or dorms. Those who endorsed the informed consent forms spent about 35 minutes completing the questionnaires and received a small gift for their contribution.

# RESULTS

Means, SDs and Cronbach's  $\alpha$  coefficients for all measures are presented in Table 1. In order to examine the potential impacts of gender, a one-way multivariate analysis of variance (MANOVA) was performed using the two filial piety factors (i.e., RFP and AFP) as dependent variables, and gender as an independent variable. MANOVA indicated a significant gender main effect (Wilks' Lambda= .97,  $F(2, 371) = 5.32 \ p < .01$ ). Follow-up univariate analyses further showed the significant main effect of gender on AFP [ $F(1, 372) = 5.55, \ p < .05$ ], indicating that males were more likely to held stronger beliefs of authoritarian filial piety than females. However, there were no significant differences in reciprocal filial piety between males and females. Therefore, we conducted the following analyses in terms of gender.

	Male ( <i>n</i> =158)		Female ( <i>n</i> =216)			Total (N=374)			
	М	SD	α	М	SD	α	М	SD	α
1. RFP	4.41	.47	.79	4.48	.40	.76	4.45	.43	.78
2. AFP	2.34	.84	.82	2.14	.82	.84	2.22	.83	.83
3. CDSE	3.51	.52	.91	3.37	.48	.91	3.43	.50	.91
4. VEC	3.36	.60	.84	3.47	.54	.83	3.43	.57	.84
5. TTF	3.14	.70	.73	2.94	.66	.79	3.03	.68	.77

Table 1. Means , Sds and Cronbach's  $\alpha$  Coefficients for all Variables

*Note:* **RFP**= reciprocal filial piety, **AFP**= authoritarian filial piety. **CDSE**= career decision selfefficacy. **VEC**= vocational exploration and commitment, **TTF**= tendency to foreclose.

Zero-order correlations were computed between two types of filial piety and CDSE, VEC and TTF by gender separately (see Table 2). For both men and women, reciprocal filial piety was significantly and moderately correlated with career decision self-efficacy (men, r=.37, p<.001; women, r=.29, p<.001), inferring that higher levels of reciprocal filial piety are associated with stronger perceived career decision self-efficacy across gender. Moreover, reciprocal filial piety was modestly but significantly related to vocational exploration and commitment for men(r=-.18, p<.05), but not for women (r=-.10, p>.05). These results

indicated that there may be gender differences in the correlations between reciprocal filial piety and vocational exploration and commitment. Authoritarian filial piety was consistently found to be significantly and moderately associated with the tendency to foreclose for both men (r=.27, p<.001) and women (r=.37, p<.001), indicating that stronger authoritarian filial piety was related to a greater tendency to foreclose regardless of gender.

Contrary to our hypotheses, for both genders, reciprocal filial piety was not significantly related to the tendency to foreclose, and authoritarian filial piety was not significantly related to either career decision self-efficacy or vocational exploration and commitment.

	1	2	3	4	5
1. RFP		.22***	.29***	10	.05
2. AFP	.19**		02	.07	.37***
3. CDSE	.37***	01		37***	.05
4. VEC	18*	03	34***		18**
5. TTF	02	.27***	.03	18*	

Table 2. Zero-Order Correlations for all Variables by Gender

Note: RFP= reciprocal filial piety, AFP= authoritarian filial piety. CDSE= career decision selfefficacy, VEC= vocational exploration and commitment, and TTF= tendency to foreclose. Zeroorder correlations for men are presented in the below diagonal, and those for women in the above. \*p<.05, \*\*p<.01, \*\*\*p<.001 (1-tailed).</p>

## CONCLUSIONS

This study examined the relationships between two types of filial piety and adaptive career decision making (i.e., career commitment and decision self-efficacy) from an indigenous perspective. The findings partially support our hypotheses that reciprocal filial piety is likely to enhance career decision self-efficacy and progress in career commitment, while authoritarian filial piety is positively associated with the tendency to foreclose. It is very important for career theorists and counseling practitioners to understand the adaptive process of career decision making among Chinese postgraduates.

For hypothesis 1, the anticipated significant and positive relationship between reciprocal filial piety and career decision self-efficacy was moderate for both men and women, indicating that those who reported stronger reciprocal filial piety perceived higher levels of self-efficacy for their career decision-making. This finding is consistent with previous research that indicated filial piety may benefit the specific-domain self-efficacy, such as academic self-efficacy in Kim and Park's(2006) research and career decision self-efficacy in this study. It is possible that respect for and a sense of indebtedness to parents, and duties to financially and emotionally support the whole family among those with collectivist values are likely to foster more confidence in personal abilities to deal with their own career decisions. In addition, less family conflict and more expressive freedom have been documented to be associated with higher career decision self-efficacy (Hargrove, Creagh, & Burgess, 2002). Since reciprocal filial piety emphasizes children's initiative to maintain the harmonious and

equal intergenerational relationships, the significant correlations of reciprocal filial piety with career decision self-efficacy would not surprise us.

The relation between reciprocal filial piety and vocational exploration and commitment was mixed. For men, higher levels of reciprocal filial piety were significantly and modestly associated with greater progress in career commitment, but not significantly for women. Such findings partially shed light on past research that family obligation (i.e., a combination of respect and care for parents) is likely to promote educational aspiration/ expectations (Fuligni, Tseng, & Lam, 1999). Here, we further provided the empirical evidence that, at least for men in this study, respect and care for parents tapped in reciprocal filial piety also benefit for young people to make progress in attaining career commitment. For women, reciprocal filial piety may only exert an indirect effect on progress in career commitment through career decision self-efficacy. However, such hypothesis merits further examination.

Contrary to our hypothesis, reciprocal filial piety was not observed to be significantly related to the tendency to foreclose across gender, although past research reported that reciprocal filial piety was positively related to openness and negatively to conventionalism (Yeh & Bedford, 2003). The confounding results are warranted to investigate more.

With respect to hypothesis set 2, as we had expected, for both men and women, those who endorsed higher authoritarian filial piety would likely show stronger tendency to foreclose. One possibility is authoritarian filial piety holders would likely submit to the authority and comply for parents' wishes. To avoid the intergenerational conflicts and parents' pressure in career choices, they tend to commit to a career choice as soon as possible out of obedience or suppression based on the requirements of social roles in the Collectivism culture, no matter whether they have clarified their own occupational preferences. This finding seems to partially support Blustein et al.'s (1989) findings that lack of autonomy was correlated with the tendency to foreclose. Another explanation is higher authoritarian filial piety has been identified to be related to higher Neuroticism personality (Yeh & Bedford, 2003), and less open-mindedness and future-orientation (Yang, 1996), which tended to lead authoritarian filial piety holders to adopt dogmatic or dualistic attitudes to approach the tasks of committing to career choices.

Contrary to our expectations, authoritarian filial piety was not significantly negatively related to career decision self-efficacy and progress in career commitment. Although previous research proposed that less bond to obey parents' expectations would facilitate Mainland Chinese college students to more commit to their career goals(Deutsch, 2004), or authoritarian filial piety holders tended to make decisions dependent on superstitious prediction, our study failed to identify that authoritarian filial piety would impede either progress in career commitment or career decision self-efficacy. The absence of significant relations may be explained by that some students would have a firm career commitment although they comply with parents' suggestions without in-depth exploration by themselves(Marcia, 1980). Given that career choice is conceived of as a means to represent interdependent self close to family (Hardin, Leong, & Osipow, 2001) or a reciprocal alternative for both self and family (Leong & Serafica, 1995) in a collectivist culture, Chinese students may have partly internalized parents' values as their own preferences, especially for those who would likely honor the family and keep the familial lineage (as partially reflected by authoritarian filial piety).

The findings of this study are very meaningful to career counseling practice. It helps us to understand the process of career decision making from a perspective of Chinese culture, above and beyond the previous angles of intrapersonal differences in cognition and personality, or adolescent development. Concerning the important effects of filial piety on career decision making, counselors are suggested to take the child-parent relationships stemming from filial piety into consideration when they design interventions for and counsel individuals endorsed the collectivist values. For example, elements regarding the enhancement of clients' gratitude and responsibilities to parents are suggested to add into career interventions in order to enhance young people's career decision self-efficacy and career commitment. It is also suggested that counselors should be sensitive to examine the antecedents of one's tendency to foreclose. If he or she commits to a premature choice purely caused by obedience to parents' authority, which in turn would reduce anxiety and pressure resulting from intergenerational conflicts, counselors had better remind the client to aware the negative possibilities of a tendency to foreclose for career success and satisfaction from a long-term career development perspective. If the tendency to foreclose is associated to one's cognitive limitations (i.e., dogmatic or dualistic thinking styles), counselors would better provide more career exposure to broaden clients' horizon.

# REFERENCES

- Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.
- Bennion, L. D., & Adams, G. R. (1986). A Revision of the Extended Version of the Objective Measure of Ego Identity Status: An identity instrument for use with late adolescents. *Journal of Adolescence Research*, 1, 183-198.
- Betz, N. E., Hammond, M. S., & Multon, K. D. (2005). Reliability and Validity of Five-Level Response Continua for the Career Decision Self-Efficacy Scale. *Journal of Career Assessment*, 13(2), 131-149.
- Betz, N. E., Klein, K. L., & Taylor, K. M. (1996). Evaluation of a short form of the Career Decision-Making Self-Efficacy Scale. *Journal of Career Assessment*, 4(1), 47-57.
- Betz, N. E., & Luzzo, D. A. (1996). Career assessment and the Career Decision-Making Self-Efficacy Scale. Journal of Career Assessment Special Issue: Assessing career beliefs, attitudes, and values, 4(4), 413-428.
- Betz, N. E., & Serling, D. A. (1993). Construct validity of fear of commitment as an indicator of career indecisiveness. *Journal of Career Assessment*, 1, 21-34.
- Blustein, D. L. (1997). A context-rich perspective of career exploration across the life roles. *The Career Development Quarterly, 45*, 260-274.
- Blustein, D. L. (2004). Moving from the inside out: Further explorations of the family of origin/career development linkage. *The Counseling Psychologist, 32*, 603-611.
- Blustein, D. L., Ellis, M., & Devenis, L. (1989). The development and validation of a twodimensional model of the commitment to career choice process [Monograph]. *Journal of Vocational Behavior*, 35, 342-378.
- Blustein, D. L., Walbridge, M. M., Friedlander, M. L., & Palladino, D. E. (1991). Contributions of psychological separation and parental attachment to the career development process. *Journal of Counseling Psychology*, 38(1), 39-50.

- Boey, K. W. (1976). *Rigidity and cognitive complexity: An empirical investigation in the interpersonal, physical, and numeric domains under task-oriented and ego-oriented conditions.* Unpublished doctoral dissertation, University of Hong Kong.
- Cheung, C. K., Lee, J. J., & Chan, C. M. (1994). Explicating filial piety in relation to family cohesion. *Journal of Social Behavior and Personality*, *9*, 565-580.
- Constantine, M. G., & Flores, L. Y. (2006). Psychological distress, perceived family conflict and career development issues in college students of color. *Journal of Career Assessment*, 14, 354-369.
- Creed, P. A., & Wong, O. Y. (2006). Reliability and validity of a Chinese version of the career decision-making difficulties questionnaire. *International Journal for Educational* and Vocational Guidance, 6, 47-63.
- Deutsch, F. M. (2004). How parents influence the life plans of graduating Chinese university students. *Journal of Comparative Family Studies*, *35*, 393-421.
- Fuligni, A. J., Tseng, V., & Lam, M. (1999). Attitudes toward family obligations among American adolescents with Asian, Latin American, and European backgrounds. *Child Development*, 70, 1030-1044.
- Gottfredson, L. S. (2002). Gottfredson's theory of circumscription, compromise, and selfcreation. In D. Brown & Associates (Eds.), *Career choice and development* (4 ed., pp. 85-148). San Francisco: Jossey-Bass.
- Hampton, N. Z. (2005). Testing for the structure of the Career Decision Self-Efficacy Scale-Short Form among Chinese college student. *Journal of Career Assessment*, 13, 98-113.
- Hardin, E. E., Leong, F. T. L., & Osipow, S. H. (2001). Cultural relativity in the conceptualization of career maturity. *Journal of Vocational Behavior, 58*, 36-52.
- Hargrove, B. K., Creagh, M. G., & Burgess, B. L. (2002). Family interaction patterns as predictors of vocational identity and career decision-making self-efficacy. *Journal of Vocational Behavior*, 61(2), 185-201.
- Harren, V. A. (1979). A model of career decision making for college students. *Journal of Vocational Behavior*, 14, 119-133.
- Harren, V. A. (1984). Assessment of career decision making. Los Angeles: Western Psychological Services.
- Ho, D. Y. F. (1990). Chinese values and behavior: A psychological study: Unpublished manuscript, University of Hong Kong.
- Ho, D. Y. F. (1994). Filial piety, authoritarian moralism, and cognitive conservatism in Chinese societies. *Genetic, Social, and General Psychology Monographs, 120*, 347-365.
- Ho, D. Y. F. (1996). Filial piety and its psychological consequences. In M. H. Bond (Ed.), *The handbook of Chinese psychology* (pp. 155-165). Hong Kong: Oxford University Press.
- Holland, J. L., Daiger, D. C., & Power, P. G. (1980). My Vocational Situation. Palo Alto, CA: Consulting Psychologists' Press.
- Ishii-Kuntz, M. (1997). Intergenerational relationships among Chinese, Japanese, and Korean Americans. *Family Relations*, *46*, 23-32.
- Jordaan, J. P. (1963). Exploratory behavior: The formation of self and occupational concepts. In D. E. Super (Ed.), *Career development: Self-concept theory* (pp. 42-78). New York: College Entrance Examination Board.
- Jordaan, J. P., & Heyde, M. B. (1979). *Vocational maturity during the high school years*. New York: Teachers College Press.

- Kim, U., & Park, Y. (2006). Indigenous psychological analysis of academic achievement in Korea: The influence of self-efficacy, parents, and culture. *International Journal of Psychology*, 41, 287-292.
- Lawrence, R. H., Bennett, J. M., & Markides, K. S. (1992). Perceived intergenerational solidarity and psychological distress among older Mexican-Americans. *Journal of Gerontology*, 47, S55-S65.
- Leal-Muniz, V., & Constantine, M. G. (2005). Predictors of the career commitment process in Mexican American college students. *Journal of Career Assessment*, 13(2), 204-215.
- Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance [Monograph]. *Journal of Vocational Behavior*, 45, 79-122.
- Lent, R. W., Brown, S. D., & Hackett, G. (2000). Contextual supports and barriers to career choice: A social cognitive analysis. *Journal of Counseling Psychology*, 47(1), 36-49.
- Lent, R. W., Brown, S. D., Schmidt, J., Brenner, B., Lyons, H., & Treistman, D. (2003). Relation of contextual supports and barriers to choice behavior in engineering majors: Test of alternative social cognitive models. *Journal of Counseling Psychology*, 50(4), 458-465.
- Leong, F. T. L., & Brown, M. (1995). Theoretical issues in cross-cultural career development: Cultural validity and cultural specificity. In W. B. Walsh, & Osipow,S.H., (Ed.), *Handbook of vocational psychology* (2nd ed., pp. 143-180). Hillsdale, NJ:: Erlbaus.
- Leong, F. T. L., & Chou, E. L. (1994). The role of ethnic identity and acculturation in the vocational behavior of Asian Americans: An integrative review. *Journal of Vocational Behavior*, 44, 155-172.
- Leong, F. T. L., & Leung, S. A. (1994). Career assessment with Asian Americans. Journal of Career Assessment, 2, 240-257.
- Leong, F. T. L., & Serafica, F. C. (1995). Career development of Asian Americans: A research area in need of a good theory. In F. T. L. Leong (Ed.), *Career development and* vocational behavior of racial and ethnic minorities (pp. 67-102). Hillsdale, NJ: Lawrence Erlbaum.
- Ma, P. W., & Yeh, C. J. (2005). Factors influencing the career decision status of Chinese American youths. *The Career Development Quarterly*, 53, 337-347.
- Marcia, J. E. (1966). Development and validation of ego-identity. *Journal of Personality and Social Psychology*, 3(5), 551-558.
- Marcia, J. E. (1980). Identity in adolescence. In J. Adelson (Ed.), Handbook of adolescent psychology (pp. 159-187). New York: John Wiley.
- Moy, S. (1992). A culturally sensitive, psychoeducational model for understanding and treating Asian-American clients. *Journal of Psychology and Christianity*, 11, 358-369.
- Osipow, S. H., Carney, C. G., Winer, J. L., Yanico, B., & Koschier, M. (1976). *The Career Decision Scale* (3rd rev. ed.). Columbus, OH: Marahon Consulting & Press and (1987) Odessa, FL: Psychological Assessment Resources.
- Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking Individualism and Collectivism : Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*, 3-72.
- Roberts, R. E. L., & Bengtson, V. L. (1990). Is intergenerational solidarity a unidimensional construct? A second test of a formal model. *Journal of Gerontology*, 45, S12-S20.

- Serling, D. A. (1988). Development of a fear of commitment inventory: Unpublished master's thesis, Ohio State University, Columbus.
- Super, D. E. (1957). *The psychology of careers*. New York: Harper & Row.
- Tang, M., Fouad, N. A., & Smith, P. L. (1999). Asian Americans' career choice: A path model to examine factors influencing their career choices. *Journal of Vocational Behavior, 54*, 142-157.
- Taylor, K. M., & Betz, N. E. (1983). Applications of self-efficacy theory to the understanding and treatment of career indecision. *Journal of Vocational Behavior*, 22(1), 63-81.
- Vondracek, F. W., Lerner, R. M., & Schulenberg, J. E. (1986a). Career development: A life span approach. Hillsdale, NJ: Lawrence Erlbaum.
- Vondracek, F. W., Lerner, R. M., & Schulenberg, J. E. (1986b). The influences of the family of origin on career development: A review and analysis. Hillsdale, NJ: Lawrence Erlbaum.
- Whiston, S. C., & Keller, B. K. (2004). The influences of the family of origin on career development. Society of Counseling Psychology, 32, 493-568.
- Yang, K. S. (1996). The psychological transformation of the Chinese people as a result of societal modernization. In M. H. Bond (Ed.), *The handbook of Chinese psychology*. Hong Kong: Oxford University Press.
- Yeh, K. H. (1997a). Changes in the Taiwanese people's concept of filial piety. In L. Y. Cheng, Y. H. Lu & F. C. Wang (Eds.), *Taiwanese Society in the 1990s* (pp. 171-214): Institute of Sociology, Academia Sinica, Taipei (in Chinese).
- Yeh, K. H. (1997b). Parent-child conflicts and their solution types: Discussion from the viewpoint of flial piety [in Chinese]. Bulletin of the Institute of Ethnology, Academia Sinica, 82, 65-114.
- Yeh, K. H. (2003). The beneficial and harmful effects of filial piety: An integrative analysis. In K. S. Yang, P. B. Hwang & P. Daibo (Eds.), Asian Social Psychology: Conceptual and Empirical Contributions: Greenwood Publishing Group, Inc., Connecticut.
- Yeh, K. H., & Bedford, O. (2003). A test of the Dual Filial Piety model. Asian Journal of Social Psychology, 6, 215-228.
- Yeh, K. H., & Bedford, O. (2004). Filial belief and parent-child conflict. International Journal of Psychology, 39, 132-144.
- Ying, Y. W., Coombs, M., & Lee, P. E. (1999). Family intergenerational relationship of Asian American adolescents. *Cultural Diversity and Ethnic Minority Psychology*, 5, 350-363.
- Yue, X., & Ng, S. H. (1999). Filial obligations and expectations in China: Current view from young and old people in Beijing. *Asian Journal of Social Psychology*, 2, 215-226.
- Zhang, W., Hu, X., & Pope, M. (2002). The evolution of career guidance and counseling in the People's Republic of China. *Career Development Quarterly*, 50, 226-236.

Chapter 24

# **REPRODUCTIVE AGENCY AMONG LOW-INCOME WOMEN IN BRAZIL<sup>\*</sup>**

# Gina Hunter de Bessa

Illinois State University, Normal, Illinois, USA

# ABSTRACT

In the United States, discussions about reproduction, contraception, and abortion often are framed within the language of "choice." Feminists, however, have long argued that the dominant ideology of choice in family planning obfuscates the larger social and institutional forces which structure and constrain the choices people can make. Many scholarly debates in social sciences, too, have shifted between theories of active and passive agency, that is, between hypotheses that posit a rational actor making calculated fertility choices (often in "modern" societies) versus those that see people passively following norms and conventions (usually attributed to so-called "traditional" societies).

In this article, I use data from ethnographic fieldwork in one low-income, urban neighborhood in Brazil to examine poor women's reproductive decision-making. Through consideration of how these women understand their reproductive options, I illustrate the social and cultural forces in the home and community that structure reproductive "choices." In particular, I highlight women's decisions to permanently end their fertility through surgical sterilization. I argue that the decision to undergo sterilization is conflicted and constrained, but is an expression of these women's active agency.

The Brazilian context is especially interesting for examining issues of reproductive choice in relation to female surgical sterilization. Brazil registers some of the highest rates of female sterilization in the world. The reasons for this are complex, but cannot be attributed directly to family planning programs or policies, as is the case in some other developing countries. While most instances of sterilization in Brazil are voluntary, Brazilian women's reproductive choices are clearly subject to myriad structural and cultural constraints. An examination of the factors women take into account when

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

making reproductive decisions, such as the decision to permanently end child-bearing, point to the centrality of social context and social networks in their lives.

### INTRODUCTION

In this article, I use data from ethnographic fieldwork in one low-income, urban neighborhood in Brazil to examine poor women's reproductive agency. In particular, I illustrate the social and cultural forces in the home and community that structure women's reproductive "choices" and I highlight women's decisions to permanently end their fertility through surgical sterilization. I argue that the decision to undergo sterilization is conflicted and constrained, but is an expression of these women's active agency, that is, their attempt to shape their lives. I conclude with the story of one woman who I will call "Tania." Her story illustrates the social context of women's decision-making and the factors that shape women's perceptions of their reproductive options.

The low-income Brazilian women with whom I worked considered it quite natural to talk about sex, childbearing, childrearing, and "women's health" (and expression associated with health center initiatives), but "reproductive decision-making" is not a category or an expression or that they would use, and in fact, it implies a similarity among distinct life events that does not necessarily exist. I use "reproductive decision-making" loosely in this article to pull together disparate points and concerns in a woman's life related to her fertility.

Situated within medical anthropological and feminist discussions of reproductive health and rights, my study focused on the factors that compel poor women to undergo surgical sterilization. While this required me to examine the various issues women took into account when managing their reproductive lives, I did not focus on decision-making per se. That is, I did not seek to come up with a model of the decision-making process. There are, however, numerous cognitive studies of decision-making within medical anthropology that do aim to uncover testable models of how people proceed when faced with alternative actions (for a brief discussion of these studies see Garro, 1998).

In what follows, I will not attempt a general review and critique decision-making studies. Rather, my point is to show that a focus on "decision-making" carries a number of assumptions that poorly capture the various struggles and constraints faced by Brazilian women as they manage their reproductive lives. I draw upon literature within the anthropology of reproduction and fertility for conceptual tools that better convey women's reproductive agency. While the discussion stays close the issues of reproduction relevant to the Brazilian case at hand, the critiques presented could be extended to other topical areas and to decision-making studies in general.

# **CONCEPTUALIZING REPRODUCTIVE AGENCY**

From an anthropological perspective, the principle limitation of a focus on decisionmaking is the primacy assigned to rational, individual actors and their "choices." This focus belies the social structures and historical contexts that constrain people's alternatives and shape their dispositions in specific ways. A focus on the individual is culturally-specific and coincides the tendency toward increasing individualization in late modern societies (Lindbladh & Lyttkens, 2002). In many cultural contexts, particular social groups, such as the family or couple, rather than individuals, may be the relevant decision-making unit, especially in issues of reproduction and fertility. So too, the assumption of a rational, "valuemaximizing" actor is culturally grounded in an ideologically-specific context. Even at the theoretical level, as some scholars have noted, the maximization required in some microeconomic models of the decision-making process is "impossibly complex, timeconsuming and costly" (Leibenstein cited in Carter, 1995, p. 56) and beyond an ordinary person's abilities. Finally, this "economic" model of decision-making also posits a utilitarian notion of rationality which disregards the role of emotion. This assumption is particularly problematic for decisions in the realm of reproduction and fertility. Like Unnithan-Kumar (2001, p. 31), I understand "women's reproductive agency (their desires and motivations) as linked to emotions...[for] the stimulus to act in particular ways often stems from emotional involvement with those who surround us and often, the decision to carry out a particular action, or even the thought that a particular course of action is the most suitable, is based on a negotiation between one's own and others' feelings on the matter." Rather than assuming actions and options a priori, social scientists need to document and account for emotional content and social context of agency. They must question when, how and under which circumstances people conceive of their options and ask who makes those decisions.

Feminist scholars have long argued that the dominant ideology of choice in fertility studies and "family planning" obfuscates the larger social and institutional forces which structure and constrain the choices people can make. Rosiland Petchesky (1981), for example, examined rising rates of female sterilization in the U.S. She argued that while women might voluntarily decide to undergo sterilization, they are motivated by a series of structural inequities that they as individuals are unable to change, such as women's disadvantageous position in the family and labor force. Specifically, Petchesky pointed to the concentration of women in low-paid jobs, the lack of publicly funded social support, and the burden of primary responsibility for birth control and childcare. Within a patriarchal sexual culture, she argues, "where men may be hostile to birth control or unwilling to take any responsibility themselves, women may prefer the method that seems less conspicuous, surest and least dependent on male cooperation" (1981, p. 66). Women's subjective perceptions of contraceptive choice therefore must be understood against the backdrop of state population policies, ideological and material interests of private and public groups in promoting a limited number of birth control methods, the structure of capitalist medical practices that profit from costly surgical intervention, and the "ideological and practical focus on women's bodies as the appropriate, and accessible, terrain of intervention" (1981, p. 75). Women are not victims of imperialistic or patriarchal forces; but nor are they rational decision-makers able to freely choose among a number of alternatives.

Anthony Carter (1995) faults many demographic and fertility studies for their separation of culture and agency (that is, theories of passive agency vs. active agency). For example, classic demographic transition theory distinguishes between controlled fertility populations (those that practice deliberate control of fertility through contraception and abortion) and natural fertility populations in which people do not control fertility except through social customs, such as marriage restrictions. The first hypothesizes a rational actor making calculated fertility choices (often in "modern" societies) whereas the second posits people passively following social norms and conventions (usually attributed to pre-industrial, "traditional" societies). Carter points out that the latter theory of human conduct in so-called traditional societies does not hold up to empirical evidence as anthropologists have long documented the active ways people in "non-contracepting" populations directly intervene in family formation. The former theory suffers too in light of ethnographic studies that show that people fail to act according to narrowly rational interests and make a variety of kinds of decisions.

To escape this dilemma, Carter argues for a "practice theory" approach that highlights that cultural and social contexts and discourses through which people make sense of their options and actions. Such a framework offers the advantages of avoiding dichotomies of belief and behavior, individual vs society, rationality vs emotion. Carter points to Anthony Gidden's theory of structuration (1979) for an understanding of human agency "not as a sequence of discrete acts of choice and planning, but rather as the reflexive monitoring and rationalization of a continuous flow of conduct" (Carter 1995, p. 61). Within these "flows of conduct," cultural principles are viewed as provisional guides that come to life as they are brought into activity. Thus we might examine where decision-making points emerge within the flow of conduct and be become aware of a variety of decision-making processes, such as routinized non-decisions "planning decisions" made in advance of activities, and "on-thespot" decisions made in the context of activities. Drawing on Jean Lave's work (1988), Carter conceives decision-making, not as a mental process, but as an "activity-in-setting" in which settings shape activities (rather than just providing a backdrop) and in which the material context of an activity, existing social relations, objects and technology, and ideas about proper conduct may all come into play as structuring resources for a person acting within that setting.

These insights can help us conceptualize low-income Brazilian women's reproductive actions as situated within a variety of contexts. Women's reproductive agency is not best envisioned as a series of once and for decisions, but rather as the ongoing monitoring and rationalization of options and actions as they emerge in particular settings which include other actors, such as family and health care providers.

# ETHNOGRAPHY OF REPRODUCTION: INTERVIEWING WOMEN IN BRAZIL

Data for this paper come from ethnographic fieldwork I conducted in 1996-1998 (with a follow-up study in 2002) in Taquaril, a low-income housing settlement located on the urban periphery of Belo Horizonte, Brazil's third largest city, located in the Southeast region of the country. Although the Taquaril area has only been occupied since the mid 1980s, the population exceeds 40,000 residents. The neighborhood is divided in thirteen sectors. The sectors closer to the edge of the city have benefited from urbanization efforts. In these areas, brick houses along line the asphalt streets and cement stairways help residents navigate the steep hillsides. More remote (and recently occupied) sectors have only dirt paths and small wood and brick shanty houses. My fieldwork was largely confined to the residents who lived along a three dirt *becos*, or small alleyways, in a "middle" sector which has experienced some urban development. Along these streets, some of the residents live in sturdy brick four-room houses on walled lots while others, like Tania, live in precarious brick shacks on marginal

pieces of land where electricity and plumbing are jerry rigged extensions from neighbors' services.

Through ethnographic interviews with twenty women in the neighborhood who were sterilized or seeking sterilization, I explored issues of contraceptive use, lay notions of reproductive health and illness, and women's trajectories toward surgical sterilization. Participant-observation in the neighborhood allowed me to witness women's everyday lives and made me privy to community members' informal conversations, the exchange of the advice between women, and neighborhood gossip.

The women I interviewed in Taquaril reflect the diversity of the community. They range in age from 18 to 50 and had completed between two and eight years of formal education. Two of the women interviewed were full-time housewives; the others held jobs outside the home as domestic and service workers. The average household income in 1998 was about R\$260 per month (two "minimum wages<sup>1</sup>" or about U.S. \$215), significantly below what it actually cost to support a family. Sixteen of the twenty women were married or cohabitating, two were single and two widowed.

Understanding the personal narratives and reproductive histories of the women I interviewed requires situating their lives in a variety of broader and narrower contextual lenses. I therefore devote much of the article to introducing the broader social, political, and historical contexts which have shaped demographic trends and women's reproductive lives in Brazil. I conclude with Tania's story to illustrate how theses forces structure and impinge upon her reproductive options.

# SITUATING REPRODUCTION: DEMOGRAPHIC TRENDS

Although levels of fertility have been declining since the 1960s throughout Latin America, Brazil is noted for the rapid pace of decline. In a single generation, from 1970 to 1991, average fertility rates in Brazil declined from 6 to 2.5 children per woman (Berquó, 1993, p. 367). This precipitous decline in fertility was not the result of an official population policy or family planning program, as has been the case in some other developing countries. Until the 1970s, the Brazilian government's position was officially pronatalist and few fertility control services where provided within the public healthcare system, reflecting in part the influence of the Catholic Church. Family planning agencies (international and domestic) and individual health care professionals, however, played important roles in providing birth control and promoting concern about population growth (Martine, 1996).

The acceleration of fertility decline in the 1960s coincided with the coming to power of a right-wing military regime (1964-1988) and widespread social change. The period was marked by rapid industrialization; miraculous economic growth followed by economic crisis; and increasing social exclusion, especially in rapidly expanding, precarious urban shantytowns. One of the most influential interpretations of Brazil's vertiginous fertility decline (Faria, 1989) posits that it was largely the unintended consequence of specific development policies implemented during the military regime. These policies included the extension of consumer credit to low-income urban populations; heavy investment in telecommunications and especially the spread of television; the expansion of hospital-based, curative medicine; and the expansion of social security coverage. Although none of the

policies had population control as the objective, all of them worked against large families by making them unnecessary for social support, or by increasing people's expectations in terms of consumption or gender roles. These policies also contributed to rapid urbanization and thus to Brazil's fertility transition (Martine, 1996).

Despite an increasing "demand" for fertility control, few contraceptive methods were available the majority of the population served by public health services. Abortion, although illegal, was probably responsible for much the early decline (Martine, 1996). Oral contraceptives became increasingly available after the mid-1960s and were sold over the counter in pharmacies (Merrick & Berquó, 1983). Given the absence of contraceptive services in public health system, Brazilian women came to rely on a restricted number of methods that were initially offered by nongovernmental organizations, individual physicians, or the market. The largest proximate determinate of fertility decline is the incidence of female surgical sterilization. By 1996, the year for which the most recent national data is available, 76.7% of married women of reproductive age used some form of contraception; but only two methods, surgical sterilization (40%) and the birth control pill (20.1%) accounted for the majority of contraceptive use. Other modern methods, such as the IUD (1.1%), injections (1.2%), and condoms (4.4%), contributed little to the contraceptive "mix"; traditional methods, such as periodic abstinence (3%) and withdrawal (3.1%) only slightly more so; and rates of vasectomy (2.6%) were insignificant compared to female sterilization (BEMFAM/DHS, 1996, p. 13).

Although fertility decline began earlier in the upper income and urban groups, it has since spread through all social strata and regions. However, important socioeconomic and regional differences exist. The total fertility rates range from 5 children per woman for those women with no formal education to 1.5 for women with 12 or more years of education (BEMFAM/DHS 1996:11). Socio-economically disadvantaged women continue to have relatively higher levels of unwanted fertility and are more likely than women of higher income and education levels to resort to surgical sterilization (Perpétuo & Wajnman, 1998). Regional differences in fertility rates and contraceptive patterns reflect dramatic socioeconomic differences, the differential impact of immigration patterns (Goldani, 2001), differential investments in family planning agents, and local cultural patterns (Caetano, 2000). Sterilization rates are highest in the North and Center-West regions of the country and lowest in the South where the pill is prevalent (BEMFAM/DHS, 1996).

The high incidence of sterilization is problematic not only because it represents a lack of contraceptive options but also due to the circumstances under which most women have been sterilized. Until 1997, voluntary sterilization was prohibited by the Brazilian Medical Ethics Code and therefore was not covered by either public or private health insurance. Sterilization was allowed only for women who presented a medical justification for sterilization (e.g. cases in which a future pregnancy presented a clear threat to a woman's health). Physicians however would routine circumvent legal restrictions by performing a tubal ligation concurrently with another abdominal surgery, most commonly a cesarean childbirth, for which hospital and anesthesia costs were covered (Potter, 1999). Seventy one percent of tubal ligations have been performed conterminously with c-sections (excluding the Northeast region where a different pattern emerged) (Potter, 1999). This arrangement for subverting legal restrictions and covering the costs of sterilization resulted in several negative consequences: a large number of unnecessary c-sections, the absence of counseling and informed consent, unequal access to sterilization, and a perverse market for clandestine

sterilization in which physicians sometimes receive side payments for performing a tubal ligation and politicians arrange sterilization in exchange for votes (Caetano, 2000; Caetano & Potter, 2004).

Under this perverse arrangement, the practice of sterilization and cesarean section became mutually reinforcing procedures (Berquó, 1993). Until the late 1970s, physicians were remunerated more for cesarean than vaginal births. Even after this incentive was eliminated, planned cesarean sections continued to offer the convenience of scheduling. Physicians became accustomed to surgical deliveries and lacked comparable training in vaginal childbirth (Potter, 1999). The rise in cesarean childbirth augmented sterilization rates due to the practice of repeat cesareans—once a woman has had a c-section physicians tend to recommend that all future pregnancies end in cesarean due to the risk of uterine rupture—in conjunction with the advice that a woman not undergo more three cesarean deliveries (Faúndes & Cecatti, 1991: Barros et al., 1991). A third c-section, in other words, provides a medical justification for performing a tubal ligation.

In the political context of re-democratization and the adoption of a new constitution in the late 1980s, a comprehensive women's health program was designed which included an expansion of contraceptive services. Lack of funding and political will however hindered the program's full implementation. As a result, neither sterilization nor cesarean rates diminished from the mid-80s to the mid-90s. In fact, rates of female surgical sterilization rose among married women from 27 to 40 per cent from 1986-1996 (BEMFAM/DHS, 1996, p. 4). Potter (1999) makes the compelling argument that by this point in time Brazil's contraceptive pattern had already become "locked-in" through path dependence and positive reinforcement for both physicians and women. Brazilian obstetricians had become specialized in both cesarean section deliveries and sterilization and their practice was widely accepted within the medical community. The implication is that cesareans and sterilization, a "tried and true" method (Berquó, 1993; Potter, 1999). Potter argues that this "culture of sterilization" sustains entrenched patterns of cesarean and sterilization despite changes in the institutional and legal context that gave rise to them in the first place.

Given the context in which women have undergone sterilization, one might expect to find to find large numbers of women who are dissatisfied or regret the procedure. Indeed levels of regret, 17% in one study (Viera & Ford, 1996), are low considering the widespread use of the procedure. Other studies have found that women are largely satisfied with the procedure (Berquó, 1993; Osis et al., 1999).

In a further effort to curb high rates of sterilization and unnecessary cesarean deliveries, sterilization was legalized in 1997. Law 9.263/96 permits voluntary sterilization for women and men who are 25 years of age or who have at least two living children. The legislation also stipulates a 70 day waiting period, counseling, and spousal consent.<sup>2</sup> Early evaluations of the program found that women still encounter obstacles accessing sterilization, including physicians' misinterpretation of the law, excessive wait time between date of request and date of service, lack of service facilities, and failure to secure spousal consent (Berquó & Cavenaghi, 2002; Luiz & Citeli, 2000). Moreover, incentives for performing unnecessary cesarean deliveries and surreptitious tubal ligations persist because the new regulations outlaw postpartum sterilizations —an aspect of the legislation meant to decouple the procedures (Potter et al., 2003). Renewed committed to family planning in the 2005-2007 National Policy on Sexual and Reproductive Rights aims to expand the supply of reversible

contraceptive services, to increase access to voluntary surgical sterilization, and to offer assisted reproduction services within SUS. More recently, the government has proposed to increase its subsidization of oral contraceptives throughout the country in an effort to curb unwanted pregnancy and illegal abortions. These measures may go some way towards improving women's reproductive health and options.

# **DEBATING THE SIGNIFICANCE OF STERILIZATION IN WOMEN'S LIVES**

Given the demographic patterns and the history of contraceptive services in Brazil, it is easy to see the number of forces at work in framing women's reproductive decisions. In many ways, high rates of sterilization seem to be the result of women's conscious choices and determination to access a service that the state long failed to provide. Many women have taken extraordinary measures to reduce their fertility, yet sterilization also appears to be an option among few alternatives and one that has subjected women to greater medical intervention. In short, scholars have wondered whether sterilization is best interpreted as a last resort taken by women in desperation or whether is signifies an attempt at greater autonomy and control (Osis et al., 1999, p. 523).

Certainly, many women turn to sterilization because they have less success using reversible methods of birth control. Perpétuo and Wajnman's (1998) analysis of the socioeconomic correlates of sterilization showed that women's education level was inversely correlated with sterilization—presumably because women with greater education levels have more information about alternative contraceptive methods and greater possibilities for using these correctly. Lower and lower-middle class women were much less successful than their more privileged counterparts to use reversible contraceptive methods effectively, and were more likely to undergo sterilization after having more than their desired number of children (Perpétuo & Wajnman 1998, p. 328). The authors conclude that women in lower economic strata have been "doubly penalized" due to restricted access to sterilization and due to the less favorable conditions (i.e. after having more than their desired number of children) under which they were sterilized. Comparing demographic factors between similar cohorts of sterilized and non-sterilized women aged 30-49 in the state of Sao Paulo, Osis et al. (2003, p. 1402) too found that sterilized women were more likely than non-sterilized women to initiate childbearing early in life and lack knowledge to plan and control their fertility.

Some authors argue that women's desire for surgical sterilization goes beyond contraceptive efficacy and reflects profound conflicts in their lives. Women's reliance on sterilization has been seen as a "survival strategy" of poor women (Giffen, 1994), as a method which allows a woman to avoid sexual negotiation with her partner (Goldstein, 1995), and as "refusal of motherhood" by poor women facing economic constraints and the solitary burden of childcare (Corrêa, 1989, p. 35-36). Several scholars have argued that women's decisions to undergo sterilization, in Brazil as elsewhere, are a response to the state's lack of investment in reproductive health care, a pervasive (although sometimes implicit) ideology of population control, and a series of social determinants (such as women's position in the workforce and patriarchal gender relations) that women have little power to change (Barroso, 1984; Lopez,

1997; Petchesky, 1981). Taken together, it is argued, these factors have created a context in which sterilization is "not exactly a choice" (Corrêa, 1994, p. 26).

The institutional arrangement of sterilization services points to collective rather than individual decision-making for sterilization. Caetano and Potter (2004) show for example the role of individual physicians and politicians in facilitating access to sterilization in the Northeast where politicians routinely provide goods and services, including sterilization, to the poor in exchange for votes. In some cases, rather than simply facilitating, physicians directly participated in women's reproductive decisions (Caetano, 2000). In addition to medical indications (e.g. repeat cesareans) for sterilization, physicians made decisions regarding whether to perform sterilization on the basis of "social indications"---"a lot of children, low income, a bad marriage" (Caetano & Potter, 2004, p. 86).

In Taquaril, women's "choices" regarding reproduction are structured in a number of ways. Here I called attention to the construction appropriate female sexuality in the community, women's experiences with motherhood and engagement with notions of "good motherhood," and the women's access to various fertility control methods.

# **GENDER AND SEXUALITY IN TAQUARIL**

A number of sometimes contradictory discourses operate in Taquaril to construct male and female gendered identities. Dominant discourses of appropriate gendered selves are linked to the traditional, patriarchal construction of the Brazilian family. Within this dominant discourse, a woman's primary identification comes from her role as mother and caretaker of the family in the home [*casa*]; whereas a man's identity is linked to being the provider and protector of the family and the mediator between the household and the impersonal, public world of the street [*rua*]. Ideally, an adolescent girl [*moça*] stays at home and remains a virgin until marriage. A woman's status, as well as the honor of her male family members, is dependent on her sexual purity and passivity. Men however are expected to be promiscuous and sexually assertive. A *moça* becomes a woman [*vira mulher*] through her first sexual intercourse and she earns respect as a woman through her marriage, her fidelity to her husband, and her confinement to the house and familial duties.

In Taquaril, this historically dominant ideology of female gender and sexuality, expressed through the idiom of *casa* and *rua*, provides the terms through which women construct their gender and sexual identities. To insult a woman, for example, one only has to suggest that she spends too much time "in the street," because good women, especially if they are married, "stay at home." Few women actually stay at home as housewives and such comments do not refer literally to women's use of the street. The main asphalt street through town is in fact well traveled by women throughout the day as they accompany small children to school, shop at vegetable stands, wait for the city bus, or stop to visit with neighbors. However, talking to men in the street, walking alone after dark, and, above all, hanging around at neighborhood bars [*botecos*] are behaviors that make a woman's sexual purity or fidelity suspect.

This discourse of appropriate female and male sexuality reflects still dominant cultural ideals, but does not necessarily reflect people's actual behavior. The moral landscape in Taquaril, as elsewhere in Brazil, is populated by variety of ways of being which are valued differently and women manipulate the discourse within certain limits to expand the range of

acceptable behavior (Goldstein, 2003; Gregg, 2003). For example, these days virginity upon marriage is a cultural ideal rather than a reflection of women's actual behavior, yet women in Taquaril found it difficult to use contraception at sexual initiation due to sexual double standards of "passive" (innocent) femininity and "active" masculinity (see also Parker, 1991). Planning ahead to use contraception placed a woman's proper femininity into question. Few women in Taquaril reported that they had used contraception before they were in a committed relationship, and even then, usually only after the birth of their first child. Although an "engagement" was sometimes the public recognition of a young couple's sexual relationship, often engagements and marriages came on the heels of a pregnancy.

Women's early sexual and reproductive experiences, however, hold several consequences for their later status and well-being. Georgia Kaufman's (1998) study of marriage, family formation, and fertility in the poor community Alto Vera Cruz, which neighbors Taquaril, showed that virginity was a key factor in whether women legally married. Women who had already had a child by a man who left found it difficult to legally marry another man; they were more likely to enter into a series of temporary unions or to cohabitate in a stable relationship. Kaufman found that legally married women were significantly better off than either cohabitating or single women; more married women worked outside the home and enjoyed greater respect from others and from their partners.

Women, more so than men, are expected to be faithful in marriage, although women said this was changing. From the perspective of women in Taquaril, the main threats to a household are "other women" and alcohol. While male infidelity is considered wrong, it is most threatening only when a man diverts money from the household. Women perceived to be "loose" or promiscuous were not respected and could be denied resources and help that they otherwise might receive. Respect is due only to virgin girls and properly married women.

Young men are expected to be sexually active and assertive—it is only "natural" for a young man to proposition a woman alone on the street. Unmarried adolescent girls, however, are supposed to remain virgin. An adolescent girl's respectability is considered vulnerable and her behavior is carefully controlled and scrutinized by parents. However, parents' work schedules, a girl's own work outside the home, and their desires for a measure of "fun" or "freedom" leave them open to advances by young men.

# GOOD MOTHERHOOD: FAMILY PLANNING AND THE SOCIAL STRUCTURING OF REPRODUCTION

Most marriages in Taquaril are common-law marriages rather than official unions registered "on paper." A couple is considered married if they reside together or the man contributes money to the household. For both a man and a woman, marriage implies "responsibility" and fulfillment of complementary roles in which men provide material support women take care of children and household. As soon as daughters are old enough, they are expected to assume household tasks and most women who work outside the home rely on help from their teen-age daughters. Given high rates of unemployment, many men find it difficult to fulfill their roles as "family provider" and it is often a woman's income that supports the families. Whereas employment for men is hard to find, women, even young girls, can find as work in domestic service. This work arrangement does not necessarily challenge

the dominant gender ideology. In her study of sexuality among low-income women in Recife, Brazil, Jessica Gregg (2003) notes that even when women earned the principal support for the family, they themselves undermined any breaks with the dominant ideology that this might engender by symbolically expecting and demanding male contributions to household maintenance and symbolically demarcating those as the principal, "providing" contributions. For some women, changing gender relations are experienced as a breakdown in the respective responsibilities of men and women. Other women affirm, however, that everyone needs to work to support the family (not only women but adolescents) and likewise that everyone should help out with housework. Still, men's work around the house is considered "help" rather than an obligation.

While motherhood remains central to the gender identity of working-class women (Neuhouser, 1998), women are adamant that families today should be smaller than those of previous generations and they see this as an outgrowth of urban life and the difficulties of raising children in the city and their desire to provide a "better life" for their children. While the experience of motherhood is shaped by the specific material conditions of Taquaril, these women are not isolated from middle class values and trends. Rather, they had ample exposure through media and especially their own personal experience as domestic workers in middle-class families. Women have taken up the discourses found in media, among health care professionals, and among the middle classes that families can be "planned" and that fertility should be controlled.

When I asked how many children women thought was ideal, the answer was invariably "one or two" or "*um casal*" (a boy and a girl). This was in explicit contrast with family size of their parents' and grandparents' generation. In the past, women explained, people had as many children, "as God sends." Nowadays this is not desirable. Despite references to "planned" families, none of the women thought they had been able to achieve a planned family. All of the women I interviewed in Taquaril had more children than their stated ideal, were ambivalent about their fertility, or said that they had gotten pregnant too young or with not enough time in between pregnancies. They were ashamed to have had "one child right after another."

Associations between poverty and (supposedly excessive) fertility are represented in the media and in everyday comments of middle-class and working-class people alike. Lowincome women judge their own "excessive" fertility and that of their neighbors in contrast to the small, planned families they attribute to the middle classes. Residents of Taquaril are very sensitive to the fact large families are associated with backwardness, rural life, and poverty. Family planning is understood as a route to achieve the modern family and social mobility. These notions were not experienced just as generalized goals. In Taquaril, a woman's sexual and reproductive life is under constant scrutiny by family, neighbors, and employers. While this may be true for women of all social strata, poor women rely heavily on their social networks for material support and access to valuable resource, making the influence of others potentially more important in their decisions.

When I asked women why they had decided to undergo sterilization, women said that they needed the *ligadura*<sup>3</sup> because they could not afford to have more children [*não tenho condições*], but they specified that they were not referring to basic necessities such as food. Women identify the primary problem as one of "raising children properly" and providing them with schooling. Sometimes women echoed a popular saying, "Where one eats, ten eat,"

meaning that food could be stretched whereas other material objects and education could not be.

Although not all households had a secure food supply, many women insisted that the problem was not feeding children but getting them to school. The need to provide children with material goods has become a justification for a housewife to seek wage work (even if her husband disapproved). Mothers want to provide nice clothing, book bags, tennis shoes, but these goods were out of reach of many people in Taquaril. A working woman usually pays all of the expenses related to her children's schooling and clothing, even if her husband is employed.

In addition to older signs of a good mother, such as keeping a child well-fed and clean, new requirements for good motherhood include limiting one's family size and the provision of schooling and "information." These statements reflect the discourse of family planners, circulated widely in Brazil as elsewhere, that women should limit fertility in order to be "better mothers" to the children they already have.

With no publicly funded day care available, women are constantly pulled between the necessities of working and caring for children. Those who do this poorly were seen as "letting their kids out loose into the street." If a woman works, she must be able to afford daycare or find someone else to look after her children. Although paid work entails extra burdens of the "double day," it also holds positive significance for many women who said they were glad that they had "learned to work" and had some control over their own money. For women in Taquaril then, fertility control and employment are more than survival strategies (Giffen, 1994); they are seen as an investment and a route to a better future.

Many women in Taquaril however complain that men do not share their concerns; that men have not "kept up" with the increasing importance of the quality of life a parent must offer children as measured in material goods and opportunities of schooling and employment. Some of my interviewees faced staunch opposition from their male partners regarding contraception, and particularly sterilization. Sterilization threatens male control over female sexuality because it effectively prevents pregnancy (in contrast to other methods which many believed could fail). For men especially, sterilization represents a breakdown of an important division between procreative sex within marriage and extramarital sex (on "the street") outside the home or with prostitutes. When a man's wife definitively ends the possibility of procreation through sterilization, she comes symbolically closer to being the "loose" woman on the street for whom sex and reproduction were not linked. So while sterilization is especially attractive for women, it can also be an especially dangerous decision for women for whom alliances with men are central to their material and social well-being. The justification that many women made was that they needed to limit their fertility to provide for the children they have. Limiting family size was not a negation of motherhood but a route to becoming a better mother.

Women's reproductive intentions are framed by shifting meanings of motherhood and gender relations. Why fertility control is rests so heavily on sterilization, however relates to women's distrust of other methods, their negative experiences with contraceptive methods, especially the pill, and a series of obstacles to reversible contraceptive use such as inconvenient access and male opposition.

# ACCESS TO FERTILITY CONTROL METHODS AND ENCOUNTERS WITH HEALTH SERVICES

The first key reproductive decision-making point for many young women in Taquaril occurs only after a delayed menstrual period leads them to suspect a pregnancy. At this point the delayed period can be interpreted in a number of ways depending on whether a woman is married or not, whether the father will recognize this potential child as his, or whether the family can afford another child. Leal and Lewgoy (1995) note for working class women in Porto Alegre, Brazil, a late menstrual cycle introduces a liminal period in which a number of different futures and trajectories are considered. A delayed period "will be dealt with as a possible 'pregnancy' only if a series of moral and material conditions endorse the decision to recognize it as such" (1995, p. 69). If those conditions are not met, a woman will likely indicate that her period is simply delayed and then take any one of a number of remedies used to "bring down the menses." These remedies are known abortifacients. Abortion is a crime and considered by many women in Taquaril to be sin on par with murder; thus most of these women are adamant that once a woman is pregnant, she (and the father) must assume full responsibility for the child. While working-class women firmly assert the "rule" against abortion in general, the same women often adopt a strategic reasoning in relation to specific real-life situations. As the anthropologists note, though, this is not a case of ignorance or hypocrisy. Women's failure to recognize and label a potential pregnancy indicates a social reading of physical alterations of which pregnancy is only one possibility. An unmarried girl may not recognize a delayed period as a sign of a pregnancy unless she is reasonably confident that she will have social support or until various remedies fail to initiate her menstruation. Many of my interviewees said they were unaware of their first pregnancy for many months.

Women's first contact with gynecological services often occurs only after she suspects she is pregnant and goes to the health center for pre-natal sessions. Family planning is frequently included as part of these prenatal classes. Concern with birth control for most women began only after the birth of one or two children. Although women learn about the pill, sterilization, and condoms earlier from other friends and relatives and from *telenovelas*, they generally do not speak to a heath care provider about these until prenatal sessions.

Birth control methods had been theoretically available to women in Taquaril for three years when I arrived in the community, and women understood that to receive birth control through the community health centers they needed to attend a family planning information session and have a gynecological exam. However, contraceptive supplies and services often run short at the health centers. One survey (Maia & Chacham, 2002, p. 12) of community health centers throughout the city of Belo Horizonte showed that no health center offered a full range of contraceptive methods and many offered only a few methods. Given inadequate services at public health centers, women often found it more convenient to purchase oral contraceptives (without a prescription) from the pharmacy. It is not surprising then that many women lacked adequate instruction on pill usage, took pills of incorrect dosage, and took the pills despite contraindications. All of these factors increase the health risks of oral contraceptives, complicate women's compliance and continuation of the pill, and decrease the pill's efficacy (Rosenberg et al., 1995).

National surveys report widespread knowledge of contraceptive methods, but women's ability to name several methods should not be conflated with the kind or quality of knowledge that women possess (Osis et al., 1999). The women I interviewed had experience with the pill, the IUD, and some traditional methods (withdrawal and the calendar [*tabela*] method). Yet, there was widespread belief in Taquaril that the birth control pill and other methods are ineffective. Even women who had used the pill successfully over many years questioned its reliability. Many women reported that they got pregnant while taking the pill. While a few conceded that they may have forgotten to take the pill, others insisted that their bodies had simply "gotten used to the pill" or had used a pill that was "too weak." Many women also quit taking the pill because if its side effects and the feeling they had that it was not "good for one's health."

Other methods were unacceptable to many women. Some women disapproved of the IUD because they considered it an abortifacient. Some told stories of babies born with an IUD embedded in its body, or that that IUDs will "get lost" inside a woman's body. None of the women I met in Taquaril had ever used a diaphragm although some had heard of it. Health workers rarely recommended it because they believed that lower-income women would find it shameful and unacceptable due to a traditional resistance to the manual manipulation it required.

When not attributed to the lack of availability of methods, women's failure to use reversible contraception effectively is almost always interpreted as their ignorance of their own bodies and lack of information about how methods work (Corrêa, 1989; Perpêtuo & Wajman 1998; Osis et al., 2003). My discussions about contraceptives methods with women in Taquaril added complexity to these arguments. Women's failure to use methods effectively should not be reduced in any simple way to a lack of information. Part of women's dissatisfaction and resistance to contraceptives can be attributed to local understandings of how the body works. This local model diverges significantly from the biomedical model and complicates women's use of contraceptive methods (see for example Nichter & Nichter, 1987).

For example, in Brazil, working class representations of the reproduction envision conception as the co-mingling of menstrual blood and semen (Leal, 1996; Victora, 1996). Thus women believed they were most likely to get pregnant on the days directly surrounding their menstrual periods. Many women told me that a traditional method of avoiding pregnancy was to refrain from intercourse for "three days before to three days after" one's menstrual period because menstrual blood both cleans and prepares the womb for conception and is also the very "substance" of a fetus. Accordingly, many contraceptive methods fail to "make sense" (Leal, 1995). For example, a woman stops taking the pill exactly during the days considered most propitious to fertilization. Traditional methods of contraception (abstinence, withdrawal), in contrast, function by either preventing the intimate interaction of bodily fluids or by interrupting a pregnancy in progress. The pill also frequently diminishes the quantity of a woman's menstrual flow which women perceived "blocked" or "retained" menstrual blood. When women complained of negative side-effects of the pill, such as included headache, nausea, fatigue and nerves, and they explained that this was due to the excesses of menstrual blood which had failed to descend (Bessa, 2006). These factors led women to periodically "rest from" the harmful effects of the pill. Side effects, such as changes in menstrual flow may be considered trivial from a biomedical standpoint but are highly unsatisfactory to women.

The inaccessibility and unacceptability of reversible methods meant that women had more children than they desired and had little success spacing their pregnancies. Women's prior use of contraceptive methods, particularly contraceptive failure and adverse effects, leads them to choose sterilization (Viera, 1999). Without recourse to legal abortion and often after a harrowing experience of clandestine abortion, sterilization presented the best option that many women could make although it often came after the failure to use the pill effectively or exhaustion from unwanted side effects.

Having decided to seek sterilization, however, a woman had no assurance that she would achieve it. Before sterilization was regulated and widely offered in the public health system, one of the biggest issues for women was how to gain access to the procedure. Many women actively sought out sterilization from the local health center and regional hospitals. Some would save money to pay out-of-pocket for the procedure; others used employers or middle-class connections to find "sympathetic" physicians willing to perform the operation for free. Some women hoped for or requested a cesarean childbirth so that they could then request a tubal ligation. Unsurprisingly, women knew about the justifications that medical professional used to evaluate women's requests for sterilization and they used these in their favor. Cecília underwent sterilization at age 26, immediately following the cesarean birth of her fourth childbirth was a planned cesarean planned so that she could have her tubes tied.

I had to invent health problems in order to be able to *ligar* through the [Military Police health plan]. At that time [before it was legal], they wouldn't *ligar* a young woman. But I had some health problems, and the doctor invented some other ones so that she could do the sterilization. I told her, 'If you don't want to give me the sterilization, and I have more children, then you can take responsibility for them. I'll bring them to your house!'

Even after legalization, women still encounter obstacles such as a lack of services or the need to obtain her husband's consent. In this context, it is the authority of a physician's recommendation of sterilization, or a physician's agreement to perform a tubal ligation, which allows some women to challenge patriarchal authority (Citeli et al., 1998). Both husbands and physicians become gatekeepers with which women negotiate fertility control.

Although some women have found creative means for accessing a service denied to them, the manner in which they undergo sterilization presents a number of negative consequences. A lack informed consent and appropriate counseling means that some women may not fully understand the irreversibility of the procedure. One common belief, for example, is that there are two distinct techniques of sterilization, "cutting" the tubes, which is permanent, and "tying" the tubes, which may be "undone" (see also Vieira & Ford, 1996). In terms of the number of women affected however, the greatest cost to women's health of clandestine sterilizations has been the link to c-sections. Cesareans have become a "price to pay" for access to tubal ligations. Vaginal births are less costly than cesareans and subject women and infants to fewer health risks. In addition women overwhelmingly prefer vaginal birth to caesarian (Perpetuo, Bessa & Fonseca, 1997), due mostly to fast recovery time. However, women's desire for tubal ligations has often made them complicit in the "surgicalization" of childbirth from which obstetricians benefit.

# **TANIA'S STORY**

Patterns of gender and sexuality, motherhood, and the organization of both gynecological and obstetric care all provide vital background without which individual women's stories are unintelligible. Having introduced these issues, we can now begin to understand Tania's story. Here, the social context of decision-making is central to how and why Tania first decided to undergo surgical sterilization and shortly thereafter came to regret it.

When I first met her in 1996, Tania was 26 years old and the mother of five children. She was employed as a domestic worker in a middle class household but earned only R\$112 (one minimum wage) per month, about half the going rate. She had come to Taquaril as a young girl with her parents and three older brothers. Her precarious two-room house sat at the bottom of a steep, dirt alleyway near the house her father built on a lot that had since been divided up among her relatives. Her mother was an alcoholic and died when Tania was 15. Two of her brothers also died; one was killed in an accident and another was murdered in gang-related violence. By 1996, Tania's relatives in the neighborhood included a widowed sister-in-law with several children and a brother, his wife and children (who lived next door). Her father, step-mother and step-siblings had moved to another neighborhood.

Tania says that she was a reckless [desmiolada] adolescent who resisted her parents' attempts to keep her close to home. At age 14 Tania's was raped by a neighborhood boy who had been her best friend when both of got her drunk on cane alcohol. She never told anyone about this experience. After this she says she had many boyfriends [*era namoradeira*]. She began working as a live-in domestic worker at 15 but says she really only wanted to go to parties [*ir para a farra*] although her father prohibited her from "running around" on the street with her friends. From her work place and without the knowledge of her parents, Tania began seeing a young man who let her stay with him and took her to fun places. When she realized that she was pregnant, she panicked about how her mother would react (she had never gotten along with her mother) and attempted to abort the fetus by jumping from high heights and drinking a concoction that included a cleaning agent in warm milk. None of these methods worked. In the same month that she realized she was pregnant, her mother died of cirrhosis. She never knew that Tania was pregnant.

Tania speculates that her mother's death softened her father's reaction because he did not immediately expel Tania from the house once he discovered that she was pregnant. He did insist that Tania marry the child's father, so she went to live with him and his family for awhile. She says however that it just did not work, she did not like him very much and his family did not like her. She said, "you can't marry just to please someone else," and besides she had already met and really liked "Mario," her current husband but was "ashamed" to approach him because she was pregnant. Tania moved back to her father's household.

After her first child, a daughter, was born, Tania and Mario began seeing each other. Her father told her that if she got pregnant again he would throw her out of the house. But he had already begun constructing a small shack for her behind his house. When Tania did get pregnant, she and Mario began living together [*amigada*] there.

Mario's only work involved drug trafficking and he had a tough demeanor and stone-face look that fit the part. Mario was an alcoholic and violent, and their relationship was tumultuous. Tania explained "from the beginning we fought all the time and when I was pregnant we fought more." He would stay with Tania for some time and then leave to stay with his mother across town. Tania set about becoming a respectable wife, but she had a reputation of being wild and irresponsible, and seems to have lived up to it when using drugs (marijuana) and alcohol.

Over the next eight years, Tania had four more children by Mario, all boys. Mario provided only sporadic contributions to the household, so Tania's had to work to support her family. On many days, she simply left her children at home because there was no one to look after them; some neighbors said it was a shame that she was "throwing her children into the street." She said community members judged her harshly for "having one child right after the other." She learned about birth control pills from a sister-in-law tried taking them after her third child was born. After six or seven months, she stopped taking them they made her nauseous and anxious [*nervosa*]. She was afraid of the IUD because she heard that it could get "lost" in the body, but she was desperate to stop having children.

By the age of 24, Tania was pregnant for the 5th time. Neighbors, family and even her employer told Tania she had had too many children, too quickly urged her to get a tubal ligation. She asked about sterilization at the family planning session but was told she would never be approved for it due to her young age. When her youngest son was born, he needed medicines she could not afford and was forced to leave him in the care of an aunt. This, for her, was the last straw; "I couldn't even keep my son with me." Tania began to look for a way to get sterilized; she said that she wanted to be able to "provide a better life" for her children and to "give them education, because it's so important today."

A while after my youngest son was born, my step-mom was working [as a janitor] at [another hospital] and got to know [a doctor who would perform the tubal ligation].

When she went to see this physician about the tubal ligation, he at first said she was too young (24), but Tania pleaded:

I explained my situation to the doctor and I think that he felt sorry for me and he told me that my husband had to sign [consent] papers, and that was a big problem because I never had a daughter with [Mario]. And [Mario] said that he didn't want to sign because he didn't have a daughter and that was his dream. And so I said, 'There's the paper. If it's not signed when I get back, then I won't stay with you because you are not being a responsible man and I don't want to put children in the world to suffer. We can't afford to have more children.' When I arrived, he asked if I was certain that this was what I wanted and I said, 'I'm absolutely sure.' He signed against his will, but he signed it. And then I took the forms and the secretary of the doctor told me that she'd contact me in 15 days.

After 15 days had passed she had heard nothing and called the secretary from her employers' home to ask about getting sterilized. They told her they could do it if she came right away. She did not have time to tell anyone and left immediately for the hospital. Arriving there she was prepped for the procedure and when she met the doctor he asked if she was sure and reminded her that she might regret it. She assured him that she was sure but was also shaking terribly from fear. He told her that if she did not calm down they would not be able to proceed. She told him that she wanted to watch the surgery because "many people said that sometimes the tubes just get 'tied'" [rather than cut] and she wanted to make sure that her tubes were cut. The physician reassured her and after the procedure, he gave her the cut pieces of tissue in a jar and told her to return in eight days, but she did not go back.

Soon after her surgery, Tania began to regret her decision to get sterilized. Tania developed physical problems that she attributed to the tubal ligation. She said her body

"rejected" the sterilization and she began to feel "tired, nervous agitation. I feel like I have too many hormones. I have sex, but cannot climax—and all this is after the tubal ligation because before I was normal. Now I also have this problem of an infection but that's a little of my own carelessness because I haven't gone to check with the doctor."

After the tubal ligation, ...there is pain here and there. [She points to her abdomen.] .....I'm not the same. If I could go back to the way I was I'd go back. I didn't do it [the surgery] because I wanted to. I did it because I couldn't stand to suffer anymore, because my husband drank too much, and so he didn't help me in any way. And so I thought that by having the *ligadura*, I would be solving my problems but unfortunately, my problems have multiplied. I think that the *ligadura* finished me off, demolished me...

By six months after the sterilization, her marriage had weakened and Mario was spending more days at his mother's home. She wished that she could have more children, especially the daughter that Mario wanted. Her step-sister got pregnant, and Tania said, "I started to enjoy her pregnancy and I wanted to be pregnant too. Now she's had the baby and it's a girl, and now I regret it much more."

She could not say however, what else she should have done.

I didn't have another choice because my husband drank so much and even hit me. I was having hard times and I needed to work, so I had to do the tubal ligation, because I couldn't go on that way. And so I thought that that was a solution but today I see that I was completely wrong.

Tânia's regret was understood by some of her neighbors as insanity [*locura*]: "How could she want more children when she can't raise the ones she's got!" they exclaimed. But her regret is not surprising given that she was only 25 years old, had no counseling, and few adequate alternatives. She struggled to raise her children in poverty and with alcoholic and abusive husband.

Tania was well aware that sterilization is irreversible and consciously embarked on a "pilgrimage for sterilization," but her subsequent regret resulted from her belief that she could prevent the further weakening of her marriage by having a daughter for her husband. Given Mario's increasing absence, Tania lost both material support and social respect. She said, "because of the way I'm living, not with my husband, people think badly of me, I'm way down low. They say I couldn't even manage to hold onto a husband. They say my kids don't have a father." To counter this she has joined the Assembly of God Church which helps her gain community respect by being a woman who "meditates on the Bible."

By 2002, Tania had all her children (including her youngest) living with her. Her 17 yearold daughter helped her keep track of the younger boys and was working herself as a parttime cleaning lady [*faxineira*]. They had made improvements to their home. Mario still visited Tania occasionally and she considered him her husband. She still hoped for marriage because "he's the only man for me" and her church told her that if she lived with him without a proper marriage then she would be "prostituting herself." She was a religious woman now and had earned some respect from her neighbors. She still regretted her decision to undergo sterilization.

# CONCLUSION

Tania's story is not meant to be representative of poor women's experiences. Even by Taquaril standards, Tania has had a hard life. Tania struggled to access sterilization only to regret her decision shortly thereafter. Her dilemma and regret begins prior to decision to undergo sterilization with her sexual experiences in adolescence and later her inability access and use other birth control methods effectively. Her story does bring into relief the kinds of issues women took into account with deciding to seek out sterilization, as well as the broader, taken-for-granted cultural assumptions related to gender and sexuality that shaped women's lives.

The language of "choice" is far too simple to capture Tania's actions. Tania's decision to undergo sterilization cannot be adequately understood if taken as a discreet event or isolated from the larger social context. To understand Tania's story requires an examination the changing patterns of gender and sexuality in Taquaril, the role of reproduction in women's lives, women's ideas about health and their bodies, as well as the availability of health services in Taquaril and the relationships between medical professional and low-income women. Addressing this lacuna in healthcare will require more than expansion of family planning services and individual counseling. Health care workers and advocates will have to consider the context of counseling and family planning available to women, and women's understanding of conception and their bodies.

An examination of women's own assessments of their contraceptive needs and the social context of their decisions does not deny any of the structural and historical forces but allows us to see the ways these larger forces impinge on women's lives and how women negotiate and make sense of these various forces. My attempt has been to account for women's perceptions of their actions in a way that grants them agency but which recognizes that decisions are not made in a social vacuum—they emerge in particular settings and involve other actors.

# **ENDNOTES**

- <sup>1.</sup> The minimum salary is a federally established monthly wage. In 1998, the minimum wage was R\$130, which corresponded to approximately US\$107. Income is usually referred to in terms of the number of minimum wages one receives
- <sup>2</sup> In the city of Belo Horizonte, a 1992 municipal law provided for voluntary sterilization but few facilities offered the service and those that did required that each case be approved by a medical committee, which assessed a woman's age and parity (both higher than those established by the 1997 federal law), medical history, spousal consent and socioeconomic factors.
- <sup>3.</sup> In addition to *esterilização* (sterilization), women in Taquaril usually referred to the procedure as a *ligadura* (ligature) or a *ligação* (ligation) as well as with the verb form *ligar* (to ligate). *Ligar as trompas* corresponds to the English "to tie the tubes."

#### REFERENCES

- Barros, F.C., Vaghan, J. P., Victora, C.G. & Huttly, S.R. A. (1991). Epidemic of cesarean sections in Brazil. *The Lancet 338*, 167-169.
- Barroso, C. (1984). Esterlização feminina: Liberdade e opressão. Revista de Saúde Público 18, 170-80.
- BEMFAM/DHS (1996). *Brasil: Pesquisa Nacional sobre Demografia e Saúde*. Rio de Janeiro: Bemfam-Departamento de Pesquisas Sociais/Demographic and Health Surveys.
- Berquó, E. (1993). Brasil: Um caso exemplar (anticoncepção e parto cirúrgicas) a espera de uma ação exemplar. *Estudos Feministas 1*, 366-381.
- Berquó, E. & Cavenaghi, S. (2002). Direitos reprodutivos de mulhers e homens face à nova legislação sobre esterilização voluntária. Paper presented at the XIII Ebcontro da Associação Brasileira de Estudos Populacionais, Ouro Preto, Minas Gerais, Brazil November 4-8.
- Bessa, G. L. H. De (2006). Ethnophysiology and contraceptive use in among low-income women in Brazil. *Health Care for Women International* 27, 1-25.
- Caetano, A. J. (2000). Sterilization for Votes in the Brazilian Northeast: The Case of Pernambuco. Ph.D. Dissertation. Department of Sociology, University of Texas at Austin, Texas.
- Caetano, A. J., & Potter J. (2004). Politics and female sterilization in Northeast Brazil. *Population and Development Review 30* (1), 79-95.
- Carter, A. (1995). Agency and fertility: For and ethnography of practice. In S. Greenhalgh (Ed.) Situating Fertility: Anthropology and Demographic Inquiry. Cambridge: Cambridge University Press.
- Corrêa, S. (1989). Uma recusa da maternidade? In S. Corrêa and A. P. Portella (Eds) *Os Direitos Reprodutivos e a Condição Feminina: Textos*. Recife: SOS CORPO Grupo de Saúde da Mulher.
- Corrêa, S. (1994). Sterilization in Brazil: Not exactly a choice. In *We Speak for Ourselves.Population and Development*. Compiled by the Panos Institute. Washington, D.C.
- Corrêa, S., & Petchesky, R. (1994) Reproductive and sexual rights: A feminist perspective. In Sen, G., Germain, A., & Chen, L. C. (Eds.) *Population Policies Reconsidered: Health Empowerment and Rights*, Harvard Series on population and International Health. Boston: Harvard School of Public Health
- Faria, V. (1989). Políticas de governo e regulação da fecundidade: Consequências não anticipadas e efeitos perversos. *Ciências Sociais Hoje*, 62-103.
- Faundes, A. & J. G. Cecatti (1993). Which policy for cesarean sections in Brazil?—An analysis of trends and consequences. *Health Policy and Planning* 8(1), 33-42.
- Garro, L. (1998). On the rationality of decision-making studies: Part 1: Decision models of treatment choice. *Medical Anthropology Quarterly 12*(2), 319-340.
- Giddens, A. (1979). Central Problems in Social Theory: Action, Structure and Contradiction in Social Analysis. Berkeley: University of California Press.
- Giffin, K. (1993). Women's health and the privatization of fertility control in Brazil. *Social Science and Medicine 39*(3), 355-360.

- Goldani, A. M. (2001). Rethinking fertility decline. Paper presented at the XXIV General Population Conference of the International Union for Scientific Studies of Population (IUSSP). Salvador, Bahia, Brazil 18-24 August, 2001.
- Goldstein, D. (1995). Aids and women in Brazil: The emerging problem. *Social Science and Medicine*, *39*(7), 919-929.
- (2003). Laughter Out of Place: Race, Class, Violence, and Sexuality in a Rio Shantytown.Berkeley: University of California Press.
- Gregg, J. (2003). Virtually Virgins: Sexual Strategies and Cervical Cancer in Recife, Brazil.Stanford: Stanford University Press.
- Kaufman, G. (1998). Gender and reproductive decision-making: The contraceptive choice of women in a Brazilian favela. In G. Martine, M. Gupta, & C. C. Lincoln (Eds.) *Reproductive Change in India and Brazil.* Delhi: Oxford University Press.
- Lave, J. (1988). Arithmetic Practices and Cognitive Theory: An Ethnographic Inquiry. Cambridge: Cambridge University Press.
- Leal, O. F. & Lewgoy, B. (1995). Pessoa, aborto e contracepção. In O. F. Leal (Ed.) Corpo e Significado: Ensaios de Anthropologia Social. Porto Alegre: Universidade Federal do Rio Grande do Sul.
- Lindbladh, E. and Lyttkens, C. H. (2002). Habit versus choice: The process of decisionmaking in health-related behavior. *Social Science and Medicine* 55, 451-465.
- Lopez, I. (1996). Agency and constraint: Sterilization and reproductive freedom among Puerto Rican women in New York City. In L. Lamphere, H. Ragoné, & P. Zavella (Eds.) Situated Lives: Gender and Culture in Everyday Lives. New York: Routledge.
- Lopez, I. (1997). An ethnography of the medicalization of Puerto Rican women's reproduction. In M. Lock & P. Kaufert (Eds.) *Pragmatic Women and Body Politics*, Cambridge: Cambridge University Press.
- Luiz, O. C. & Citeli M. T. (1999). Esterilização cirúrgica: Lei que fica no papel. *Jornal da Rede Feminista de Saúde. No. 20.*
- Maia, M. B. & Chacham, A. (2002). Grupos de planejamento familiar dos centros de saúde do município de Belo Horizonte: Proposta de avaliação de qualidade do serviço ofertado na rede. Paper presented at the XIII Encontro da Associação Brasileira de Estudos Populacionais held in Ouro Preto, Minas Gerais, Brazil, from November 4-8, 2002.
- Martine, G. (1996). Brazil's fertility decline, 1965-1995: A fresh look at key factors. In G. Martine, M. Das Gupta & C. L. Chen (Eds.) *Reproductive Change in India and Brazil*. Delhi: Oxford University Press.
- Merrick, T. & E. Berquó (1983). *The determinants of Brazil's recent rapid decline in fertility*. Washington, D.C.: National Academy of Science Press.
- Neuhouser, K. (1998). "If I had abandoned my children:" Community mobilization and commitment to the identity of mother in Northeast Brazil. *Social Forces* 77(1), 331-58.
- Nickter, M., & Nickter, M. (1987). Cultural notions of fertility in South Asia and their impact on Sri Lankan family planning practices. *Human Organization 46*(1), 18-28.
- Osis, M. J. D., Faúndes, A., de Sousa, M. H. & Bailey, P. (1999). Consequências do uso de métodos anticoncepcionais na vida das mulheres: O caso da laqueadura tubária. *Cadernos de Saúde Pública 15*(3), 521-532.
- Osis, M. J. D., Faúndes, A, Sousa, M.H., Duarte, G. A., & Bailey. P. (2003). Fertility and reproductive history of sterilized and non-sterilized women in Campinas, São Paulo, Brazil. *Cadernos de Saúde Pública 19*(5), 1399-1404.

- Parker, R. G. (1991). Bodies, Pleasures, and Passions: Sexual Culture in Contemporary Brazil. Boston: Beacon Press.
- Perpetuo, I. H.O., Bessa, G. H. de, & Fonseca, M. C. (1997). Parto cesáreo: Uma análise da perspectiva das mulheres de Belo Horizonte. Anais do XI Encontro Nacional de Estudos Populacionais, ABEP Vol. 4, 95-119.
- Perpetuo, I. & Wajnman, S. (1998). Socioeconomic correlates of female sterilization in Brazil. Paper presented at CICRED- ISUNAM Seminar on Poverty, Fertility and Family Planning, June 2-4, Mexico City.
- Petchesky, R. (1981). "Reproductive choice" in the contemporary United States: A social analysis of female sterilization. In K. L. Michaelson (Ed.) And the Poor Get Children: Radical Perspectives on Population Dynamics. New York and London: Monthly Review Press.
- Potter, J. E. (1999). The persistence of outmoded contraceptive regimes: The cases of Mexico and Brazil. *Population and Development Review 25*(4), 703-739.
- Potter, J. E., Perpétuo, I.H. O., Berquó, E., Hopkins, K., Leal, O.F. Formiga, M. C. C., & Souza, M.R. (2003). Frustrated demand for postpartum female sterilization in Brazil. *Contraception* 67, 385-390.
- Rosenberg, M. J., Burnhill, M. S., Waugh, M. S., Grimes, D. A. & Hilard, P. J. A. (1994) Compliance and oral contraceptives: A review. *Contraception* 52(3), 137-141.
- Unnithan-Kumar, M. (2001). Emotion, agency, and access to healthcare: Women's experiecnes of reproduction in Jaipur. In S. Tremayne (Ed.) *Managing Reproductive Life*. Fertility, Reproduction and Sexuality, Vol 1. New York: Berghahn Books.
- Victora, C. G. (1995). As imagens do corpo: Representações do aparelho reprodutor femino e reapropriações dos modelos medicos. In O. F. Leal (Ed.) Corpo e Significado: Ensaios de Antropologia Social. Porto Alegre: Universidade Federal do Rio Grande do Sul.
- Vieira, E. M. (1999). Do women's attitude towards abortion and contraceptive methods influence their option for sterilization? *Cadernos de Saúde Pública 15*(4), 739-747.
- Vieira, E. M., and Ford, J. N. (1996). Regret after female sterilization among low-income women in São Paulo, Brazil. *International Family Planning Perspectives*. 22(1), 32-40.

Chapter 25

# COMPARING RISKY AND INTER-TEMPORAL DECISIONS: VIEWS FROM PSYCHOLOGY, ECOLOGY AND MICROECONOMICS\*

Tobias Kalenscher<sup>\*1</sup> and Philippe N. Tobler<sup>2</sup>

 <sup>1</sup>Animal Physiology and Cognitive Neuroscience
 Neurobiology Section of Swammerdam Institute for Life Sciences (SILS) University of Amsterdam, Faculty of Science
 Kruislaan 320, 1098 SM, Amsterdam, The Netherlands
 <sup>2</sup>Department of Physiology, Development and Neuroscience University of Cambridge, Downing Street, Cambridge CB2 3DY, United Kingdom

# ABSTRACT

When making decisions between different options, we often consider two basic properties of these options, how risky they are and when they will occur. For example, we may choose to gamble or to wait for a larger reward. Decisions under risk refer to decisions among known probabilistic options, inter-temporal decisions refer to choices between options that will be realized at known future timepoints.

Risky and inter-temporal decisions have been captured theoretically primarily by Ecology and Microeconomics but findings from Behavioral Economics, Psychology and Neuroscience often contradicted theoretical predictions. As a consequence, a wealth of more descriptive models has emerged to explain the findings. A subset of these models has stressed the similarities between risky and inter-temporal decisions. In this chapter we review both core theoretical approaches and empirical findings. We discuss possible explanations for discrepancies and identify key behavioral experiments.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

Ph: +31 (0)20 525 7658; Fax: +31 (0)20 525 7709: Email T.Kalenscher@uva.nl Web: http://home.medewerker.uva.nl/t.kalenscher/

## **1. INTRODUCTION**

When we make decisions, the outcomes of our choices rarely occur with certainty, and often we have to wait some time for the consequences to happen. For example, investing time and money into a good education makes more likely, but doesn't guarantee, a successful professional career and a high income; in other words, the outcome of your investment decision is probabilistic. Likewise, when paying the high tuition fees for your education, you invest resources now for benefits that are yet to come, because you will only be able to harvest the fruits of your labor once you finish your education in a couple of years. Choices between probabilistic outcomes are called 'risky decisions' and choices between outcomes that will be realized at different instants in the future are called 'inter-temporal decisions'. Both types of decisions have been extensively discussed in several scientific disciplines, including biology, ecology, micro- and macroeconomics, psychology and cognitive neuroscience. In this chapter, we review some of the most influential theories on risky and inter-temporal decision making, and will outline the theoretical and empirical differences in the different approaches. We will then discuss to what degree attempts to unify the two research fields are and can be successful and identify key behavioral and neuroscientific experiments. We conclude with highlighting the importance of cooperation between the various disciplines in elucidating the effects of risk and time on choice.

# **2. DECISIONS WITHOUT RISK**

Decisions between certain, immediate, but quantitatively different choice outcomes appear easy: You just compare which of the two outcomes results in the higher gain or the smaller loss, and choose accordingly. However, how do you compare two qualitatively different commodities, for example, apples and pears? In economics, this problem is solved by assuming that different commodities are translated into a common currency, the subjective value, or the utility of a prospect. Utility is a measure of relative satisfaction or gratification which allows to rank-order and therefore compare the different possible outcomes (Montague and Berns, 2002). Although frequently used in financial contexts (also in this chapter), utility does not exclusively refer to monetary gains (and losses), but also to more abstract benefits, such as obtaining pleasure from engaging in a favorite recreational activity, or enjoying one's favorite food, or the like. Although embracing essentially the same solution, behavioral ecology has given a biological twist to the common currency problem. It replaces utility with fitness which, depending on the model, may correspond to e.g. rate of energy gained per unit of time spent foraging (Charnov, 1976) or to reproductive success (Hamilton, 1964).

The utility of an outcome is not a linear function of its objective (e.g., monetary) value, but a function of the current level of wealth (Friedman and Savage, 1948; Bernoulli, 1954; Kahneman and Tversky, 1979; Tobler et al., 2007a). More precisely, it has been argued that each additional unit in the utility function, the so-called marginal utility (Mankiw, 2004), is smaller than the previous unit, resulting in a progressive decrease in marginal utility with increasing assets or energy reserves (Friedman and Savage, 1948; Bernoulli, 1954; Sibly and McFarland, 1976; Kahneman and Tversky, 1979; Kacelnik, 1997; Kacelnik and Bateson, 1997; Tobler et al., 2007a). As a consequence, the utility of a commodity is presumed to be a

decelerating concave function of this commodity (figure 1A). For example, winning \$100 would be more valuable to you when you are poor than when you are a millionaire.

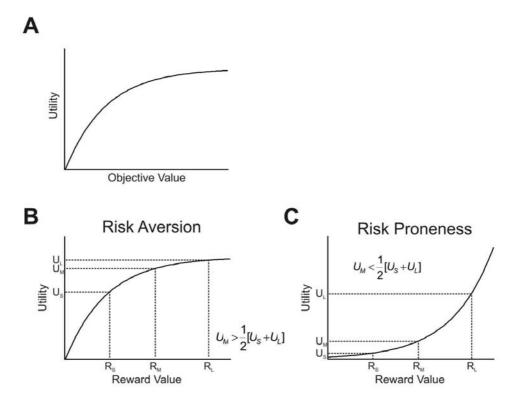


Figure 1. Utility functions. (A) Utility as a function of the objective value of a commodity. The utility curve is a concave function of the current level of wealth because the marginal utility, i.e., the utility increment with each additional unit, decreases with increasing level of wealth. (B) A concave utility function predicts risk aversion when choosing between a medium-sized, certain and a large and small risky reward. In a multi-choice situation, the average utility of the certain rewards exceeds the average utility of the risky rewards because the utility of the large reward is sublinearly larger compared to the utilities of the other rewards. (C) A convex function predicts risk proneness.

## **3. DECISIONS UNDER RISK**

Risky decisions are decisions between probabilistic outcomes. The level of riskiness is equivalent to the spread from the mean (variance) of the risky outcomes. For example, a gamble that pays either \$ 9 or 11 with 50% probability is less risky than a gamble that pays either \$ 2 or 18 with 50 % probability (although both gambles have the same mean payoff of \$ 10). This notion of risky decision making is important, as it differs from the common, folk-psychological conception of risky behavior, which frequently implies that a person engaging in risky choices is consciously willing to accept high losses ("the gambler who bets his house and family"). However, although the readiness to accept large losses may certainly play a role in biasing an individual's risk attitude, academic research of risky decision making focuses on the formal impact of outcome variance on choice.

Neither humans nor animals are risk-neutral (Friedman and Savage, 1948; Kacelnik and Bateson, 1996, 1997; Bateson and Kacelnik, 1998): When deciding between a certain and a risky option, one is often chosen more frequently than the other, even when the expected values of the respective options (their probabilities multiplied with their objective values) are identical. Most animals, including humans, are risk-averse, but show occasional risk-proneness (Friedman and Savage, 1948; Kacelnik and Bateson, 1996, 1997; Bateson and Kacelnik, 1998; Glimcher, 2002; McCoy and Platt, 2005a, 2005b; Hayden and Platt, 2007; Tobler et al., 2007b). Importantly, this suggests subjective differences in the valuation of the same objective options. Utility theories have been put forward in order to capture subjective valuation with a particular focus on risk.

#### Economic Utility Theories

Utility theories assign numbers to preferences with both a descriptive and a normative purpose. Specially psychology is interested in explaining choice behavior and has thus focused on the descriptive aspect whereas economy and statistics, more so in the past, have put stronger weight on the normative aspect of characterizing consistent, coherent and optimal choice and ecology is interested in both descriptive and normative aspects. Utility theories all make basic assumptions (axioms) about the elements of the decision space and the preference relations of the decision maker with respect to these elements. From these axioms they deduce statements (theorems) for example about how the preference relations as observed from choices can be transformed into utility relations (numbers).

Von Neumann and Morgenstern (1944) axiomatized utility theory by requiring completeness, transitivity, continuity and (not explicitly stated but necessary) independence of preferences. Thereby they founded expected utility theory (EUT), which provides the most prominent normative framework for the analysis of decisions under risk. The completeness axiom requires that the decision maker has preferences across options, transitivity that preferences are in a basic hierarchical order, continuity that for each option there is a better and worse one and independence that preferences do not change by adding common outcomes to all options. If these axioms are fulfilled then a number u (x) can be assigned to each x so that:

$x \le y$ if and only if $u(x) \le u(y)$ .	(1)
--	-----

The preference relation  $x \le y$  may be read as "alternative x is not preferred to alternative y", <= corresponds to the standard "smaller than or equal to" and u(x) and u(y) refer to the utilities of x and y. If this is true then utility function u preserves the ordering of  $\le$  and allows translating utilities to preferences and vice versa.

As long as the axioms described above hold EUT proposes that a decision whether to accept or reject a choice option should be made by multiplying the utility of all possible outcomes of the option with their probabilities, integrating across products, and choosing the option with the larger sum. Thus, the utility of a choice option with risky outcomes corresponds to its expected utility. For example, if you consider that it is relatively likely to receive a high income following a good education, and that with the expected income you

will be better than with your current situation, then you should decide to invest into a good education. In general, benefits represent positive utility values, costs negative values.

The shape of the utility curve can be related to people's risk attitudes. A risk neutral person has a linear utility curve, and for such a person the expected utility of a gamble is equivalent to the utility of the mean of the gamble. Convex utility curves correspond to risk-proneness, concave curves to risk aversion (figure 1B and 1C). To use a simple example, imagine a situation in which an agent chooses between a certain option, offering a medium-sized reward, and a risky option, offering a large and a small reward with a 50% chance each. The expected value of both options is identical. According to EUT, the expected utility of a given choice option is the sum of the utilities of each possible outcome multiplied with their probability:

$$E[U_{option}] = \sum p_{outcome} \bullet U_{outcome}$$
(2)

The expected utility of the certain option would accordingly be computed as:

$$E[U_{certain}] = 1 \bullet U_{medium-reward}$$
<sup>(3)</sup>

and the expected utility of the risky option would be computed as:

$$E[U_{risky}] = (0.5 \bullet u_{small-reward}) + (0.5 \bullet u_{large-reward})$$
<sup>(4)</sup>

Due to the concavity of the utility function of risk averse agents, the utility of the large reward,  $u_{large-reward}$ , is sublinearly larger than the utilities of the medium or small rewards,  $u_{medium-reward}$  and  $u_{smaller-reward}$ . Thus,  $EU_{risky}$  will be smaller than  $EU_{certain}$  and agents will avoid the risky option (see figure 1B). Put in simpler words, an agent receives the large and small rewards with equal probability when choosing the risky option. Because the utility of the large reward is sublinearly smaller than the utilities of the other rewards, the mean utility of the large and small rewards (risky option) would be smaller than the utility of the gamble's expected value and thus the utility of the medium-sized certain option. As a consequence, the agent will avoid the risky option. Risk-proneness can be explained by assuming a convex, accelerating utility curve, in which the utility of the large reward is supralinearly larger than the utility of a medium or a small reward (figure 1C).

#### Violations of Preference Axioms

Empirical research showed violations of most of the normative axioms of EUT. Reports of violations of the independence axiom appeared relatively soon after von Neumann and Morgenstern's seminal work (Allais, 1953; Ellsberg, 1961). As an example, consider the following two decisions (Kahneman and Tversky, 1979):

Decision 1)

- a) \$2500, P = 0.33; \$0, P = 0.67 (Read as: a 33% chance of winning \$2500 and a 67% chance of winning nothing)
- b) \$ 2400, P = 0.34; \$0, P = 0.66

Decision 2)

- a) \$ 2500, P = 0.33; \$ 2400, P = 0.66; \$ 0, P = 0.01
- b) \$ 2400, P = 1.0; \$ 0, P = 0.0

Most people choose a) in decision 1) and b) in decision 2). However this pattern of preference reversals violates the independence axiom because 2a and 2b result from adding (\$ 2400, P = 0.66) to 1a and 1b and therefore either a) or b) should be chosen in both cases.

Also the transitivity axiom can be systematically violated such that decision makers show cyclic preferences (A $\geq$ B, B $\geq$ C, C $\geq$ A; e.g. Loomes et al., 1991; Shafir, 1994; Waite, 2001). For example (Waite, 2001), blue jays prefer one raisin, 28 cm into a tube (option A) over two raisins, 42 cm into a tube (option B). They also prefer option B over three raisins 56 cm into a tube (option C) but when given the choice between options A and C, they do not prefer A. Humans also show systematic violations of transitivity in certain choice situations (Tversky, 1969), for example, when the choice options are composed of several features that vary along different dimensions. Models of context-dependent choice such as regret theory (Loomes and Sugden, 1982) suggest that violations of the transitivity axiom arise from changes in utility because decision makers evaluate options not in isolation but consider also the outcomes of unchosen alternatives.

Kahneman and Tversky (1979) have pointed out another problem of EUT in that it does not account for differences in how decision problems are described (framed). As an example, consider the following two decisions (Tversky and Kahneman, 1986):

Decision 3) a) \$ 240, P = 1.0 b) \$ 1000, P = 0.25; \$0, P = 0.75

Decision 4) a) \$-750, P = 1.0 b) \$-1000, P = 0.75; \$0, P = 0.25

Most people choose 3a and 4b. However, the combination of 3a and 4b is dominated by (has a lower expected value than) the combination of 3b and 4a.

Findings of axiom violations have provoked different reactions. Some theorists have relaxed one or more of the axioms (e.g. Machina, 1982; Fishburn, 1982) others have given up on the project of axiomatising utility theory and proposed purely descriptive models. The most famous of these latter approaches is prospect theory (Kahneman and Tversky, 1979; 1992). Prospect theory suggests that the subjective value function is concave for gains and convex for losses and steeper for losses than gains (figure 2). This reflects the finding that decision makers are usually risk seeking for losses, risk averse for gains and reluctant to accept a fair bet on the toss of a coin (in fact, potential gains have to exceed potential losses by a factor of about 2 in order to achieve indifference). The different steepness for gains and

losses introduces a "kink" in the value function which makes it difficult to treat mathematically and which has been termed "reference point". Although not formally defined, the reference point often corresponds to the status quo or the current wealth level. Moreover, decision weights modulate probabilities according to an inverted-S-shaped probability weighting function. This reflects the finding that many decision makers overweigh small and underweigh large probabilities, at least when making hypothetical decisions.

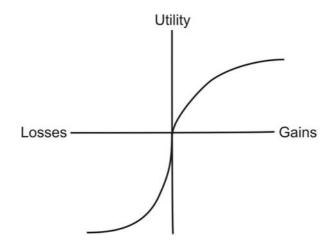


Figure 2. Utility function as proposed in Kahneman and Tversky's (1979) prospect theory. The utility function is concave in the domain of gains, and convex in the domain of losses, and steeper for losses than for gains. The crossing of the axes corresponds to the reference point against which the prospects are contrasted, for example, the current level of wealth.

#### Contribution from Behavioral Ecology Risk-Sensitive Foraging Theory

The importance of the utility function is to relate subjective to objective value. This is useful because subjective value is often not a linear function of objective value. However, what causes this non-linearity? Bernoulli (1954) suggested that wealth renders the utility function concave. Behavioral ecology suggests additional factors such as upper boundaries on how much energy can be stored and lower boundaries on how much energy is needed for survival can also introduce curvature or even "kinks" into the utility function. Consider for example birds foraging for the night. Because of their small body size and high metabolic rate they face the possibility of starving over night if they fail to accumulate enough resources during the day. As a consequence, a normally risk averse bird might become risk-prone towards the end of the day or at low temperatures, if its energy requirements for the night are not yet met (Caraco et al., 1990). Thus, in addition to varying over individuals, risk attitudes vary also over situations and time.

Energy requirements change not only over the course of a single day but also over the year, for example in the pre-migratory period of migratory birds (Moore and Simm, 1986). In that period birds must acquire sufficient reserves for the migration and they behave in a risk-prone manner until they reach maximal body size. Conversely, birds that are not in the pre-migratory period or pre-migratory birds that have reached maximal body size avoid risky foraging options.

Risk-sensitive foraging theory considers decision optimality given reserve constraints. Stephens (1981) showed that it is optimal for birds on a negative energy budget to choose options with higher variance. Conversely, they should avoid risky options when the less risky options provide the birds with a mean rate of intake that exceeds the starvation threshold, i.e. they are on a positive energy budget. For example, a bird almost starved to death (negative energy budget) should avoid the certain option if the certain food amount is not sufficient to guarantee its survival (certain death), and the only chance to survive would be to obtain the large, risky reward (possible survival). On the other hand, a bird on a higher energy budget should be risk-averse if the certain option is sufficient to guarantee survival (certain survival), and the small food amount of the risky option would be insufficient for survival (possible death). Current models of risk-sensitive foraging incorporate the possibility of sequential choices and the costs of foraging (e.g. McNamara and Houston, 1986).

In addition, life-history trade-offs may have favoured the evolution of risk-attitude in a similar fashion (Wolf et al., 2007). Animals differ in behaviours that affect their future fitness, for example, in their exploration effort during foraging. Animals that put more emphasis on future than immediate fitness returns have higher expectations regarding their future reproductive success than others. Investment into the future only pays out if the animal survives until it is able to realize the upcoming opportunities. Because survival is thus important for the strategy to work out, evolution may have favoured the development of risk-aversion towards predators and aggressive conspecifics in individuals with high fitness/reproductive expectations who have more to lose, whereas individuals with low expectations who have little to lose should be less risk-averse.

In summary, the minimal energy requirement (and by extension the maximal reserves) and/or the life-history of an animal may provide an inflection point ("kink"), where the curvature of its utility function changes.

#### A Normative Framework for Inter-Temporal Decisions

The second type of decisions making discussed in this chapter concerns decisions between outcomes that can only be realized at different instants in the future: Inter-temporal decisions. Humans and non-human animals prefer immediate over delayed rewards (so-called temporal discounting): Provided that the costs for both options are identical, the preference for an immediate or a temporally remote expected outcome is a function of the value of the respective outcomes and their delays, i.e., the time until the outcomes can be realized (McDiarmid and Rilling, 1965; Rachlin and Green, 1972; Ainslie, 1975; Mazur, 1984, 1987, 1988; Grossbard and Mazur, 1986; Logue, 1988; Benzion et al., 1989; Green et al., 1994, 1997; Evenden and Ryan, 1996; Evenden, 1999; Frederick et al., 2002; Reynolds et al., 2002; Kalenscher et al., 2005, 2006a).

Inter-temporal decisions have been extensively studied in psychology and ecology, but were and still are also of great interest in economic models of choice. As with decisions under risk, many of the normative models in economics are based on several theoretical assumptions and theorems, including preference monotonicity, stationarity, and maximization of utility rate.

A fundamental assumption in rational choice theories is that preference orders should be consistent across time. Preference monotonicity and stationarity are directly related to this assumption. Monotonicity means that a prospect X1 that is preferred over another prospect X2 will also be assigned a higher utility than X2 as long as the utility function is monotonic. Monotonicity of time preference (Lancaster, 1963) holds that

$X(t_1) \ge X(t_2)$ , if, and only if, $t_2 \ge t_1$	(	5)	

This means that commodity X, available at timepoint  $t_1$ , will be preferred over X, available at timepoint  $t_2$ , only when  $t_2$  occurs later than  $t_1$ .

Stationarity is related to the axiom of monotonicity of time preference and posits that

If $X(t) \sim Y(t+\tau)$ then $X(s) \sim X(s+\tau)$	(6)

This means that if an agent is indifferent (~) between commodity X, delivered at timepoint *t*, and commodity Y, delivered at timepoint  $t+\tau$ , he would still be indifferent when X was delivered at timepoint *s* and Y at timepoint  $s+\tau$  (Strotz, 1955; Koopmans, 1960; Fishburn and Rubinstein, 1982). Indifference means that the utility of both options is identical, and thus the frequency of choosing A or B is about 50%. Thus, if both options were deferred by the same time interval, preference orders should be preserved. In other words, if you desire to receive \$10 in 5 days as much as receiving \$50 in 20 days, then you will still desire to receive \$10 in 15 days as much as receiving \$50 in 30 days, i.e., when both delays are prolonged by 10 days.

It was proposed that the discounting rate by which future commodities are delivered should be constant (Samuelson, 1937), for instance resulting in a linear or exponential discount function. Many theories, therefore, assumed exponential discounting (Lancaster, 1963; Fishburn and Rubinstein, 1982; Benzion et al., 1989; cf., Ainslie, 1975; Loewenstein, 1992; Fehr, 2002). Combining exponential discounting with stationarity yields (Lancaster, 1963):

$(A,t) \sim Ae^{-k(t-t_0)}$ (7)	$(A,t) \sim A e^{-k(t-t_0)}$	(7)
---------------------------------	------------------------------	-----

This expression states that a reward with the amount A, delivered at timepoint t, is equally valuable (~) as a reward amount A at  $t_0$  (i.e., now), exponentially discounted for the interval t- $t_0$ , with  $t_0$  referring to the present timepoint, and k being an individually different discount value. In other words, the utility of a future outcome can be expressed as an exponential function of the same outcome realized today.

#### Violation of Stationarity

As outlined above, stationarity predicts that the ranking of preferences between several future outcomes should be preserved when the choice outcomes are deferred by the same time interval. This has been investigated in an empirical study where human subjects chose between pairs of monetary rewards available after different delays (Green et al., 1994). Subjects preferred a small, short-delayed over a large, long-delayed reward, but their preference reversed away from the small towards the large reward when the delays to both

rewards were advanced by the same time interval. Notably, the prolongation of the delays resulted in a preference reversal even though the difference in the delays remained identical (Green et al., 1994). This finding therefore represents a violation of stationarity. Numerous other studies with human subjects (Ainslie, 1975; Logue, 1988; Benzion et al., 1989; Loewenstein, 1992; Kirby and Herrnstein, 1995; Green et al., 1997; Frederick, Loewenstein and O'Donoghue, 2002; McClure et al., 2004; Rohde, 2005), pigeons (Chung and Herrnstein, 1967; Rachlin and Green, 1972; Ainslie, 1974; Green et al., 1981) and rats (Ito and Asaki, 1982; Bennett, 2002) replicated and confirmed these results (cf., Kalenscher and Pennartz, *in preparation*). Thus, human and non-human animals systematically violate the crucial assumption of inter-temporal consistency of choice. Note that many studies in the animal literature do not defer both choice outcomes equally, but only one outcome is increasingly delayed, whereas the delay to the other outcome remains constant (see e.g. figure 3A). Preference reversals in those cases do not challenge stationarity, as changes in valuations would be predicted for the increasingly delayed outcome, but not the constant outcome.

The fact that human and animal subjects prefer the small, short-term reward over the large, delayed reward when the receipt to the small reward is near, but not when it is in the relatively far future, suggests that short-term rewards are discounted more steeply than long-term rewards. Such asymmetric discounting poses a strong challenge for the postulation of exponential discounting (Lancaster, 1963).

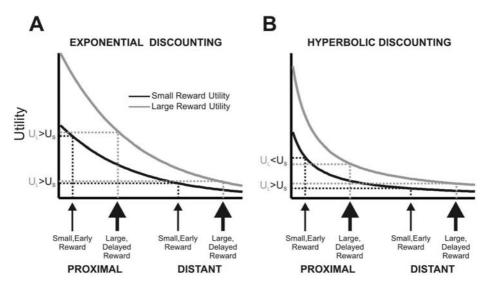


Figure 3. Exponential vs. hyperbolic discounting of future events. (A) Exponential utility curve of a large, delayed (grey line) and small, short-term reward (black line). With exponential discounting, stationarity holds because the utility of the large reward (UL) is always higher than the utility of the small reward (US). This is true when both rewards are temporally proximal, or when they are deferred by the same time interval into the future (distant rewards). (B) Hyperbolic discounting can explain preference reversals and the violation of stationarity. Due to the steeper decay for short delays, the utility of the small, short-term reward is higher than the large, delayed reward for temporally proximal rewards, but the utility order reverses when both rewards are deferred into the future.

Accordingly, as theoretically suggested by Ainslie (1975), and later empirically shown by Mazur (Mazur, 1984, 1987, 1988; Grossbard and Mazur, 1986) and others (Rachlin et al., 1991; Myerson and Green, 1995; Green and Myerson, 1996; Rohde, 2005; Jones and Rachlin,

2006; Laibson, 1997), discounting curves can be better approximated with hyperbolic than exponential or other constant discount functions, as outlined in figure 3. Why do humans and other animals systematically violate such crucial laws in economics?

#### Why Do We Discount the Future?

An assumption underlying most economic theories is utility maximization. This assumption is shared by many ecological theories of choice, namely the postulation that evolution favors choice mechanisms that maximize fitness levels, and minimize fitness losses. Applied to inter-temporal decision making, this means that the decision maker should act so as to maximize the utility rate, or in ecological terms, the energy intake rate per time unit (Stephens and Krebs, 1986). Rate maximization can explain why humans and animals sometimes prefer a less attractive, but temporally proximal outcome over a more attractive, but temporally remote outcome.

For example, we consider an inter-temporal choice task in which an animal has to choose between a small, always immediate reward and a large reward that is initially also delivered immediately, but that is delayed further as the experimental session progresses. Let's further assume that the large reward is 1.5 times as big as the small reward, and that the next choice opportunity always follows immediately after the animal has consumed its previous reward. The rate maximization hypothesis would predict that the animal should begin the session by preferring the large reward. However, the delay preceding the large reward gets longer and longer over the course of the session. So, at some point, the waiting time for the large reward, and thus the time until the next reward can be realized, will be more than twice as long as the delay preceding the small reward. Naturally, it would make sense now to prefer the small reward, as the hungry animal would be able to consume two small rewards in the same time that it would have to wait for only one large reward. Because two small rewards represent a larger food quantity than one large reward, the animal would maximize its energy intake per time unit by shifting its preference to the small reward once the delay preceding the large reward gets too long. In formal terms, optimal foraging theory assumes that organisms maximize, at least on the long run, the ratio of food intake and the time needed to obtain or consume the food, as described by the following quantity (Stephens and Krebs, 1986):

$\sum_{i=1}^{\infty}G_{i}$	
$\max \frac{i=1}{\infty}$	(8)
$\sum_{i=1} t_i$	

where  $G_i$  represents the net energy gain obtained from consuming the *i*th food item (here basically corresponding to its amount), and  $t_i$  represents the time between food item *i* and the previous food item *i*-1.

#### Animals Do Not (Always) Maximize Intake Rate

The above example implies that the next choice opportunity follows immediately after receipt of the reward. In many studies (Rachlin and Green, 1972; Ainslie, 1974; Grossbard and Mazur, 1986; Mazur, 1988; Evenden and Ryan, 1996; Cardinal et al., 2000; Isles et al., 2003, 2004; Winstanley et al., 2004, 2006; Kalenscher et al., 2005; Hwang et al., 2006; Louie and Glimcher, 2006), however, the inter-trial interval between reward and next choice was adjusted so that the total trial length was identical in all trials and independent of delay length and other factors. In such a scenario, the rate maximization hypothesis predicts that subjects should always choose the large reward, independent of the delay between response and reward, because only then would the animals maximize the total energy intake per trial, or per experimental session respectively. However, neither pigeons (Rachlin and Green, 1972; Ainslie, 1974; Grossbard and Mazur, 1986; Mazur, 1988; Kalenscher et al., 2005), nor rats (Evenden and Ryan, 1996; Cardinal et al., 2000; Winstanley et al., 2004, 2006; Roesch et al., 2006), mice (Isles et al., 2003, 2004), or monkeys (Hwang et al., 2006; Louie and Glimcher, 2006) show the predicted perseverance on the large reward alternative, but instead reverse their preference to the small, immediate reward once the large reward delay exceeds an individual threshold limit. This shows that the animals' choices depended on the waiting time preceding the rewards, but not on the ratios of reward amount and duration between the rewards, as would be predicted from rate maximization.

In fact, rate maximization models predict that amount and/or delay variations shouldn't play any role in the animals' decisions, because the choices should be only and exclusively directed towards maximizing the rate on the long-term. If, for example, an animal chooses between a fixed medium-term reward and or a variable-delay reward with either short or long delays (variable interval schedule), animals should always choose the option yielding the higher average reward rate. If the average reward rate is identical, animal should be indifferent between both options. However, contrary to this prediction, they usually prefer variable-interval over fixed schedules (Kacelnik and Bateson, 1996), indicating that delay variance does influence an animal's reward preference in addition to other factors, such as reward rate. This variance-proneness is interesting as animals are usually variance-averse if reward magnitude, and not delay, is variable, as explained above. Proneness to delay variance can be explained with hyperbolic discounting (see below).

In summary, animals do not make their choices according to the predictions of rate maximization models. They seem to employ rather short-sighted, waiting-time sensitive choice heuristics, and have a preference for delay variability.

#### Preference for Delay Variability

Hyperbolic discounting, as outlined in figure 3B, can explain the preference for variable over fixed interval schedules. Since, due to the hyperbolic decay, the utility of short-term rewards is disproportionally higher than the utility of medium-term or delayed rewards, but the difference in utility of medium-term and delayed rewards is negligible, the average expected utility of short-term and delayed rewards (variable interval schedules) will exceed the expected utility of fixed medium-term rewards. Hence, animals should prefer variable over fixed delays.

An alternative hypothesis, scalar expectancy theory (SET), can account for both variance aversion when reward magnitude is variable, and variance proneness when reward delay is variable. SET refers to the subjective time and magnitude representation which is normal around the actual means, but as a consequence of Weber's law, has a constant coefficient of variation (ratio of standard deviation to mean). Thus, the combination of an early and a late distribution results in a positively skewed integral, which explains preference for variation in delay (Reboreda and Kacelnik, 1991). Evidence for or against either SET or the hyperbolic discounting account is equivocal (Kacelnik an Bateson, 1997; Bateson and Kacelnik, 1998) and needs to be further tested in the future.

#### Ecological Models of Inter-Temporal Decisions: Ecological Rationality

In addition to the unclear support, neither of these accounts can explain why animals developed delay sensitivity in the first place: Why does evolution favor choice heuristics that produce suboptimal results in many cases by over-emphasising the delay to the next reward (e.g., through hyperbolic discounting), and ignoring the long-term relevance of time/amount sequences? Obviously, animals may equate delay with collection risk, as outlined in greater detail below. If delays are mentally treated as risks, a risk-averse animal will naturally avoid long delays. However, this doesn't provide an acceptable answer because the question remains why evolution has favoured suboptimal decision rules, be they related to risk avoidance or delay aversion. The first answer that comes to mind is that short-sighted rules have higher fitness values than long-sighted rules because the animals' constitutions do not allow them to tolerate too long waiting periods. For example, animals with a high metabolism cannot afford to wait too long for a large amount of food, or, in other words, what is the use of high quality, high amount of food if the animal has starved to death while waiting for it? Thus, short-sighted rules may have a certain evolutionary advantage over long-sighted rules.

This would certainly hold if the waiting times were close to the animals' starvation thresholds. However, mice, rats, pigeons, monkeys or other animals used in inter-temporal choice experiments shift their preference away from the economically more advantageous reward even when the waiting time to the larger reward exceeds less than a few seconds, and not hours or days (cf., McDiarmid and Rilling, 1969). Certainly, all those animals would be able to survive longer waiting periods than just a few seconds without food, but nevertheless, they prefer the short-term option over the long-delayed option, even if the long-delayed reward is a multiple of the short-term reward. Such extremely myopic decision patterns are difficult to explain with a fitness advantage of faster available food items. Why does evolution favor such extremely myopic choice heuristics?

Bounded rationality or ecological models, such as the ecological rationality hypothesis (Stephens et al., 2004) claim that choice heuristics that fail to produce maximum fitness in artificial experimental settings do, in fact, perform well in more ecologically valid contexts. For example, Stephens and colleagues (2004) argued that a more ecologically valid choice context entails decisions about limited food resources. A typical decision would consist of whether to entirely exploit all food resources in a given food patch, or leave the food patch early before having consumed all resources, and search for a new patch. The difference between the patch situation and the standard inter-temporal choice task is that, in a standard inter-temporal choice task, an animal has a binary choice between a large, delayed or a small,

immediate reward, whereas in the patch situation, it chooses whether to continue to stay in a given patch, or to leave and search for a new patch. Figure 4 illustrates an inter-temporal choice situation, often also referred to as a 'self-control' task (4A), and a patch situation (4B).

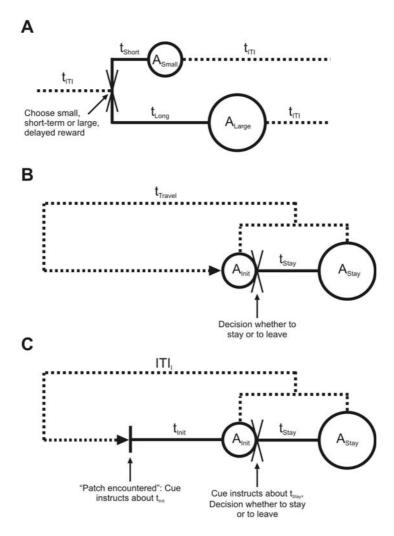


Figure 4. Inter-temporal choice and patch situation. (A) displays a schematic drawing of a standard inter-temporal choice task used in animal research, often also referred to as a 'self-control' task. Following an inter-trial interval (tITI), the animal makes a binary choice between a large and a small reward amount (ALarge and ASmall), delivered after a long or a short time delay (tLong and tShort). Reward consumption is followed by another ITI until the next choice opportunity. The length of the ITI, tITI, is adjusted to compensate for differences in delay length and choice, so that every trial is of identical duration. (B) Patch situation. An animal travels until it encounters a food patch, where it consumes an initial reward amount, AInit. After consumption, the animal has to decide whether to leave the patch and initiate travel time tTravel, until it finds the next patch, or whether to stay in the patch, wait time tStay for an additional food amount, AStay, and then leave the patch and initiate the travel time tTravel to figure 1B, but includes an additional initial waiting time, tInit, preceding the initial reward AInit. In all panels, the cross of thin lines indicates the timepoint where the animal makes its decision.

Put in more formalized terms, an animal consuming a reward in a food patch (the initial food amount,  $A_{Init}$ ) has to decide whether to stay in the current patch and wait for further rewards until the patch is completely depleted, or whether to leave early and initiate a new travel time to the next patch. If it decides to leave, it has to travel for time  $t_{Travel}$  until the next patch is encountered where it receives a new initial reward amount  $A_{Init}$ . If it decides to stay, it obtains additional food rewards of amount  $A_{Stay}$  delivered after a certain waiting time  $t_{Stay}$  until the patch is depleted. It then has to leave the patch as well, and initiate a further travelling time  $t_{Travel}$  until it encounters a new patch and obtains reward amount  $A_{Init}$  in the new patch (see figure 4B).

Staying is more time-consuming  $(t_{Stay}^{Total} = t_{Stay} + t_{Travel})$ , but yields higher reward amounts  $(A_{Stay}^{Total} = A_{Stay} + A_{Init})$ , leaving is less time-consuming  $(t_{Leave}^{Total} = t_{Travel})$  because the animal doesn't need to wait for the additional reward in the old patch, but it also misses out on that additional reward, and thus receives lower reward quantities  $(A_{Leave}^{Total} = A_{Init})$ . A farsighted decision rule based on rate maximization would predict that the animal prefers to leave if it gained more rewards per time unit in the leave alternative than in the stay alternative, and it would stay in case it gained more rewards per time in the stay alternative than in the leave alternative. Assuming that travelling to a new patch takes always the same time, and that the yield in a new patch is always identical, an animal would leave if

$$\frac{A_{Init}}{t_{Travel}} > \frac{(A_{Add} + A_{Init})}{(t_{Stay} + t_{Travel})}$$
(9)

and it would predict to stay in the opposite case.

A short-sighted, waiting-time sensitive rule, as observed in most experimental settings on inter-temporal decision-making, would predict that an animal considers only the delay until the next reward in its decision. That is, the rule would predict that an animal prefers to leave if

$\frac{A_{lnit}}{A_{lnit}} > \frac{A_{Add}}{A_{Add}}$	(10)
$t_{Travel}$ $t_{Stay}$	(10)

and it would predict to stay in the opposite case. Because travel time,  $t_{Travel}$ , and initial reward amount,  $A_{Init}$ , are identical in the leave and the stay case, the long-term, rate-maximising rule in expression (9) is algebraically equivalent to the short-term, waiting-time-sensitive, impulsive rule in expression (10). Thus, in the above described patch situation, a short-sighted impulsive choice rule (expression 10) would approximate long-term rate maximization, consistent with predictions from optimal foraging theory.

If this theoretical line of argument was true, then the same short-sighted choice heuristic should produce rate maximization in the patch-situation, but not in the standard intertemporal choice task. To test this prediction, Stephens and Anderson (2001) trained blue jays in a 'self-control' situation and a patch situation. The 'self-control' situation was essentially equivalent to the inter-temporal choice task sketched in figure 4A: Blue jays chose between a small, immediate or a large, delayed reward,  $A_{Small}$  and  $A_{Large}$ , delivered after  $t_{Short}$  and  $t_{Lorg}$ , by hopping on a perch on the left or right side in a training box. After reward delivery and consumption, they had to leave the perch, and an inter-trial interval (ITI, equivalent to travel time in figure 4B) was initiated, after which they could make their next choice. Instead of a binary choice between large and small rewards, the patch situation (figure 4C) consists of a sequence of choices between smaller rewards potentially summing up to a large reward: a cue instructed the animals about the delay  $t_{lnit}$  to an initial reward of  $A_{lnit}$  amount. By hopping on a randomly activated perch in the box, the initial reward was delivered after the indicated delay. Afterwards, the animals could choose whether to stay and wait for the additional reward by remaining on the perch (a second cue indicated the delay length  $t_{Stay}$  to the additional reward of amount  $A_{Stay}$ .), or whether to leave the perch, miss out on the additional reward, and initiate the ITI, and a new waiting time  $t_{lnit}$  to the next small reward. Hence, the patch task resembled the situation illustrated in figure 4B, with the exception that there was an additional waiting time  $t_{lnit}$  preceding the initial food reward  $A_{lnit}$ . Moreover, in the patch situation, the travel time needed to leave the initial patch and move to the next one corresponds to the ITI of the standard inter-temporal choice task. Note that defined this way, the ITI becomes an integral part of the decision in the patch situation but not in the standard choice situation. This is because a short-term decision rule taking into account only the delay to the next reward would predict that, in the 'self-control' situation, animals consider the delays to the large and the small reward ( $t_{Short}$  versus  $t_{Lone}$ ), and that, in the patch situation, they consider the delays to the different rewards in the stay or leave cases  $(t_{Stay}$  vs.  $t_{ITI} + t_{Init})$ .

It is now possible to choose the task parameters so that the 'self-control' and the patch situation are economically equivalent: The sum of the two rewards in the patch situation equals the large reward amount in the 'self-control' situation, and the sum of the delays to the first and second rewards in the patch situation is equivalent to the delay preceding the large reward in the 'self-control' situation. This means that animals would receive the same amount of food within the same time when they chose to stay in the patch situation or when they chose the large, delayed reward in the 'self-control' situation

Because of this economical equivalence, the animals should always show identical preferences across both situations if their decision rule was purely economical, e.g., farsighted. If, on the other hand, the animals indeed used a short-sighted decision rule, then they should show inconsistent choices under most circumstances, and should maximize their intake rate in the patch, but not in the 'self-control' situation. Such inconsistencies would arise because, first, the ITI is an integral part of the decision in the patch situation (remember that in the leave case, the total waiting time for the next reward would be  $t_{Init}^{Leave} + ITI$ ), but not in the 'self-control' situation (the delay between choice and reward-delivery is not dependent on the ITI), and, second, because the large reward in the patch situation consists of a series of smaller rewards, and not of a one-shot delivery of one single large reward as in the 'self-control' situation. In their study, Stephens and Anderson (2001) systematically varied the delays, amounts and ITIs, and tested the preference patterns of their blue jays. As predicted, they reported inconsistencies in choice between the two situations. In particular, they found that the blue jays were overall more ready to maximize their intake rate in the patch situation than in the 'self-control' situation. This supports the main notion of the ecological rationality hypothesis that one and the same short-sighted decision rule results in rate maximization in one, but not the other choice situation.

In summary, theoretical considerations and empirical evidence suggests that evolution may have favored the development of short-sighted choice heuristics because such rules produce long-term rate maximization in many natural situations, but not necessarily in artificial laboratory settings.

#### Ecological Models of Inter-Temporal Decisions: Feeding Ecology

Although all animals have in common that they discount the future, the rate by which future events are discounted differs dramatically between species. Mice, for example, seem to tolerate waiting times up to only a few seconds (Isles et al., 2003, 2004), capuchin monkeys wait for several minutes (Ramseyer et al., 2006), and humans can wait for months or even years for a relatively attractive reward (Green et al., 1994, 1997). The ecological rationality hypothesis can explain why evolution has favored the development of impulsive decision rules, but it cannot readily account for these large inter-individual and inter-species differences in delay-tolerance.

Another theory, the feeding ecology hypothesis, aims to explain those differences. It departs from the comparison of discount rates between very similar monkey species: Cottontop tamarins and common marmosets. These two new-world monkeys are similar in terms of social behavior, mating system, life span, life history, home range size, parental care, body size and weight, brain size and weight, and other factors (cf., Stevens et al., 2005a). However, despite those similarities, the animals show very different choice behavior when tested in an adjusting delay procedure. In an adjusting delay procedure (Mazur, 1987), animals choose between a small, short-term reward and a large, delayed reward. After large reward choices, the delay to its receipt is increased in the next trials, after small reward choices, the delay preceding the large reward is decreased. This procedure allows the measurement of indifference points, i.e., the delay length at which the large, delayed reward has equal value to the small, short-term reward. Stevens and colleagues (2005a) found that the marmosets waited considerably longer for a large reward than tamarins. However, when tested in a spatial version of the adjusting feature task, in which travelling distance, but not delay, to a large reward was varied, the pattern reversed: Whereas the tamarins preferred the large reward independent of the travel distance to the large reward, the marmosets preference for the large reward continuously decreased with increasing travel distance (Stevens et al., 2005b). Taken together, space and time affected the monkeys' decisions differentially: compared to tamarins, marmosets were more patient when waiting for a delayed reward, but discounted spatially distant rewards steeper than tamarins.

How come that the two monkeys have evolved so different discounting patterns? Because of their striking similarity in many aspects, differences in metabolism, physical condition, starvation threshold or the like can be ruled out as possible explanations. Stevens et al. (2005a, 2005b) point out that one of the main differences between the two New World monkeys is their diet: Although both species eat fruits, marmosets additionally feed on plant exudates, such as gum and sap, and tamarins feed on insects.

Feeding ecology plays a major role in shaping cognitive and neural functions (e.g., Clayton and Krebs, 1995; Basil et al., 1996; Emery and Clayton, 2001). Accordingly, the differences in foraging behavior between the tamarins and marmosets may prove to be the key evolutionary pressure for the differential development of the temporal and spatial discount rates: Marmosets feed on localized, immobile food sources that do not require far-distance travels. But feeding on gum and sap requires to scratch the bark of the tree and then wait for the sap to exude. For the marmosets, it is therefore essential to be patient (in time), but not necessarily mobile (in space) in order to get most of the slowly exudating sap. Conversely, for the insectivore tamarins, it is crucial to be constantly alert, and react quickly in order not to miss any passing-by insects. Moreover, they feed on dispersed food sources and have to cover rather large territories to find insects. Therefore, in contrast to marmosets, tamarins must be quick and impulsive, and willing to travel relatively far distances to find enough food to survive.

In summary, given the individual differences in foraging behavior, it may be more advantageous for the marmosets to be patient in time, but impulsive in space, and for the tamarins to be impulsive in time, but patient in space. The individual differences in foraging ecology may therefore explain the differential evolution of temporal and spatial discounting.

# Commonalities and Differences Between Risky and Inter-Temporal Decisions

The preceding parts of this chapter have treated inter-temporal and risky decisions as separate. However, several authors argued that there might not be a real difference between delay and risk because each dimension can be expressed in terms of the other (Mischel, 1966; Stevenson, 1986; Rachlin et al., 1986, 1987, 1991; Mazur, 1989, 1995, 1997; Green and Myerson, 1996, 2004; Sozou, 1998; Hayden and Platt, 2007). A delayed reward might be less likely to occur (at least in natural situations) and therefore its expected value might be lower than that of earlier rewards. Moreover, as the state of the agent might change, the value of a later reward might also be more uncertain due to the unpredictability of the subject's own state, including its own survival. For example, our annual mortality risk is about 1% but was considerably higher in our evolutionary past. There is little use in waiting for a large, delayed reward if we will never experience the reward. Therefore a delayed reward may be equivalent to a risky reward and decision-makers may equate temporal distance with collection-risk (Kacelnik and Bateson, 1996; Sozou, 1998).

Theoretically, the proposal that delay and risk are processed similarly boils down to models that incorporate only either risk attitude or discounting as subjective weighting factors for utility but not both. For example, consider the following model (Mazur, 2007):

$U = \Sigma P_i (A/I + KD_i) \tag{11}$	)
--	---

Here, utility is a function of objective probability (P) but subjective time discounting (K). A denotes reward amount and D delay. Conversely, the following model comprises three subjective weighting factors determining utility (Kheramin et al., 2003, adapted):

$$U = \Sigma (1/(1 + Q/q_i) \times 1/(1 + Kd_i) \times 1/(1 + H\theta_i))$$
(12)

where q denotes reward amount, d reward delay and  $\theta$  odds against reward ( $\theta = [1/P_i] - 1$ ). The subjective weighting factors Q, K and H denote subjective sensitivity to reward magnitude, delay and probability, respectively. H>1 corresponds to risk aversion, H<1 to risk seeking, Note, that this model treats probability and delay similarly by using a hyperbolic form for both.

Conversely, instead of treating delayed rewards as uncertain, it has also been proposed that risky rewards may be treated as variably delayed rewards (Rachlin et al., 1986; Mazur, 1989; Hayden and Platt, 2007). For example, take a gamble between two options with equal expected values, but where one option yields a medium-sized, certain reward and the other one yields either a large or a small reward with a 50% chance each. Animals are usually not risk-neutral in those types of tasks (Kacelnik and Bateson, 1996, 1997; Bateson and Kacelnik, 1998; McCoy and Platt, 2005b; Hayden and Platt, 2007). Macaques, for example, generally prefer the risky over the certain option (McCoy and Platt, 2005b; Hayden and Platt, 2007). To explain this risk-proneness, it has been argued that, if an animal consistently sticks with the risky option offering a 50% chance of a large pay-off, they will almost certainly receive the large reward eventually: If not now, then on a future trial. Therefore, the risky option gives a practically guaranteed, though potentially delayed large pay-off. Thus, because probabilistic rewards may be treated as large and delayed rewards, they may recruit similar cognitive mechanisms (Rachlin et al., 1986; Mazur, 1989; Hayden and Platt, 2007).

The empirical evidence supporting commonalities of delay and risk sensitivity is equivocal (Mazur, 1989; Rachlin et al., 1986, 1991; Green and Myerson, 1996, 2004; Estle et al., 2006; Hayden and Platt, 2007). Commonalities are entertained by the occurrence of similar preference reversals when the delay or the probability of reward is increased for both options. Thus, violations of the independence axiom (Allais paradox) with probability are similar to violations of the stationarity axiom with delay (both described above). In other words, the utility of sooner and more probable rewards increases more than that of later and less probable rewards as reward immediacy and probability increase (reviewed in Green and Myerson, 2004). Accordingly, both probability and delay are amenable to hyperbolic discounting functions.

Conversely, there is considerable evidence that risk and delay are processed differentially. For example, humans discount smaller delayed rewards more steeply than larger delayed rewards but discount smaller probabilistic rewards less steeply than larger probabilistic rewards (Du et al., 2002; Estle et al., 2006). Inflation affects decisions involving delayed but not risky monetary rewards (Ostaszewski et al., 1998). Drug addiction appears to affect delay discounting more than risk processing (Reynolds et al., 2004) whereas problem gambling might have the opposite effect (Holt et al., 2003). Even culture appears to influence probability and delay differentially, with Japanese graduate students discounting probabilistic rewards more steeply and delayed rewards less steeply than Chinese students (Du et al., 2002). Further reinforcing the notion that delay differs from risk, earlier rewards may be

preferred for several reasons over later rewards over and above to the later reward being riskier or the future subjective state more uncertain (Kacelnik and Bateson, 1996):

- Earlier rewards can be put to use and earn compound interest (corresponding to offspring's offspring) before later rewards arrive.
- Waiting for a delayed reward may prevent an agent from pursuing other courses of action. This might diminish the value of delayed rewards.
- It might be easier to learn about action-reward and stimulus-reward contingencies with earlier rewards because at their arrival the mental representation of the causally relevant antecedents has decayed to a lesser degree compared with later rewards.
- With long delays to the later reward and fixed intertrial duration, choosing earlier rewards can maximize the energy intake per unit time even if the later reward is larger than the earlier reward (as explained above on the chapter on rate maximization).

# CONCLUSION

In the preceding sub-chapters, we have outlined a selection of different attempts from different disciplines to explain decisions under risk and inter-temporal choices. Broadly speaking, normative approaches, such as EUT or optimal foraging (rate maximization), focus on how an animal should behave in order to meet formulated choice criteria, such as optimal decision-making, utility maximization or consistency of choice. Descriptive and empirical approaches, on the other hand, challenge many of the predictions and implications of normative models, and show that the normative analysis of decision making may not always be consistent with the empirical reality of choosing and acting. For example, humans do not always choose according to the predictions of EUT, and, contrary to the prediction of optimal foraging theory, animals frequently fail to maximize their intake rate. Ecological models, such as the budget rule or the ecological rationality hypothesis, deal with the question why evolution favored the development of choice patterns that often violate the predictions of the normative approaches. In particular, they provide an analysis within an ecologically valid framework of the sense and non-sense of the way animals (and humans) make their decisions. Last but not least, psychological approaches attempt to identify the choice-mediating cognitive mechanisms, such as whether animals employ far-sighted vs. myopic choice heuristics.

We hope to have illustrated that these different approaches have strong limitations when isolated from each other. In particular, we believe that normative models have high explanatory power, but are of questionable validity if not substantiated with empirical results. Empirical studies are of potentially higher validity per se, but it is difficult, if not impossible in many cases, to generalize single experimental results to a common framework of choice. They are therefore often of little use to answer the question how we generally make decisions when presented outside a theoretical context. Moreover, although empirical studies can be used to identify the short-comings of theoretical models, such de-construction is only useful when followed by the formulation of a better theoretical model. In conclusion, neither normative, nor empirical, nor psychological approaches alone can produce useful results, but only the combination of all approaches yields fruitful outcomes.

#### ACKNOWLEDGMENTS

We greatly appreciate the comments of Martin Kocher on an earlier version of the manuscript. TK is supported by a grant from SenterNovem (BSIK-03053).

## REFERENCES

Ainslie G (1974) Impulse control in pigeons. J. Exp. Anal. Behav. 21:485-489.

- Ainslie G (1975) Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychol. Bull.* 82:463-496.
- Allais M (1953) Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine. *Econometrica* 21:503-546.
- Basil JA, Kamil AC, Balda RP, Fite KV (1996) Differences in hippocampal volume among food storing corvids. *Brain Behav. Evol.* 47:156-164.
- Bateson M, Kacelnik A (1998) Risk-sensitive foraging: decision-making in variable environments. In: *Cognitive Ecology* (Dukas R, ed), pp 297-341. Chicago, Illinois: University of Chicago Press.
- Bennett SM (2002) Preference reversal and the estimation of indifference points using a fastadjusting delay procedure with rats. Dissertation, University of Florida.
- Benzion U, Rapoport A, Yagil J (1989) Discount rates inferred from decisions: An experimental study. *Management Sci.* 35:270-284.
- Bernoulli D (1954) Exposition of a new theory on the measurement of risk. *Econometrica* 22:23-36.
- Caraco T, Blanckenhorn WU, Gregory GM, Newman JA, Recer GM, Zwicker SM (1990) Risk-sensitivity: Ambient temperature affects foraging choice. *Anim. Behav.* 39:338-345.
- Cardinal RN, Robbins TW, Everitt BJ (2000) The effects of d-amphetamine, chlordiazepoxide, alpha-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology* (Berl) 152:362-375.
- Charnov EL (1976) Optimal foraging: The marginal value theorem. *Theor. Popul. Biol.* 9:129-136.
- Chung SH, Herrnstein RJ (1967) Choice and delay of reinforcement. J. Exp. Anal. Behav. 10:67-74.
- Clayton NS, Krebs JR (1995) Memory in food-storing birds: from behaviour to brain. *Curr. Opin. Neurobiol.* 5:149-154.
- Du W, Green L, Myerson J (2002) Cross-cultural comparisons of discounting delayed and probabilistic rewards. *Psychol. Rec.* 52:479–492.
- Ellsberg D (1961) Risk, ambiguity, and the Savage axioms. Q. J. Econ. 75:643-669.
- Emery NJ, Clayton NS (2001) Effects of experience and social context on prospective caching strategies by scrub jays. *Nature* 414:443-446.

- Estle SJ, Green L, Myerson J, Holt DD (2006) Differential effects of amount on temporal and probability discounting of gains and losses. *Mem. Cognit.* 34:914-928.
- Evenden JL (1999) Varieties of impulsivity. Psychopharmacology (Berl) 146:348-361.
- Evenden JL, Ryan CN (1996) The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology* (Berl) 128:161-170.
- Fehr E (2002) The economics of impatience. Nature 415:269-272.
- Fishburn PC (1982) Nontransitive measurable utility. J. Math. Psychol. 26:31-67.
- Fishburn PC, Rubinstein A (1982) Time Preference. Internat Econ. Rev. 23:677-694.
- Frederick S, Loewenstein G, O'Donoghue T (2002) Time discounting and time preference: a critical review. *J. Econ. Lit.* 40:351-401.
- Friedman M, Savage LJ (1948) The utility analysis of choices involving risk. *J. Polit. Econ.* 56:279-304.
- Glimcher PW (2002) Decisions, decisions; choosing a biological science of choice. *Neuron* 36:323-332.
- Glimcher PW, Rustichini A (2004) Neuroeconomics: the consilience of brain and decision. *Science* 306:447-452.
- Green L, Fisher EB Jr, Perlow S, Sherman L (1981) Preference reversal and self-control: Choice as a function of reward amount and delay. *Behav. Anal. Lett.* 1:43-51.
- Green L, Fristoe N, Myerson J (1994) Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonom. Bull. Rev.* 1:383-389.
- Green L, Myerson J (1996) Exponential versus hyperbolic discounting of delayed outcomes: Risk and waiting time. *Amer. Zool.* 36:496-505.
- Green L, Myerson J (2004) A discounting framework for choice with delayed and probabilistic rewards. *Psychol. Bull.* 130:769-792.
- Green L, Myerson J, McFadden E (1997) Rate of temporal discounting decreases with amount of reward. *Mem. Cognit.* 25:715-723.
- Grossbard CL, Mazur JE (1986) A comparison of delays and ratio requirements in selfcontrol choice. J. Exp. Anal. Behav. 45:305-315.
- Hamilton W D (1964) The genetical evolution of social behaviour. J. Theor. Biol. 7:1-52.
- Hayden BY, Platt ML (2007) Temporal discounting predicts risk sensitivity in rhesus macaques. *Curr. Biol.* 17:49-53.
- Holt DD, Green L, Myerson J (2003) Is discounting impulsive? Evidence from temporal and probability discounting in gambling and nongambling college students. *Behav. Processes* 64:355–367.
- Hwang J, Kim S, Lee D (2006) Neuronal signals related to delayed reward and its discounted value in the macaque dorsolateral prefrontal cortex. *Annual Meeting of the Society for Neuroscience*, Program No. 71.7.
- Isles AR, Humby T, Walters E, Wilkinson LS (2004) Common genetic effects on variation in impulsivity and activity in mice. *J. Neurosci.* 24:6733-6740.
- Isles AR, Humby T, Wilkinson LS (2003) Measuring impulsivity in mice using a novel operant delayed reinforcement task: effects of behavioural manipulations and damphetamine. *Psychopharmacology* (Berl) 170:376-382.
- Ito M, Asaki K (1982) Choice behavior of rats in a concurrent-chains schedule: amount and delay of reinforcement. *J. Exp. Anal. Behav.* 37:383-392.
- Jones, B. and Rachlin, H. (2006) Social discounting. Psychol. Sci. 17, 283-286.

- Kacelnik A (1997) Normative and descriptive models of decision making: time discounting and risk sensitivity. *Ciba Found Symp* 208:51-67; discussion 67-70.
- Kacelnik A, Bateson M (1996) Risky theories the effects of variance on foraging decisions. *Amer. Zool.* 36:402-434.
- Kacelnik A, Bateson M (1997) Risk-sensitivity: crossroads for theories of decision-making. *Trends Cogn. Sci.* 1:304-309.
- Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47:263-291.
- Kalenscher T, Ohmann T, Gunturkun O (2006) The neuroscience of impulsive and selfcontrolled decisions. Int. J. Psychophysiol. 62:203-211.
- Kalenscher T, Windmann S, Diekamp B, Rose J, Güntürkün O, Colombo M (2005) Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. *Curr. Biol.* 15:594-602.
- Kalenscher T, Pennartz CMA (2007) A bird in the hand is worth two in the future. *The neuroeconomics of inter-temporal decision making*. In preparation.
- Kheramin S, Body S, Ho M, Velazquez-Martinez DN, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2003) Role of the orbital prefrontal cortex in choice between delayed and uncertain reinforcers: a quantitative analysis. *Behav. Processes.* 64:239-250.
- Kirby KN, Herrnstein RJ (1995) Preference reversals due to myopic discounting of delayed reward. *Psychol. Sci.* 6:83-89.
- Koopmans TC (1960) Stationary ordinal utility and impatience. Econometrica 28:287-309.
- Laibson DI (1997) Golden eggs and hyperbolic discounting. Q. J. Econ. 42: 861-871.
- Lancaster K (1963) An axiomatic theory of consumer time preference. *Internat. Econ. Rev.* 4:221-231.
- Loewenstein G (1992) The fall and rise of psychological explanations in the economics of intertemporal choice. In: *Choice over time*, pp. 3-34. Eds G. Loewenstein, J. Elster. Russell Sage Foundation: New York.
- Logue AW (1988) Research on self-control: An integrating framework. *Behav. Brain. Sci.* 11:665-709.
- Loomes G, Starmer C, Sugden R (1991) Observing violations of transitivity by experimental methods. *Econometrica* 59:425-439.
- Loomes G, Sugden R, (1982) Regret theory: An alternative theory of rational choice under uncertainty. *Econ. J.* 92:805-824.
- Louie K, Glimcher PW (2006) Temporal discounting activity in monkey parietal neurons during intertemporal choice. *Annual Meeting of the Society for Neuroscience*, Program No. 605.5.
- Machina MJ (1982) "Expected utility" analysis without the independence axiom. *Econometrica* 50:277-323.
- Mankiw, NG (2002) Macroeconomics. 5th edition, Worth Publishers: New York.
- Mazur JE (1984) Tests of an equivalence rule for fixed and variable reinforcer delays. J. Exp. Psychol: Anim. Behav. Process 10:426-436.
- Mazur JE (1988) Estimation of indifference points with an adjusting-delay procedure. J. Exp. Anal. Behav. 49:37-47.
- Mazur JE (1989) Theories of probabilistic reinforcement. J. Exp. Anal. Behav. 51:87-99.
- Mazur JE (1995) Conditioned reinforcement and choice with delayed and uncertain primary reinforcers. J. Exp. Anal. Behav. 63:139–150.

- Mazur JE (1997) Choice, delay, probability, and conditioned reinforcement. *Anim. Learn. Behav.* 25:131–147.
- Mazur JE (2007) Species differences between rats and pigeons in choices with probabilistic and delayed reinforcers. *Behav. Processes* 75:220-224.
- Mazur JE, Commons ML, Mazur JE, Nevin JA, Rachlin H (1987) *An adjusting procedure for studying delayed reinforcement*. In, pp 55-73. Hillsdale, NJ: Erlbaum.
- McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004) Separate neural systems value immediate and delayed monetary rewards. *Science* 306:503-507.
- McCoy AN, Platt ML (2005a) Expectations and outcomes: decision-making in the primate brain. J. Comp. Physiol. A Neuroethol. Sens Neural. Behav. Physiol. 191:201-211.
- McCoy AN, Platt ML (2005b) Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* 8:1220-1227.
- McDiarmid CG, Rilling ME (1965) Reinforcement delay and reinforcement rate as determinants of schedule preference. *Psychon. Sci.* 2:195-196.
- McNamara JM, Houston AI (1986) The common currency for behavioral decisions. *Am. Nat.* 127:358-378.
- Mischel W (1966). Theory and research on the antecedents of self-imposed delay of reward. In BA Maher (Ed.), *Progress in experimental personality research*: Vol. 3 (pp. 85–132). New York: Academic Press.
- Moore FR, Simm PA (1986) Risk-sensitive foraging by a migratory bird (Dendroica coronata). *Experientia* 42:1054-1056.
- Myerson J, Green L, (1995) Discounting of delayed rewards: Models of individual choice. J. *Exp Anal. Behav.* 64, 263-276.
- Ostaszewski P, Green L, Myerson J (1998) Effects of inflation on the subjective value of delayed and probabilistic rewards. *Psychon. Bull. Rev.* 5:324–333.
- Powell K (2003) Economy of the mind. PLoS Biol 1:E77.
- Rachlin H, Castrogiovanni A, Cross D (1987) Probability and delay in commitment. J. Exp. Anal. Behav. 48:347–353.
- Rachlin H, Green L (1972) Commitment, choice and self control. J. Exp. Anal. Behav. 17:15-22.
- Rachlin H, Logue AW, Gibbon J, Frankel M (1986) Cognition and behavior in studies of choice. *Psychol. Rev.* 93:33-45.
- Rachlin H, Raineri A, Cross D (1991) Subjective probability and delay. J. Exp. Anal. Behav. 55:233-244.
- Ramseyer A, Pele M, Dufour V, Chauvin C, Thierry B (2006) Accepting loss: the temporal limits of reciprocity in brown capuchin monkeys. *Proc. Biol. Sci.* 273:179-184.
- Reboreda JC, Kacelnik A (1991) Risk sensitivity in starlings: Variability in food amount and food delay. *Behav. Ecol.* 2:301-308.
- Reynolds B, de Wit H, Richards J (2002) Delay of gratification and delay discounting in rats. *Behav. Processes* 59:157-168.
- Reynolds B, Richards JB, Horn K, Karraker K (2004) Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behav. Processes*, 65:35–42.
- Rohde K (2005) The hyperbolic factor: a measure of decreasing impatience. Research memoranda 044. *Maastricht Research School of Economics of Technology and Organization: Maastricht.*
- Samuelson PA (1937) A note on measurement of utility. Rev. Econ. Stud. 4:155-161.

- Shafir S (1994) Intransitivity of preferences in honey bees: support for 'comparative' evaluation of foraging options. *Anim. Behav.* 48:55–67.
- Sibly RM, McFarland DJ (1976) On the fitness of behaviour sequences. Am. Nat. 110:601-617.
- Sozou PD (1998) On hyperbolic discounting and uncertain hazard rates. *Proc. R. Soc. Lond. B.* 265:2015-2020.
- Stephens DW (1981) The logic of risk-sensitive foraging preferences. *Anim. Behav.* 29:628-629.
- Stephens DW, Anderson D (2001) The adaptive value of preference for immediacy: when shortsighted rules have farsighted consequences. *Behav. Ecol.* 12:330-339.
- Stephens DW, Kerr B, Fernandez-Juricic E (2004) Impulsiveness without discounting: the ecological rationality hypothesis. *Proc. R. Soc. Lond B. Biol. Sci.* 271:2459-2465.
- Stephens DW, Krebs JR (1986) Foraging theory. *Monographs in behavior and ecology*. Princeton University Press: New Jersey.
- Stevens JR, Hallinan EV, Hauser M (2004) The ecology and evolution of patience in two New World monkeys. *Biol. Lett.* 1:223-226.
- Stevens JR, Rosati AG, Ross KR, Hauser MD (2005) Will travel for food: Spatial discounting in two new world monkeys. *Curr. Biol.* 15:1855-1860.
- Stevenson MK (1986) A discounting model for decisions with delayed positive and negative outcomes. J. Exp. Psychol: General 115:131–154.
- Strotz RH (1955) Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* 23:165-180.
- Tobler PN, Fletcher PC, Bullmore ET, Schultz W (2007a) Learning-related human brain activations reflecting individual finances. *Neuron*. 54:167-175.
- Tobler PN, O'Doherty JP, Dolan RJ, Schultz W (2007b) Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. J. Neurophysiol. 97:1621-1632.
- Tversky A (1969) Intransitivity of preferences. Psych. Rev. 76: 31-48.
- Tversky A, Kahneman D (1986) Rational choice and the framing of decisions. J. Business 59:S251-S278.
- Tversky A, Kahneman D (1992) Advances in prospect theory: Cumulative representation of uncertainty. J. Risk Uncert. 5:297–323.
- Von Neumann J, Morgenstern O (1944) *Theory of games and economic behavior*. Princeton. Princeton University Press: New Jersey.
- Waite TA (2001) Intransitive preferences in hoarding gray jays (Perisoreus canadensis). *Behav. Ecol. Sociobiol.* 50:116-121.
- Winstanley CA, Theobald DE, Cardinal RN, Robbins TW (2004) Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. J. Neurosci. 24:4718-4722.
- Winstanley CA, Theobald DE, Dalley JW, Cardinal RN, Robbins TW (2006) Double dissociation between serotonergic and dopaminergic modulation of medial prefrontal and orbitofrontal cortex during a test of impulsive choice. *Cereb. Cortex* 16:106-114.
- Wolf M, van Doorn GS, Leimar O, Weissing FJ (2007) Life-history trade-offs favour the evolution of animal personalities. *Nature* 447:581-584.

Chapter 26

# FROM ABSTINENCE TO SAFER SEX: AMBIGUITIES AND DIALECTICS\*

Donna E. Howard, Robin Sawyer and Jessica Brewster-Jordan

Department of Public and Community Health University of Maryland, College Park, MD USA

# ABSTRACT

Despite a steady decrease in teen pregnancy, birth rates and abortion since 1991, adolescents in the United States continue to experience unintended pregnancy at higher rates than their peers in nearly all other industrialized nations and are at high risk for Sexually Transmitted Infections (STIs). At the same time, debate rages over whether comprehensive sexuality education or abstinence-only curricula should be taught in schools. Additionally, there is growing concern, among adolescent sexual behavior researchers, that our limited ability to affect changes in sexually risky behavior may be related to discordance between the survey questions pertaining to adolescent sexual behavior and actual youth practices. That is, research suggests that a significant proportion of data purporting to accurately gauge adolescent participation in specific acts of sexual behavior may be not only incomplete but flawed. This chapter will review the debate over sexuality education, examine the language used on current national health surveys that assess the sexual behavior of adolescents and young adults and provide recent findings from an exploratory study which examined the denotative meaning of sexual terms among a convenience sample of late adolescent university students. In this study, it was found that there still exists little consensus among adolescents about the behavioral referents for the terms *safe* and *safer* sex. Information in this chapter will help inform the development of rigorous investigations to further examine ambiguities surrounding the language and practice of sexual behavior among adolescents and consider how this imprecision affects policy and programmatic decision making.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

# INTRODUCTION

# The Debate over Sexuality Education

Although birth rates and abortion have decreased in the United States since 1991, American adolescents continue to experience unintended pregnancy at higher rates than their peers in nearly all other industrialized nations (Guttmacher Institute, 2004). Although such reductions are obviously a positive development, recent data from the national Youth Risk Behavior Survey (YRBS, 2005) suggest that adolescents continue to be sexually active and engage in unprotected sexual behavior. According to a 2003 study of adolescent risk behavior, 33% of ninth graders reported that they had experienced sexual intercourse; the proportion of sexually active adolescents in the older grades were as follows: 44% of tenth graders, 53% of eleventh graders and 62% of twelfth graders (CDC, 2003). In addition to unplanned pregnancy, adolescent sexual behavior can lead to the acquisition and communicability of sexually transmitted infections (STIs). The Centers for Disease Control and Prevention (CDC) estimates the transmission of over 19 million STIs annually, with almost half occurring among young people, ages 15-24 (CDC, 2004). Recent estimates suggest the direct medical costs to diagnose and treat STIs in the United States is approaching 13 billion dollars annually (Chesson, Blandford, Gift, Tao, and Irwin, 2004). The consequences of early sexual activity are all too apparent with implications for both adolescents and the larger society. Outcomes associated with teen childbearing include lower educational achievement, higher rates of poverty, greater welfare use, lower rates of school completion, less stable employment, lower rates of marital stability and additional nonmarital births (Chase-Lansdale, 1995; Chesson, Blandford, Gift, Tao, and Irwin, 2004).

Despite the widespread recognition of serious health, psychosocial and economic consequences of adolescent sexual behavior there exists little consensus in the United States regarding the appropriate approach to this issue. The current battle being waged between proponents of abstinence-only-until-marriage education and advocates of a more comprehensive approach to sex education is indicative of other political schisms apparent in the United States today. Issues related to human sexuality, i.e., school-based sex education, gay marriage, mandatory HPV (Human Papilloma Virus) vaccination, stem cell research, and emergency contraception (the "morning after pill") have become highly politicized and polarized. Furthermore, they have become lightning rods for political wrangling over issues of religion, morality, and ethics.

Attempting to restrict comprehensive sexuality education is not new, and in fact, began in 1981. In that year the Office of Population Affairs began administering the Adolescent Family Life Act (AFLA) that was designed to prevent teen pregnancy by advocating chastity and self-discipline (Saul, 1998). In 1993, as a result of a suit by the American Civil Liberties Union, AFLA programs were forced to adjust their curricula to: not include religious references, be medically accurate, and to respect the principle of self-determination regarding contraceptive referral for teenagers (Saul, 1998). The AFLA continues to fund abstinence-only programs today, despite two decades of programming without rigorous evidence to suggest that the programs are effective at changing adolescent sexual behavior (Bartels et al., 1996; Kirby, 2002). The abstinence-until-marriage only movement began to flourish in 1996, when the federal government attached a provision to a welfare reform law establishing a

federal entitlement program for abstinence-only-until-marriage education. This program, Section 510(b) of Title V of the Social Security Act appropriated 250 million dollars over five years for state initiatives promoting abstinence-only sex education (Social Security Act, 1996). To ensure that the abstinence-only message was not diluted, the law (PL 104-193, Section 510 of the Social Security Act) stipulated that the term "abstinence education" refers to an educational or motivational program that includes the following eight elements:

- has as its exclusive purpose, teaching the social, psychological, and health gains of abstaining from sexual activity;
- (2) teaches abstinence from sexual activity outside of marriage as the expected standard for all school-age children;
- (3) teaches that abstinence is the only certain way to avoid out-of-wedlock pregnancy, STDs, and associated health problems;
- (4) teaches that a mutually faithful monogamous relationship within marriage is the expected standard of human sexual activity;
- (5) teaches that sexual activity outside of marriage is likely to have harmful psychological and physical effects;
- (6) teaches that bearing children out-of-wedlock is likely to have harmful consequences for the child, the child's parents, and society;
- (7) teaches young people how to reject sexual advances and how alcohol and drug use increases vulnerability to sexual advances;
- (8) teaches the importance of attaining self-sufficiency before engaging in sex. (Section 510, p. 1).

Regardless of the requirement's lengthy and overtly moralistic undertone, the wording fails to accurately define the term *abstinence*. This obvious flaw creates methodologic and pragmatic issues when attempting to evaluate the efficacy of abstinence-only programming. How can researchers truly evaluate what has not been defined? Indeed, there is a paucity of existing research examining the effects of the abstinence-only approach to sexuality education. In 2002, Advocates for Youth performed an evaluation of findings from ten states that were implementing abstinence-only sex education. In addition to definitional problems, the evaluation was complicated by poor research design and inconsistent statistical analysis. Their review was unable to establish any positive outcomes for the abstinence-only programs (Kirby, 2002). In 2005, the Department of Health and Human Services published a report on the first year impact of abstinence among students participating in the program, there was no available evidence of a change in intention to abstain from sexual activity (Mathematica Policy Research, 2005).

The Department of Health and Human Services attempted to refine its definition of abstinence, when in June of 2006, a notice of federal grant availability saw the abstinence education themes expand from eight to thirteen and, for the first time, the terms abstinence and sexual activity were defined: "Abstinence means voluntarily choosing not to engage in sexual activity until marriage. Sexual activity refers to any type of genital contact or sexual stimulation between two persons including, but not limited to, sexual intercourse" (Administration for Children and Families, 2006). This more specific definition of abstinence may have been intended as a response to data suggesting that adolescents were turning to

"alternate" sexual activity such as oral or anal sex to maintain their abstinent status (Lewin, 1997; Stepp, 1999). In 2006, 46 states applied for the abstinence education funds, with only California, Maine, New Jersey and Pennsylvania rejecting the opportunity (Kaiser, 2007).

Newly revised federal guidelines for the 2007 abstinence education program included the expansion of the program's scope to include unmarried adults up to age 29 (Health and Human Services, 2007). The logic for this quite dramatic increase in target audience seems to rest on the fact that the age group most likely to bear children out of wedlock is 19-29 year olds; thus, the Department of Health and Human Services considers this group a prime target for abstinence education. Once again, alarm has been raised that common sense is being trumped by extra-scientific, political motives rooted in religious doctrinaire. Attempting to proselytize abstinence to an age group in which 90% of individuals in their late twenties are already sexually active may appear not only impractical but paternalistic. That is, this approach ignores the committed, monogamous couple who choose not to wed, it narrowly defines healthy sexuality and could be construed as a negation of sexual behavior that is not linked to procreation. Recently, the Society for Adolescent Medicine (SAM) published a position paper regarding their stance on abstinence-only education(Santelli, Ott, Lyon, Rogers, and Summers, 2006a). The SAM statement supported abstinence education, but not as the sole option; indeed they argued that "abstinence only" education lacked both ethical and scientific integrity. A comprehensive health education program, however, which includes, among other things, information on abstinence, contraceptives and protection from STIs, is endorsed by the SAM (Santelli, Ott, Lyon, Rogers, and Summers, 2006a; Santelli, Ott, Lyon, Rogers, Summers, and Schleifer, 2006b).

Disagreement exists not only among policy makers and scientists but between legislators and the constituencies that are affected by abstinence-only educational policy. For example, several nationwide polls provide evidence that a significant percentage of Americans support comprehensive sexuality education, rather than abstinence-only curriculum (Kaiser Family Foundation, 2004; Hickman-Brown Research, Inc., 1999). In fact, in a 2000 national poll conducted by Kaiser Family Foundation, the majority of parents, teachers, principals, and students were in favor of offering sexuality classes to high school students covering a variety of sexual behavior related topics (Kaiser Family Foundation, 2000). In a recent study in which parents of North Carolina youth were asked for feedback on the health education provided in their public schools (Ito, Gizlice, Owen-Dowd, Foust, Leone, and Miller, 2006), the viewpoints of the parents regarding comprehensive sexuality education did not correspond with the legislated abstinence curriculum for the state of North Carolina. That is, most parents were against legislators setting the sexuality education curriculum and thought parents and health professionals should determine the content of such courses. In addition, parents identified the following sexuality education topics as being important: transmission and prevention of STIs, HIV, and AIDS; how to handle the pressure to have sex; what to do if someone is sexual assaulted; and how to talk with partner about not having sex (Ito, Gizlice, Owen-Dowd, Foust, Leone, and Miller, 2006).

# An Examination of Sexual Health Behavior Terminology

As noted above, a core ambiguity surrounding abstinence education is the lack of an accepted definition for terms like abstinence and sex, and perhaps, more importantly, a total

lack of consensus among adolescents as to their interpretation of these terms. In 1998, the American public received an unexpected presidential primer on the definition of sex as President Clinton adamantly defended his "non-sexual" relationship with Monica Lewinsky. In the wake of this highly publicized event, a magazine survey of 15-19 year olds reported that 40% of respondents felt that oral sex was not "sex" (Seventeen, 2000), while a 2000 Internet survey indicated that 18% of 13-19 year-old girls believed that oral sex was something you did with your boyfriend before you were ready to have sex (Birnbaum, 2000).

Elsewhere, a group of Midwestern adolescents, who were receiving abstinence education, were asked how they interpreted the term abstinence. The students had great difficulty arriving at a definition and responses ranged from "kissing is OK" to "anything but intercourse" (Bell, 2000). Another study conducted with Midwestern adolescents, concluded that these individuals conceptualized abstinence as a "waiting period" and part of a continuum of natural development (Ott, Pfeiffer, and Fortenberry, 2006). In a 1998 study of over 1,000 college freshmen and sophomores, 61% of respondents considered mutual masturbation to be abstinence, 37% thought oral sex was also abstinence and 24% believed that anal intercourse was not sex (Horan, Phillips and Hagan, 1998). Even health professionals have shown little evidence of consensus with 30% of a sample of health educators responding that oral sex was abstinent behavior while 30% believed that mutual masturbation was not abstinence (Mercer, 1999). In a recent study where a four-item measure to assess sexual abstinence behavior was being developed, the researchers grappled with the definitional meaning of the term abstinence in the context of a person's virginity status (Norris, Clark and Magnus, 2003). These authors argued that those committed to sexual abstinence are more often virgins, yet virginity is not a prerequisite for sexual abstinence.

As the federal government came to understand, the promotion of abstinent behavior necessitates definition of the terms abstinence and sexual activity. A number of studies have been conducted to clarify the meaning of these terms among adolescents and young adults. A 1999 study of college students (Sanders and Reinisch, 1999) demonstrated that penile-vaginal intercourse was the only sexual activity that respondents could agree constituted "sex"; 60% reported that oral-genital contact was not considered as having "sex." To add further complexity to this definitional quandary, research suggests that adolescents' perceptions of what is considered "sex" often differs by contextual factors (Bogart, Cecil, Wagstaff, Pinkerton, and Abramson, 2000; Cecil, Bogart, Wagstaff, Pinkerton, and Abramson, 2002). Bogart et al. (2000) provided undergraduate students with a series of scenarios involving sexual behavior in an effort to glean their interpretation of the sexual acts performed between two hypothetical male and female actors. Over 90% of the sample felt both actors would perceive vaginal and anal intercourse as sex and the female actor was projected to identify a broader range of sexual behaviors to constitute sex. Cecil et al. (2002) also distributed sexual behavior items to college students to illicit feedback on what actions between the male and female actors signify their status as sexual partners. Couples who had engaged in vaginal or anal intercourse rather than oral sex were more often labeled as sexual partners. In addition, findings suggested that the frequency of the sex behavior and the dating status of the couple also played a in role in judgments regarding their status as sexual partners.

One of the most common items used on sexuality surveys is the term "sexual intercourse." Although most researchers may *infer* penis-vagina intercourse when they use this term, (Bell, 2000; National College Health Risk Behavior Survey, 1995; NBC/People, 2005) survey respondents may *perceive* the meaning in a different way. For example, among

a sample of undergraduate students asked to define the term, "sexual intercourse," 36% defined it as vaginal intercourse, 36% characterized it as including vaginal and anal intercourse, while 20% of the sample explicitly excluded oral intercourse from their definition (Sawyer, Howard, Brewster-Jordan, Gavin, and Sherman, 2007). Although the term, "vaginal intercourse" is slightly more specific than "sexual intercourse" possible ambiguity still exists. National behavioral risk surveillance systems routinely survey adolescents and young adults about a host of risk behaviors, including sexual risk (Youth Risk Behavior Survey, 2005). A common survey item reads, "Have you had sexual intercourse in the past three months?" (NBC/People, 2005). Without any clarifying definition, the difficulty with accurately interpreting responses to this question is all too obvious: does a "yes" response indicate penisvagina intercourse, oral intercourse, and/or anal intercourse? The ability to make meaningful inferences, let alone programmatic decisions based on responses to an ambiguous item simply asking about "sexual intercourse" appear to be highly compromised.

The concept of virginity has also been found to be highly ambiguous. There is a high degree of subjectivity among individuals in terms of what embodies the transition from virgin to non-virgin status. Carpenter (2001) explored virginity loss through in-depth interviews with 61 men and women ages 18 to 35. She found consensus among participants that first time vaginal-penile intercourse constituted a loss of virginity; however, discrepancies between other sexual acts that individuals could engage in and still be considered a virgin were apparent (i.e., oral sex, same-sex intercourse, anal sex). She concluded that virginity can be viewed alternatively as a gift, stigma, or part of an inevitable or advantageous rite of passage in life depending on other contextual factors discussed in the interviews. Elsewhere (Stevens-Simon, 2001), virginity and abstinence have each been described as possibly a "state of mind" rather than of body, in that many young people who report being abstinent or a virgin may in fact be sexually active. Clearly, further exploration into the denotative and connotative meaning of the term virginity is warranted in order to better understand its relationship to adolescent sexual beliefs, intentions and behaviors.

Another important area of research surrounds the study of oral sex or oral intercourse. During the past decade, sexuality researchers seem to have been so preoccupied with studies related to unintended pregnancy prevention and its concomitant focus on penis-vagina intercourse that research on other important sexual behaviors, including oral sex, has largely been ignored. Despite much anecdotal information in the popular press suggesting an increase in oral sex among adolescents (Stepp, 1999), there exists very little empirical data in peerreviewed research journals. This is despite the clinical findings linking oral sex to the transmission of STIs like Gonorrhea, Herpes, and Chlamydia (Hook, and Handsfield, 1999; Lafferty, Downey, Celum, and Wald, 2000; Kashima, Shah and Lyles, 1992). In a study of ninth-grade adolescents, more respondents reported having experienced oral sex (19.6%) than penis-vagina intercourse (13.5%), viewing oral sex as less risky (Schuster, Bell and Kanouse, 1996). A 2002 study reported that 55% of males and 54% of females aged 15-19 in the U.S. had engaged in oral sex, (Mosher, Chandra, Jones, 2005) while a United Kingdom study reported an almost identical rate of experience with oral sex in a population of students aged 16-18 years (Stone, Hatherall, Ingham, and McEachran, 2006). Notwithstanding the high prevalence of oral sex among adolescents, major surveys purported to monitor adolescent engagement in health-risk behaviors, such as the national Youth Risk Behavior Survey, fail to include a single item on this behavior. Developing a complete and useful picture of adolescent sexual behavior would seem difficult without the inclusion of items that examine

one of the most common adolescent sexual activities. However, recent research would seem to suggest that simply including an item that reads, "Have you had oral sex?" would be insufficient. A recent study found almost 25% of the sample disagreed as to the meaning of the term oral sex, leading the authors to conclude that more often the person performing oral sex was viewed as *not* having oral sex in comparison to the person on the receiving end (Sawyer, Howard, Brewster-Jordan, Gavin, and Sherman, 2007).

Recently Sawyer et. al., (2007) conducted an exploratory study to examine the denotative meaning of sexual terms in a convenience sample of late adolescent university students. Three anonymous surveys were developed to capture students' interpretations of the terms sexual intercourse, virgin, abstinence, and oral sex. Upon receiving Institutional Review Board approval from the university, questionnaires were distributed to students enrolled in three large survey courses offered by a department of public and community health.

#### Sexual Intercourse

Slightly over one-third of the sample (36%, n=117) defined sexual intercourse as meaning vaginal sex. A similar proportion of the sample interpreted sexual intercourse to include vaginal and/or anal and oral sex (35%, n=112). Just under 20% (n=63) of the sample did not include oral sex as part of the definition of sexual intercourse, but did include vaginal and/or anal sex. Only a small proportion of the sample specifically used the term penetration (6%, n=18) in their definition. A smaller percent perceived oral sex alone as synonymous with sexual intercourse (2%, n=5). There were gender differences regarding the designation of oral sex in that males were more likely than females to list oral sex as a sexual activity included in the term sexual intercourse; these differences, however, were non-significant.

## Virgin

Vaginal intercourse was the singular behavior that was most likely to negate a person's virginity status by both females and males. Five percent (n=16) of the sample included penetration as a behavior associated with loss of virginity. About 19% (n=62) also considered a virgin to be a person who has never participated in vaginal or anal sex. Another 14% (n=46) included oral sex, in addition to vaginal and/or anal sex, as an action that would terminate one's status as a virgin. There were no significant gender differences in the behaviors that would constitute a non-virgin identity.

## Abstinence

A large proportion of students indicated that if anal or vaginal *penetration* occurred, they did not consider themselves abstinent (anal 157, 86%, vaginal 174, 95%,); however, over a third of the sample recorded that anal or vaginal *touching* by a penis still signified abstinent behavior. The sample was nearly split as to whether performing or receiving oral sex, defined either as a tongue penetrating or touching another's vagina or penis penetrating or touching

another's mouth, could be practiced while still considering oneself abstinent. Sexual behaviors that could be practiced while still considering oneself abstinent displayed minimal gender variability with one exception. More males (14%) believed that if their penis penetrated a female's anus they could remain abstinent while only about 6% of females believed they could be abstinent if they were on the receiving end of penis-anal penetration.

#### Oral Sex

In this survey, university students were presented with the statement, "Jamie performs oral sex on Chris." The students were then asked to respond to the following questions: "In your opinion, has Jamie had oral sex? In your opinion has Chris had oral sex?" Although 76% of the sample felt that both individuals had experienced oral sex, fully 20% of respondents believed that Jamie (the person performing oral sex) had *not* experienced oral sex. In addition, 5% of respondents believed that the person receiving oral stimulation had not experienced oral sex (Sawyer, Howard, Brewster-Jordan, Gavin, and Sherman, 2007). That is, almost 25% (n=38) of the sample differentiated between "giving" and "receiving" when defining oral sex. In summary, the person on the receiving end was more often seen as having oral sex than was the person performing the oral sex.

These findings have implications for research, clearly bringing into question the accuracy and reliability of data gathered from questionnaires. When as many as one-fifth of a survey population responds "no" to an item regarding oral sex because they had performed and not received oral sex, then the interpretability of survey results are obviously questionable. The findings from this study suggest little consensus among adolescents about the behavioral referents of the term abstinence and related terms such as virginity, sexual intercourse, and oral sex. Regardless of whether elicitation of definitions occurred via open-ended or closedended probes, these terms appear to have highly personalized and often contradictory interpretations among college students, with one important exception. Findings from this study support previous research that has determined penis-vagina intercourse as the one sexual behavior that an individual cannot perform and still consider him or herself to be abstinent (Sanders and Reinisch, 1999; Bogart, Cecil, Wagstaff, Pinkerton, and Abramson, 2000; Pitts and Rahman, 2001). Penis-anus penetration was the only other sexual behavior that a large majority of the sample in this study believed could not be practiced for one to still be abstinent. Adding to the uncertainty surrounding the definition of terms such as abstinence, sex and virginity are issues related to the actual practice of oral sex, which was evident in the results reported by Sawyer et al., (2007). Many of the nationally representative adolescent surveys do not include items on oral sex, and those that do, only ask "Have you had oral sex?" without providing a definition for the term; indeed, the ambiguities surrounding its meaning may call into serious question the results obtained from such surveys (Youth Risk Behavior Survey, 2005; National College Risk Behavior Survey, 1995; American College Health Association College Health Assessment, 2003; Halpern-Fesher, Cornell, Kropp, and Tschann, 2005).

In summary, the need to accurately define what is meant by abstinence, and other related terms such as sexual intercourse and virginity is critically important on many levels. If adolescents are encouraged to be abstinent, then they need a precise definition of what that really means. Furthermore, if program planners and researchers are to implement and evaluate abstinence-only interventions, operational definitions must be developed that are robust. Unless survey items are constructed with precise definitions of the referenced sexual behavior terms, interpretation of data may be difficult and prevalence estimates of abstinence may be inflated. Inevitably, related terms like sexual intercourse and virginity will also need precise definitions as they are relevant to our understanding of abstinence. Furthermore, without such specificity, adolescents will continue to be confused about the nomenclature surrounding their sexual behavior and health professionals will have great difficulty evaluating sexuality data, in general, and the success of abstinence-only education, in particular.

The lack of consensus surrounding definitions of sexual behavior may well have a considerable impact on the ability of researchers to collect meaningful data. Although surveillance reports measuring the prevalence of unintended pregnancy and rates of STI are considered to be accurate, due in large measure to the clinical nature of the data and the generation of accurate medical records, adolescent sexuality data collected through large surveys with a self-report format may well be flawed or incomplete. Accuracy of data can only be achieved if the researcher and respondent have a common understanding of the definitional meaning of individual terms in survey questions.

## An Examination of Sexual Language Used on National Surveys

The ambiguities discussed in the prior sections raise a host of concerns regarding the sexual terminology incorporated into national surveys which aim to monitor adolescent health-related risk behaviors, particularly sexual behaviors. As a result, Howard, Sawyer and Brewster-Jordan conducted a comprehensive literature review to examine the specific language used on current national health surveys that assess the sexual behavior of adolescents and young adults. The aim was to evaluate the extent to which that language is vague or limited in scope. The decision to examine national surveys was based on the fact that they are representative in nature, rigorously designed, and often guide policy and funding decisions at various levels within the public health infrastructure.

The literature review was conducted using multiple large research databases containing published journal articles and abstracts. A total of eight national surveys meet eligibility criteria and underwent thorough review. These eight surveys included: American College Health Association - National College Health Assessment (ACHA-NCHA); Kaiser Family Foundation and Seventeen Magazine National Survey of Teens: Safer Sex, Condoms and the Pill; Middle School Youth Risk Behavior Survey (MSYRBS); 1995 National College Health Risk Behavior Survey (NCHRBS); National Longitudinal Study of Adolescent Health (Add Health); National Survey of Family Growth (NSFG); NBC News/People Magazine: National Survey of Young Teens Sexual Attitudes and Behaviors; and Youth Risk Behavior Survey (YRBSS).

A review of the surveys found inconsistent sexual health behavior terminology being used with most lacking descriptions of the behaviors being queried. All the surveys asked about intercourse, either using the specific phrase "vaginal intercourse" or the more general phrase "sexual intercourse". Very few surveys asked about "anal sex" and/or "oral sex", however, one of the surveys asked about both sexual behaviors and included explanations of these acts. Surveys administered through an interview methodology, as opposed to a selfadministered questionnaire, were more apt to delve into detailed sexual behavior questions. The findings from this study further exhibit the need for researchers to provide specific descriptions of adolescent sexual health terms in order to enhance the validity and reliability of measures designed to investigate and understand an adolescent sexual behavior.

## Exploratory Study of the Connotative Meaning of Safer Sex

As previously mentioned, nearly half of all HIV infections and two thirds of all sexually transmitted STIs occur among youth and young adults under the age of 25, yet there is currently no federal mandate to teach sex education in public school (Starkman, and Rajani, 2002). Indeed, federal funding for programs that focus on abstinence-only, which exclude information on sexuality and contraception, has increased. Teachers appear stymied in their ability to adequately educate youth about human sexuality; they report being more and more handicapped in their ability to provide answers to sexuality questions that are not included in the more restrictive abstinence-only educational curricula (Darroch, Lanfdry, and Singh, 2000). This shift comes at a time where a majority of Americans say they favor courses that teach contraception and condom use, in addition to abstinence (Wilson, 2000). Findings from the Annenberg National Health Communication Survey, conducted in 2006 to gauge U.S. adults preference for the teaching of sex education in public schools, found that a plurality of adults (82% of respondents) favor abstinence-plus programs which emphasize not only abstinence but other methods of preventing pregnancy and STIs (Bleakley, Hennessy, and Fishbein, 2006). One of the potential downsides from the aforementioned substantive changes in sexuality education is growing confusion, among young adults, as to whether one can practice safe sex and what constitutes safe sex.

The term, safe sex, whose coinage was associated with the emergence of a global AIDS crisis in the 1980's (Hillier, Harrison, and Warr, 1998), has evolved into the language of safer sex; perhaps in acknowledgement of an ongoing dialectic between risk elimination, a seemingly unattainable goal, and risk *reduction*, a more modest approach but pragmatically more achievable (Moore, 1997). There is an implication that the term safe sex is an oxymoron; in essence, the only way to avoid risks associated with sexual contact is to abstain completely from sexual activity (Wikipedia, 2006). The intent of developing the terms safe and safer sex has been largely an attempt to reduce potentially risky sexual behavior and its negative consequences; practical recommendations for safe(r), versus risky activities have become commonplace in sexuality communications and interventions during the past two decades. Unfortunately, despite the development of various definitions for this term (Medline Plus, 2006), the nomenclature of safe(r) sex appears susceptible to the same ambiguities as other terms that reference sexual behavior. Indeed, as noted by Moore (1997), "within the categories of "possible safe" practices, ideas about what constitutes safer sex are constantly being reformulated" and exist on a continuum rather than seen as a polarity. This confusion clearly impedes the development of interventions that seek to address effective measures to limit sexual risk-taking.

In some contexts, safe sex is seen as synonymous with both safer sex and protected sex and refers to a set of practices that are intended to reduce the risk of acquiring a STI. Yet, at times, the term is associated with a series of seemingly divergent messages that aim to transform sexual practices (Moore, 1997). Hillier et. al., (1998) contend that usage of the term safe sex has been overly simplified and sanitized and has been reduced to a simple equation: safe sex equals condom use. By contrast, unsafe sex refers to unprotected sexual behavior, i.e., sexual intercourse without use of any barrier contraception or prevention against STIs (Wikipedia, 2006). Yet, it also has been argued that sexual behavior must not be placed in the same context as other risk behavior, since human sexuality is both normative and healthy- it is only certain sexual practices that constitute risk (Dailard, 2001). In fact, there is some concern that endorsement and promotion of abstinence only practices such as virginity-until-marriage pledges may have the unintended consequence of leaving youth vulnerable to unplanned pregnancy and STI if and when they do break these pledges, because evidence suggests they are less likely to engage in protected sex (Bearman and Bruckner, 2000). The significance of this finding may lie in the fact that by 1995 more than 2 million adolescents, or 12% of all adolescents, had taken such pledges (Bruckner and Bearman, 2005). Based on data drawn from the National Longitudinal Study of Adolescent Health (Add Health), Bruckner and Bearman (2005) examined the effectiveness of virginity pledges in decreasing STD infection among youth aged 8-24. They found that while those who pledge may initiate sexual intercourse later than youth who do not make such pledges most will eventually engage in premarital sex; more importantly, when they do they are less likely to use a condom at first intercourse. For those youth who do not have premarital sex, they are likely to substitute oral and/or anal sex for vaginal sex.

It has been argued that there is a direct relationship between the high rate of negative sexual health outcomes among youth in the U.S. and negative attitudes toward sexuality, an "erotophobia" of sorts, as promulgated by ideologically conservative and religious fundamentalists (Dodge, Sandford, Yarber, and de Wit, 2005; Weeks, 2002; Weeks, 1989). Furthermore, evidence is accumulating that the message of safe sex is equally as potent as an abstinence message in reducing the amount of unprotected sex, yet the political climate may dictate that abstinence-only programs remain a cornerstone to educational efforts for youth (AIDS Alert,1998). Finally, the best evidence available, i.e., based on rigorously designed and evaluated trials, suggests that abstinence-only curricula have failed to demonstrate efficacy in delaying initiation of sexual intercourse (Santelli, Ott, Lyon, Rogers, and Summers, 2006a; Santelli, Ott, Lyon, Rogers, Summers, and Schleifer, 2006b). In fact, Santelli et al. (2006b), maintain that the problem with abstinence- only programs is not abstinence but *only* abstinence and raise important scientific, human rights and ethical concerns with the federal funding requirements surrounding sexuality education.

Research conducted over the past 30 years suggests that the meaning of sexual terms changes over time, there are often gender differences in the how sexual terms are used and interpreted and what may constitute sexual behavior may also differ by demographics (Bogart, Cecil, Wagstaff, Pinkerton, and Abramson, 2000). Bogart et al., (2000) found that broad cultural definitions influenced subjective judgments regarding sexual activity and virginity status along with contextual considerations such as the nature of the behavior, whether and who experienced orgasm. This has been echoed by Dodge et al., (2005) who argue that human sexual behavior involves a complex socio-cultural set of dynamics. Indeed, when applying terms such as safer sex and risk in the context of sexual health messages and prevention efforts, it is appropriate to consider personal and situational factors if the goal is to facilitate decision-making to reduce risk (Moore, 1997). Perhaps because the concept of safe sex derived from a model designed for gay men, it neglects certain psychosocial implications of sexual activity that specifically apply to women, including a sullied reputation and the

concomitant social sanctions that are associated with such behavior, particularly when the relationship is not sustained (Hillier, Harrison, and Warr, 1998).

In light of these ambiguities and contradictions, the following exploratory study aimed to examine the denotative meaning of the term "*safer sex*".

# **METHODS**

Data used for this study were gathered from administration of an opened-ended survey on safer sex to students enrolled in two undergraduate courses at a large mid-Atlantic university. The two courses where participants were recruitment from were a 100 level Personal and Community Health course and a 300 level Psychology of Women course. Students were provided with a brief explanation of the purpose of the study followed by instructions for survey completion. Study team members emphasized the anonymity of the collected responses and that participation was completely voluntary.

# Survey Administration

A 1-item survey was designed to capture open-ended feedback from students about their interpretation of the concept *safer sex*. The item asked students, "*When someone says he or she is practicing "safer sex*", what do YOU think he or she means? The open-ended responses for this item were systematically coded by the research team using a master coding dictionary developed specifically for that purpose. Responses were reviewed and coded twice by different study team members to establish inter-rater reliability.

The development of the master coding dictionary began with the identification of six coding categories. These categories reflected an initial review of written responses collected from the Personal and Community Health course. The six coding categories were: (1) general description (not specifying a particular form of contraception or behavioral protection), (2) behavioral act of protection, (3) contraception, (4) protection, (5) abstinence, and (6) other. Each coding category was labeled and included examples taken directly from student surveys; collectively they constituted the coding dictionary. Using these six categories, surveys from the Psychology of Women course were reviewed to determine whether additions and/or changes to the coding dictionary were required. While no additional categories were needed, additional examples were included into the master coding dictionary to further illustrate differences across the six categories. Surveys from both courses were independently reviewed twice to confirm the final coding. Collected data from this survey was entered into SPSS 13.0 for analysis (Shad, 1997). Overall frequency distributions were computed along with sub-analyses by gender.

# RESULTS

An overwhelming majority of students listed either contraception or protection as the behavioral referent for the term "practicing safer sex" (see Table 1). Males were more likely

to use the more narrow term contraception, while females were more apt to define the practice of safer sex in a broader sense, i.e., protection. Only 1 student defined safer sex as abstinence.

Sexual Behaviors	% (N)	Females	Males
General description, no behavioral referents <sup>4</sup>	6.5 (8)	0.07 (6)	0.07 (2)
Activation of behavioral protection <sup>5</sup>	1.6 (2)	0.02 (2)	0 (0)
Contraception <sup>6</sup>	39.0 (48)	0.34 (31)	0.50 (14)
Protection <sup>7</sup>	48.8 (60)	0.55 (50)	0.36 (10)
Abstinence <sup>8</sup>	0.8 (1)	0	0
Other	3.3 (4)	0.02 (2)	0.07 (2)

# Table 1. Denotative meaning of the term "practicing safer sex"<sup>1</sup> $(N=123; 91 \text{ females}, 28 \text{ males})^{2,3}$

<sup>1</sup> Actual question asked: "When someone says he or she is practicing "safer sex", what do *YOU* think he or she means?"

<sup>2</sup> Data were gathered from an introductory health class, HLTH 140- Personal and Community Health (n = 53) and a midlevel psychology class, PSYCH 336- Psychology of Women (n = 70)

<sup>3</sup> 4 participants did not indicate gender

<sup>4</sup> Examples included: "Being more moderate and less dangerous", "Means doing something you weren't doing before, but doesn't mean practicing safe sex", "Taking more precautions"

<sup>5</sup> Examples included: "Getting checked out", "Limiting sexual activities to just one person, not engaging in casual sex"

<sup>6</sup> Examples included: "Condoms", "Birth Control", "Spermicides", "More than one form of protection", "Birth Control", "Using condoms more often", "One or more forms of contraception"

<sup>7</sup> Definitions incorporated the term protection; category further divided into 9 sub-domains

<sup>8</sup> Definitions incorporated the term abstinence.

# Table 2. Meaning of "Protection" as regards the practice of safer sex1(N= 60; 50 females, 10 males)

Sexual Behaviors	% (N)	Females	Males
Using precautionary measurements/using > 1 method	30 (18)	0.30 (15)	0.30 (3)
Having control over actions + contraceptives	8.3 (5)	0.10 (5)	0
Contraceptives + other means of having "safe sex"	3.3 (2)	0.02(1)	0.10(1)
Using protection + getting tested	5.0 (3)	0.06 (3)	0
Improving own personal methods of protection	20.0 (12)	0.18 (9)	0.30(3)
Birth control + condoms + pulling out	1.7(1)	0.02(1)	0
Working towards being less sexually active +	23.3 (14)	0.24 (12)	0.20 (2)
contraception			
Checking up on sexual health of partners +	5.0 (3)	0.04 (2)	0.10(1)
contraceptive			
Taking precautions (i.e., using condoms) some of the	3.3 (2)	0.04 (2)	0
time			

<sup>1</sup> Pertains to the subset of students who answered question in Table 1 and provided specification on the meaning of protection as a strategy for practicing "safer sex

<sup>2</sup> Table 1 data were gathered from an introductory health class, HLTH 140- Personal and Community Health (n = 53) and a midlevel psychology class, PSYCH 336- Psychology of Women (n = 70).

For those students who used the term protection to describe the practice of safer sex (50.4%) many provided additional detail as to the meaning of this term (see Table 2). The most common definition of protection included using precautionary measures, that is, more than 1 method of protection. Next frequent were statements indicating the use of contraception, but also sentiments related to "working towards being less sexually active". Frequent reference was also made to improving one's personal methods of protection. Some interesting gender differences emerged. Only females mentioned "having control over one's actions" or "getting tested" as a measure of safer sex. Additionally, females were much more likely to indicate that "working toward being less sexually active" was a means of being protected.

# DISCUSSION

The findings from this exploratory study suggest that being protected and/or using contraception is seen as synonymous with the practice of safer sex. Only 1 student equated safer sex with abstinence. Examples of contraception provided by students included using condoms, spermicides or birth control. Clearly, these methods are not equivalent in terms of preventing both pregnancy and STIs. Furthermore, in elaborating on the use of contraception, some participants interpreted the practice of safer sex to mean more frequent use of contraception or use of more than one form of contraception. This raises questions regarding the regularity of contraception use and contextual factors that might determine whether and which additional forms of contraception are used. Clearly, more needs to be learned about the decision-making process and factors that influence the nature, extent and choice of contraception use among college students.

The meaning of protection, in the context of practicing safer sex, was wide ranging. Some students discussed protection in terms of behavioral referents, i.e., "using more than one method" and "getting tested". Some females referred to cognitive factors such as "having more control over one's actions". For some males and females, being protected included "working toward being less sexually active", conceivably endorsing implicitly the notion that safe sex equates to abstinence.

# CONCLUSION

As noted by Abraham and Sheeran more than a decade ago (1993), perhaps the thrust, so to speak, of future adolescent sexuality research should be away from biology to social psychology and a search for a psychology of safer sex promotion. While emphasizing the complex social skills that are embodied in the practice of safe sex attention must also be placed on how environmental and cultural forces shape and reinforce sexual attitudes and behavior. As part of an advocacy agenda for 2010, Wilson (2000) argues that proponents of responsible sex education must continue to remind the public that an overwhelming majority support the teaching of contraception and disease prevention in addition to abstinence and point out to politicians that there does not exist scientifically credible evidence that abstinence-only education has delayed the onset of sexual intercourse or reduced sexual

activity. Researchers, as well, must be pressed to be more explicit and exacting when using sexual terminology, assessing sexual behavior and interpreting sexuality-related survey data.

Surely, we have our work cut out for ourselves.

# REFERENCES

- Abraham, C., Sheeran, P. (1993). In search of a psychology of safer-sex promotion; beyond beliefs and texts. (1993). *Health Education Quarterly*, 8, 245-254.
- Administration for Children and Families. (2006). Request for proposals for communitybased abstinence education programs, Family and Youth Services Bureau. Accessed from: http://www.acf.hhs.gov/grants/open/HHS-2006-ACF-ACYF-AE-0099.html
- American College Health Association. (2003). National College Health Assessment. Accessed from: http://www.acha.org/projects programs/sample ncha.pdf.
- Bartels C, Wilcox B, Limber S, O'Bierne H. (1996). Adolescent Abstinence Promotion Programs: An Evaluation of Evaluations. Paper presented at the Annual Meeting of the American Public Health Association, No. 18, New York, NY.
- Bell, H.A. (2000). Just because you see their privates doesn't mean you're not a virgin: adolescents' understanding of sexual terminology. Unpublished thesis, Iowa State University, Ames, IA.
- Bearman, P.S., and Bruckner, H. (2000). Promising the future: Virginity pledges as they affect transition to first intercourse. *Institute for Social and Economic Theory and Research*, 1-63.
- Birnbaum, C. (2000). The love and sex survey, Twist, Oct/Nov. 2000, 54-56.
- Bleakley, A., Hennessy, M., and Fishbein, M. (2006). Public Opinion on Sex Education in US Schools. Archives of Pediatric and Adolescent Medicine, 160, 1151-1156.
- Bogart, L.M., Cecil, H., Wagstaff, D.A., Pinkerton, S.D., and Abramson, P.R. (2000). Is it sex? College students' interpretations of sexual behavior terminology. *Journal of Sex Research*, 37, 108-116.
- Bruckner, H., and Bearman, P. (2005). After the promise: the STD consequences of adolescent virginity pledges. *Journal of Adolescent Health*, 36, 271-8.
- Carpenter, L. (2001). The ambiguity of "having sex": The subjective experience of virginity loss in the United States. *The Journal of Sex Research*, 38, 127-139.
- Cecil, H., Bogart, L.M., Wagstaff, D.A., Pinkerton, S.D., and Abramson, P.R. (2002). Classifying a person as a sex partner: The impact of contextual factors. *Psychology and Health: An International Journal*, 17, 221-234.
- Centers for Disease Control and Prevention. (2003). Youth Risk Behavior Surveillance MMWR, (53) SS-2,: United States, 2003, May 21, 2004. Accessed from: http://www.cdc.gov/mmwr/PDF/SS/SS5302.pdf.
- Centers for Disease Control and Prevention. (2004). STD Surveillance. Trends in Reportable Sexually Transmitted Diseases in the United States. Accessed from: http://www.cdc.gov/std/stats/trends 2004.htm.
- Chase-Lansdale, P.L. (1995) Who's Responsible? An historical analysis of the changing roles of mothers, fathers, and society. In: PL Chase-Lansdale and J Brooks-Gunn (Eds.), *Escape from Poverty: What makes a difference for children?* (pp11-37) New York: Cambridge University Press.
- Chesson, H.W., Blandford, J.M., Gift, T.L., Tao, G., and Irwin, K.L. (2004). The estimated cost of STD in American youth, Abstract P075. 2004 National STD Prevention Conference, Philadelphia, PA.

- Dailard, C. (2001). Recent Findings from The 'Add Health' Survey: Teens and Sexual Activity. *The Guttmacher Report on Public Policy*, 4, 1-3.
- Darroch, J.E., Singh, L., and Singh, S. (2000). Changing emphases in sexuality education in U.S. public schools, 1988-1999. Family Planning Perspectives, 32, 204-211 and 265.
- Dodge, B., Sandfort, T.G.M., Yarber, W.L., and de Wit, J. (2005). Sexual health among male college students in the United States and the Netherlands. *American Journal of Health Behavior*, 29, 172-182.
- Editorial staff. (1998). Safe sex vs abstinence: which is more effective. AIDS Alert, 78-79.
- Guttmacher Institute. (2004) U.S. Teenage Pregnancy Statistics: Overall trends, trends by race and ethnicity and state-by-state information, updated February 19, 2004. Accessed from: http://www.guttmacher.org/pubs/state\_pregnancy\_trends.pdf
- Halpern-Felsher, B.L., Cornell, J.L., Kropp, R.Y., and Tschann, J. (2005). Oral versus vaginal sex among adolescents: perceptions, attitudes and behavior. *Pediatrics*, 115, 845-851.
- Health and Human Services. (2007). FY 2007 Program announcement, section 510, Abstinence Education Program. Accessed from: www.acf.hhs.gov/grants/open/HHS-2007-ACF-ACYF-AEGP-0143.html., November 3, 2006.
- Hickman-Brown Research Inc. (1999). Public Support for Sexuality Education Reaches Highest Level. SIECUS and Advocates for Youth. Accessed from: http://www.siecus.org/parent/pare 0003.html
- Hillier, L., Harrison, L., and Warr, D. (1998). "When you carry condoms all the boys think you want it": negotiating competing discourses about safe sex". *Journal of Adolescence*, 21, 15-29.
- Hook, E.W. and Handsfield, H.H. (1999). Gonococcal infections in the adult. In: Holmes K.K., Mardh, P.A., Sparling, P.F., et al., eds: *Sexually Transmitted Diseases*. 3rd ed., New York,NY:McGraw-Hill:p.456.
- Horan, P.F., Phillips, J., and Hagan, N.E. (1998). The meaning of abstinence for college students. *Journal of HIV/AIDS Prevention and Education for Adolescents and Children*, 2, 51-66.
- Ito, K.E., Gizlice. Z., Owen-O'Dowd, J., Foust, E., Leone, P.A., and Miller, W.C. (2006). Parent opinion of sexuality education in a state with mandated abstinence education: does policy match parental preference? *Journal of Adolescent Health*, 39, 634-641.
- Kaiser. (2007). Daily women's health policy: Federal guidelines expand the scope of abstinence education funds to include people up to age 29. Accessed from: http://www.Kaisernetwork.org. October 31, 2006.
- Kaiser Family Foundation. (2000). Sex Education in America: A View from Inside the Nation's Classrooms. Accessed from: http://www.kff.org/youthhivstds/3048-index.cfm
- Kaiser Family Foundation. (2004). Sex Education in America: General Public/Parents Survey. Accessed from, http://www.kff.org/newsmedia/7017.cfm
- Kaiser Family Foundation and Seventeen Magazine National Survey of Teens: Safer Sex, Condoms and the Pill (2000). Accessed from: http://www.kff.org/entpartnerships/ upload/SexSmarts-Survey-Safer-Sex-Condoms-and-the-Pill-Toplines.pdf
- Kashima, H.K., Shah, F., and Lyles, A.A. (1992). A comparison of risk factors in juvenileonset and adult-onset recurrent respiratory papillomatosis. *Laryngoscope*, 102, 9-13.
- Kirby, D. (2002). Do Abstinence-Only Programs Delay the Initiation of Sex Among Young People and Reduce Teen Pregnancy? Washington DC: Campaign to Prevent Teen Pregnancy, October 2002.
- Lafferty, W.E., Downey, L., Celum, C., and Wald, A. (2000). Herpes simplex virus type 1 as a cause of genital herpes: impact on surveillance and prevention. *Journal of Infectious Disease*, 81, 1454-1457.

- Lewin, T. (1997). Teenagers alter sexual practices, thinking risks will be avoided, *New York Times*, April 5, 1997, A1.
- Mathematica Policy Research, Inc. (2005, June). First Year Impacts of Four Title V, Section 510 Abstinence Education Programs. Princeton, NJ.
- Medline Plus. "Safe" sex. updated June 6, 2006. Accessed from: http://www.nlm.nih.gov/medlineplus/ ency/article/001949.htm.
- Mercer, J.G. (1999). Defining and teaching abstinence: an e-mail survey of health educators. Unpublished thesis, North Carolina State University, Raleigh, NC.
- Middle School Youth Risk Behavior Survey. (2005). Accessed from: http://www.cdc.gov/ HealthyYouth/yrbs/pdf/questionnaire/2005middleschoolquestionnaire.pdf
- Moore, L. (1997). "It's like you use pots and pans to cook. It's the tool": The technologies of safer sex. *Science, Technology, and Human Values*, 22, 434-471.
- Mosher, W., Chandra, A., Jones, J. (2005). Sexual behavior and selected health measures: men and women 15-44 years of age, United States, 2002. *Advance Data from Vital and Health Statistics* September 15, 2005; No. 362:21-26.
- Nanchahal, K., Wellings, K., Barrett, G. Copas, A., Mercer, C., Macmanus, S., Macdowall, W., Fenton, K., Erens, B., and Johnson, A. (2005). Changes in the circumstances of young mothers in Britain: 1990 to 2000. *Journal of Epidemiology Community Health*, 59, 828-833.
- National College Health Risk Behavior Survey (1995) Updated November 14, 1997. Accessed from: http://www.cdc.gov/mmwr/preview/mmwrhtml/00049859.htm.
- The National Longitudinal Study of Adolescent Health (Add Health). Accessed from: http://www.cpc.unc.edu/addhealth
- National Survey of Family Growth (NSFG). (2002). Accessed from: http://www.cdc .gov/nchs/nsfg.htm
- NBC/People (2005) National survey of young teens' sexual attitudes and behaviors, 2005, updated January 31, 2005. http://www.msnbc.msn.com/id/6839072.
- NBC /People: National Survey of Young Teens Sexual Attitudes and Behaviors. (2004). Accessed from: http://msnbcmedia.msn.com/i/msnbc/Sections/TVNews/Dateline %20NBC/ NBCTeenTop line.pdf
- Norris, A. E., Clark, L. F., and Magnus, S. (2003). Sexual abstinence and the Sexual Abstinence Behavior Scale. *Journal of Pediatric Health Care*, 17, 140-144.
- Ott, M.A., Pfeiffer, E. J., and Fortenberry, J.D. (2006). Perceptions of sexual abstinence among high- risk early and middle adolescents. *Journal of Adolescent Health*, 39, 192-198.
- Pitts, M., and Rahman, Q. (2001). Which behaviors constitute "having sex" among university students in the UK? *Archives of Sexual Behavior*, 30, 169-176.
- Sanders, S.A. and Reinisch, J.M. (1999). Would you say you "had sex" if ...? Journal of the American Medical Association, 281, 275-277.
- Santelli J., Ott, M.A., Lyon, M., Rogers, J., Summers D.. (2006a). Abstinence-only education policies and programs: a position paper of the Society for Adolescent Medicine, *Journal* of Adolescent Health, 38, 83-87.
- Santelli J., Ott, M.A., Lyon, M., Rogers, J., Summers D., and Schleifer, R. (2006b). Letters to the editor: Abstinence and Abstinence-Only Education, *Journal of Adolescent Health*, 39, 152-154.
- Saul, R. (1998). "Whatever Happened to the Adolescent Family Life Act?" *Guttmacher Report on Public Policy*, 1, 5-11.
- Sawyer, R.G., Howard, D.E., Brewster-Jordan, J., Gavin, M., and Sherman, M. (In press, 2007). "We didn't have sex ... did we?" College students' perceptions of abstinence. *Journal of Health Studies*.

- Schuster M.A., Bell R.M., and Kanouse D.E. (1996). The sexual practices of adolescent virgins: genital sexual activities of high school students who have never had vaginal intercourse. *American Journal of Public Health*, 86, 1570-1576.
- Seventeen. (2000, February 28). News Release: National survey conducted by *Seventeen* finds that more than half of teens ages 15-19 have engaged in oral sex.
- Shad, B.W., Barnwell P.S., and Bieler G.S. (1997). Software for the statistical analysis of correlated data: User's manual. Research Triangle Park, NC: Research Triangle Institute.

Social Security Act, P.L. (1996), 104-193, Section 510(b) of Title V.

- Starkman, N. and Rajani, N. (2002). The case for comprehensive sex education. AIDS Patient Care STDS, 16, 313-318.
- Stepp, L.S. (1999). Parents are alarmed by an unsettling new fad in middle schools: oral sex. Washington Post, July 8, 1999, p.A1; and Stepp, LS, Talking to kids about sexual limits, Washington Post, July 8, 1999, p. C4.
- Stevens-Simon, C. (2001). Virginity: a state of mind...but not necessarily of body. Journal of School Health, 71, 87-88.
- Stone, N., Hatherall, B., Ingham, R., McEachran, J. (2006). Oral sex and condom use among young people in the United Kingdom. *Perspectives on Sexual and Reproductive Health*, 38, 6-12.
- Weeks, J. (2002). The social construction of sexuality. In Peiss K, (ED). Major Problems in the History of American Sexuality. Boston: Houghton Mifflin Company, 2-9.
- Weeks, J. (1989). Sex, politics and society: the regulation of sexuality since 1800 (2nd edition). London: Longman, 1-326.
- Wikipedia. (2006). Safe sex. Accessed from: http://en.wikipedia.org/wiki/Safe\_sex
- Wilson, S. (2000). Sexuality education: Our current status, and an agenda for 2010. Family Planning Perspectives, 32:252-254.
- Wilson, S.N. (2000). Sexuality education: Our current status, and an agenda for 2010. *Family Planning Perspectives*, 32, 252-254.
- Youth Risk Behavior Surveillance United States, 2005. Morbidity and Mortality Weekly Report 2006; 55(SS-5):1-108.
- Youth Risk Behavior Surveillance Survey. (2005). Accessed from: http://www.cdc.gov/ HealthyYouth/yrbs/pdf/questionnaire/2005/

Chapter 27

# COMPARING RISKY AND INTER-TEMPORAL DECISIONS: VIEWS FROM PSYCHOLOGY, ECOLOGY AND MICROECONOMICS\*

Tobias Kalenscher<sup>1,†</sup> and Philippe N. Tobler<sup>2,‡</sup>

<sup>1</sup>Animal Physiology & Cognitive Neuroscience Neurobiology Section of Swammerdam Institute for Life Sciences (Sils) University of Amsterdam, Faculty of Science Kruislaan 320 1098 sm, Amsterdam, The Netherlands

<sup>2</sup>Department of Physiology, Development and Neuroscience University of Cambridge Downing street Cambridge cb2 3dy United Kingdom

# ABSTRACT

When making decisions between different options, we often consider two basic properties of these options, how risky they are and when they will occur. For example, we may choose to gamble or to wait for a larger reward. Decisions under risk refer to decisions among known probabilistic options, inter-temporal decisions refer to choices between options that will be realized at known future timepoints. Risky and inter-temporal decisions have been captured theoretically primarily by ecology and microeconomics but findings from behavioral economics, psychology and neuroscience often contradicted theoretical predictions. As a consequence, a wealth of more descriptive Models has emerged to explain the findings. A subset of these models has stressed the Similarities between risky and inter-temporal decisions. In this chapter we review both

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Corrrespondence to: University of Amsterdam, Faculty of Science, Kruislaan 320 1098 sm, Amsterdam The Netherlands, Ph +31 (0)20 525 7658 Fax +31 (0)20 525 7709 Email t.kalenscher@uva.nl, Web http://home.medewerker.uva.nl/t.kalenscher/

<sup>&</sup>lt;sup>‡</sup> Correspondence to: <sup>2</sup>Department of Physiology, Development and Neuroscience, University of Cambridge Downing street Cambridge cb2 3dy United Kingdom, Ph +44 1223 339 544 Fax +44 1223 333 786 Email pnt21@cam.ac.uk

core theoretical approaches and empirical findings. We discuss possible explanations for discrepancies and identify key behavioral experiments.

# MAIN TEXT

# 1. Introduction

When we make decisions, the outcomes of our choices rarely occur with certainty, and often we have to wait some time for the consequences to happen. For example, investing time and money into a good education makes more likely, but doesn't guarantee, a successful professional career and a high income; in other words, the outcome of your investment decision is probabilistic. Likewise, when paying the high tuition fees for your education, you invest resources now for benefits that are yet to come, because you will only be able to harvest the fruits of your labor once you finish your education in a couple of years. Choices between probabilistic outcomes are called 'risky decisions' and choices between outcomes That will be realized at different instants in the future are called 'inter-temporal decisions'. Both types of decisions have been extensively discussed in several scientific disciplines, Including biology, ecology, micro- and macroeconomics, psychology and cognitive neuroscience. In this chapter, we review some of the most influential theories on risky and inter-temporal decision making, and will outline the theoretical and empirical differences in the different approaches. We will then discuss to what degree attempts to unify the two research fields are and can be successful and identify key behavioral and neuroscientific experiments. We conclude with highlighting the importance of cooperation between the various disciplines in elucidating the effects of risk and time on choice.

# 2. Decisions Without Risk

Decisions between certain, immediate, but quantitatively different choice outcomes appear easy: you just compare which of the two outcomes results in the higher gain or the smaller loss, and choose accordingly. However, how do you compare two qualitatively different commodities, for example, apples and pears? In economics, this problem is solved by assuming that different commodities are translated into a common currency, the subjective value, or the utility of a prospect. Utility is a measure of relative satisfaction or gratification which allows to rank-order and therefore compare the different possible outcomes (montague

And berns, 2002). Although frequently used in financial contexts (also in this chapter), utility does not exclusively refer to monetary gains (and losses), but also to more abstract benefits, Such as obtaining pleasure from engaging in a favorite recreational activity, or enjoying one's favorite food, or the like. Although embracing essentially the same solution, behavioral ecology has given a biological twist to the common currency problem. It replaces utility with fitness which, depending on the model, may correspond to e.g. rate of energy gained per unit of time spent foraging (charnov, 1976) or to reproductive success (hamilton, 1964). The utility of an outcome is not a linear function of its objective (e.g., monetary) Value, but a function of the current level of wealth (friedman and savage, 1948; bernoulli, 1954; kahneman and tversky, 1979; tobler et al., 2007a). More precisely, it has been argued

That each additional unit in the utility function, the so-called marginal utility (mankiw, 2004), Is smaller than the previous unit, resulting in a progressive decrease in marginal utility with Increasing assets or energy reserves (friedman and savage, 1948; bernoulli, 1954; sibly and Mcfarland, 1976; kahneman and tversky, 1979; kacelnik, 1997; kacelnik and bateson, 1997; tobler et al., 2007a). As a consequence, the utility of a commodity is presumed to be a Decelerating concave function of this commodity (figure 1a). For example, winning \$100 would be more valuable to you when you are poor than when you are a millionaire.

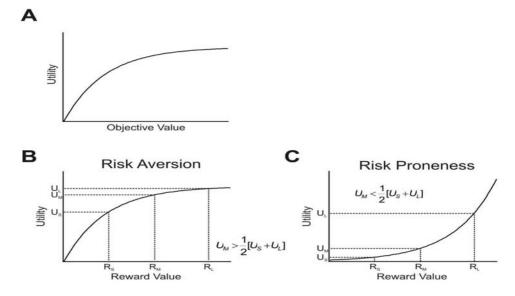


Figure 5. Utility functions. (a) utility as a function of the objective value of a commodity. The utility curve is a concave function of the current level of wealth because the Marginal utility, i.e., the utility increment with each additional unit, decreases with increasing level of wealth. (b) a concave utility function predicts risk aversion when choosing between a medium-sized, certain and a large and small risky reward. In a multi-choice situation, the average utility of the certain rewards exceeds the average utility of the risky rewards because the utility of the large reward is sublinearly larger compared to the utilities of the other rewards. (c) a convex function predicts risk proneness.

# 3. Decisions Under Risk

Risky decisions are decisions between probabilistic outcomes. The level of riskiness is equivalent to the spread from the mean (variance) of the risky outcomes. For example, a gamble that pays either \$ 9 or 11 with 50% probability is less risky than a gamble that pays either \$ 2 or 18 with 50 % probability (although both gambles have the same mean payoff of \$ 10). This notion of risky decision making is important, as it differs from the common, folk-psychological conception of risky behavior, which frequently implies that a person engaging in risky choices is consciously willing to accept high losses ("the gambler who bets his house and family"). However, although the readiness to accept large losses may certainly play a role in biasing an individual's risk attitude, academic research of risky decision making focuses on the formal impact of outcome variance on choice.

Neither humans nor animals are risk-neutral (friedman and savage, 1948; kacelnik and bateson, 1996, 1997; bateson and kacelnik, 1998): when deciding between a certain and a risky option, one is often chosen more frequently than the other, even when the expected values of the respective options (their probabilities multiplied with their objective values) are identical. Most animals, including humans, are risk-averse, but show occasional risk-Proneness (friedman and savage, 1948; kacelnik and bateson, 1996, 1997; bateson and Kacelnik, 1998; glimcher, 2002; mccoy and platt, 2005a, 2005b; hayden and platt, 2007; Tobler et al., 2007b). Importantly, this suggests subjective differences in the valuation of the aame objective options. Utility theories have been put forward in order to capture subjective valuation with a particular focus on risk.

# Economic Utility Theories

Utility theories assign numbers to preferences with both a descriptive and a normative purpose. Specially psychology is interested in explaining choice behavior and has thus focused on the descriptive aspect whereas economy and statistics, more so in the past, have put stronger weight on the normative aspect of characterizing consistent, coherent and optimal choice and ecology is interested in both descriptive and normative aspects. Utility theories all make basic assumptions (axioms) about the elements of the decision space and the preference relations of the decision maker with respect to these elements. From these axioms they deduce statements (theorems) for example about how the preference relations as observed from choices can be transformed into utility relations (numbers).

Von neumann and morgenstern (1944) axiomatized utility theory by requiring completeness, transitivity, continuity and (not explicitly stated but necessary) independence of preferences. Thereby they founded expected utility theory (eut), which provides the most prominent normative framework for the analysis of decisions under risk. The completeness axiom requires that the decision maker has preferences across options, transitivity that preferences are in a basic hierarchical order, continuity that for each option there is a better and worse one and independence that preferences do not change by adding common outcomes to all options. If these axioms are fulfilled then a number u (x) can be assigned to each x so that:

$$X = y \text{ if and only if } u(x) \le u(y).$$
(1)

The preference relation x = Y may be read as "alternative x is not preferred to alternative y", <= corresponds to the standard "smaller than or equal to" and u(x) and u(y) refer to the utilities of x and y. If this is true then utility function u preserves the ordering of = and allows translating utilities to preferences and vice versa.

As long as the axioms described above hold eut proposes that a decision whether to accept or reject a choice option should be made by multiplying the utility of all possible outcomes of the option with their probabilities, integrating across products, and choosing the option with the larger sum. Thus, the utility of a choice option with risky outcomes Corresponds to its expected utility. For example, if you consider that it is relatively likely to receive a high income following a good education, and that with the expected income you will be better than with your current situation, then you should decide to invest into a good education. In general, benefits represent positive utility values, costs negative values.

The shape of the utility curve can be related to people's risk attitudes. A risk neutral person has a linear utility curve, and for such a person the expected utility of a gamble is equivalent to the utility of the mean of the gamble. Convex utility curves correspond to risk-proneness, concave curves to risk aversion (figure 1b and 1c). To use a simple example, imagine a situation in which an agent chooses between a certain option, offering a medium-sized reward, and a risky option, offering a large and a small reward with a 50% chance each. The expected value of both options is identical. According to eut, the expected utility of a given choice option is the sum of the utilities of each possible outcome multiplied with their probability

$$E[uoption] = spoutcome \bullet Uoutcome$$
(2)

The expected utility of the certain option would accordingly be computed as:

 $E[ucertain] = 1 \cdot Umedium \cdot reward$ (3)

and the expected utility of the risky option would be computed as:

$$E[urisky] = (0.5 \cdot Usmall - reward) + (0.5 \cdot Ul \operatorname{Arg} E - reward)$$
(4)

Due to the concavity of the utility function of risk averse agents, the utility of the large reward, ularge-reward, is sublinearly larger than the utilities of the medium or small rewards, umedium-reward and usmaller-reward. Thus, eurisky will be smaller than eucertain and agents will avoid the risky option (see figure 1b). Put in simpler words, an agent receives the large and small rewards with equal probability when choosing the risky option. Because the utility of the large reward is sublinearly smaller than the utilities of the other rewards, the mean utility of the large and small rewards (risky option) would be smaller than the utility of the gamble's expected value and thus the utility of the medium-sized certain option. As a consequence, the agent will avoid the risky option. Risk-proneness can be explained by assuming a convex, accelerating utility curve, in which the utility of the large reward is supralinearly larger than the utility of a medium or a small reward (figure 1c).

#### Violations of Preference Axioms

Empirical research showed violations of most of the normative axioms of eut. Reports of violations of the independence axiom appeared relatively soon after von neumann and morgenstern's seminal work (allais, 1953; ellsberg, 1961). As an example, consider the following two decisions (kahneman and tversky, 1979):

Decision 1) A) \$ 2500, p = 0.33; \$ 0, p = 0.67 (read as: a 33% chance of winning \$ 2500 and a 67% chance of winning nothing) B) \$ 2400, p = 0.34; \$0, p = 0.66

Decision 2) A) \$ 2500, p = 0.33; \$ 2400, p = 0.66; \$ 0, p = 0.01 B) \$ 2400, p = 1.0; \$ 0, p = 0.0 Most people choose a) in decision 1) and b) in decision 2). However this pattern of preference reversals violates the independence axiom because 2a and 2b result from adding (\$ 2400, p = 0.66) to 1a and 1b and therefore either a) or b) should be chosen in both cases.

Also the transitivity axiom can be systematically violated such that decision makers show cyclic preferences (a=b, b=c, c=a; e.g. loomes et al., 1991; shafir, 1994; waite, 2001). For example (waite, 2001), blue jays prefer one raisin, 28 cm into a tube (option a) over two raisins, 42 cm into a tube (option b). They also prefer option b over three raisins 56 cm into a tube (option c) but when given the choice between options a and c, they do not prefer a. Humans also show systematic violations of transitivity in certain choice situations (tversky, 1969), for example, when the choice options are composed of several features that vary along different dimensions. Models of context-dependent choice such as regret theory (loomes and sugden, 1982) suggest that violations of the transitivity axiom arise from changes in utility because decision makers evaluate options not in isolation but consider also the outcomes of unchosen alternatives.

Kahneman and tversky (1979) have pointed out another problem of eut in that it does not account for differences in how decision problems are described (framed). As an example, consider the following two decisions (tversky and kahneman, 1986):

Decision 3) A) \$ 240, p = 1.0 B) \$ 1000, p = 0.25; \$0, p = 0.75 Decision 4) A) \$ -750, p = 1.0 B) \$ -1000, p = 0.75; \$ 0, p = 0.25

Most people choose 3a and 4b. However, the combination of 3a and 4b is dominated by (has a lower expected value than) the combination of 3b and 4a.

Findings of axiom violations have provoked different reactions. Some theorists have relaxed one or more of the axioms (e.g. machina, 1982; fishburn, 1982) others have given up on the project of axiomatising utility theory and proposed purely descriptive models. The most famous of these latter approaches is prospect theory (kahneman and tversky, 1979; 1992). Prospect theory suggests that the subjective value function is concave for gains and convex for losses and steeper for losses than gains (figure 2). This reflects the finding that decision makers are usually risk seeking for losses, risk averse for gains and reluctant to accept a fair bet on the toss of a coin (in fact, potential gains have to exceed potential losses by a factor of about 2 in order to achieve indifference). The different steepness for gains and losses introduces a "kink" in the value function which makes it difficult to treat mathematically and which has been termed "reference point". Although not formally defined, The reference point often corresponds to the status quo or the current wealth level. Moreover,

Decision weights modulate probabilities according to an inverted-s-shaped probability Weighting function. This reflects the finding that many decision makers overweigh small and underweigh large probabilities, at least when making hypothetical decisions. Insert figure 2 around here

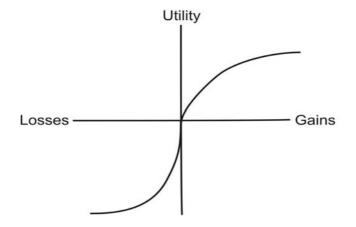


Figure 6. Utility function as proposed in kahneman and tversky's (1979) prospect Theory. The utility function is concave in the domain of gains, and convex in the domain of Losses, and steeper for losses than for gains. The crossing of the axes corresponds to the Reference point against which the prospects are contrasted, for example, the current level of Wealth.

#### Contribution from Behavioral Ecology Risk-sensitive for Aging Theory

The importance of the utility function is to relate subjective to objective value. This is useful because subjective value is often not a linear function of objective value. However, what causes this non-linearity? Bernoulli (1954) suggested that wealth renders the utility function concave. Behavioral ecology suggests additional factors such as upper boundaries on how much energy can be stored and lower boundaries on how much energy is needed for survival can also introduce curvature or even "kinks" into the utility function. Consider for example birds foraging for the night. Because of their small body size and high metabolic rate they face the possibility of starving over night if they fail to accumulate enough resources during the day. As a consequence, a normally risk averse bird might become risk-prone towards the end of the day or at low temperatures, if its energy requirements for the night are not yet met (caraco et al., 1990). Thus, in addition to varying over individuals, risk attitudes vary also over situations and time.

Energy requirements change not only over the course of a single day but also over the year, for example in the pre-migratory period of migratory birds (Moore and Simm, 1986). In that period birds must acquire sufficient reserves for the migration and they behave in a risk-prone manner until they reach maximal body size. Conversely, birds that are not in the pre-migratory period or pre-migratory birds that have reached maximal body size avoid risky foraging options.

Risk-sensitive foraging theory considers decision optimality given reserve constraints. Stephens (1981) showed that it is optimal for birds on a negative energy budget to choose options with higher variance. Conversely, they should avoid risky options when the less risky options provide the birds with a mean rate of intake that exceeds the starvation threshold, i.e. they are on a positive energy budget. For example, a bird almost starved to death (negative energy budget) should avoid the certain option if the certain food amount is not sufficient to guarantee its survival (certain death), and the only chance to survive would be to obtain the large, risky reward (possible survival). On the other hand, a bird on a higher energy budget should be risk-averse if the certain option is sufficient to guarantee survival (certain survival), and the small food amount of the risky option would be insufficient for survival (possible

death). Current models of risk-sensitive foraging incorporate the possibility of sequential choices and the costs of foraging (e.g. McNamara and Houston, 1986).

In addition, life-history trade-offs may have favoured the evolution of risk-attitude in a similar fashion (Wolf et al., 2007). Animals differ in behaviours that affect their future fitness, for example, in their exploration effort during foraging. Animals that put more emphasis on future than immediate fitness returns have higher expectations regarding their future reproductive success than others. Investment into the future only pays out if the animal survives until it is able to realize the upcoming opportunities. Because survival is thus important for the strategy to work out, evolution may have favoured the development of risk-aversion towards predators and aggressive conspecifics in individuals with high fitness/reproductive expectations who have more to lose, whereas individuals with low expectations who have little to lose should be less risk-averse.

In summary, the minimal energy requirement (and by extension the maximal reserves) and/or the life-history of an animal may provide an inflection point ("kink"), where the curvature of its utility function changes.

#### A Normative Framework for Inter-Temporal Decisions

The second type of decisions making discussed in this chapter concerns decisions between outcomes that can only be realized at different instants in the future: Inter-temporal decisions. Humans and non-human animals prefer immediate over delayed rewards (so-called temporal discounting): Provided that the costs for both options are identical, the preference for an immediate or a temporally remote expected outcome is a function of the value of the respective outcomes and their delays, i.e., the time until the outcomes can be realized (McDiarmid and Rilling, 1965; Rachlin and Green, 1972; Ainslie, 1975; Mazur, 1984, 1987, 1988; Grossbard and Mazur, 1986; Logue, 1988; Benzion et al., 1989; Green et al., 1994, 1997; Evenden and Ryan, 1996; Evenden, 1999; Frederick et al., 2002; Reynolds et al., 2002; Kalenscher et al., 2005, 2006a).

Inter-temporal decisions have been extensively studied in psychology and ecology, but were and still are also of great interest in economic models of choice. As with decisions under risk, many of the normative models in economics are based on several theoretical assumptions and theorems, including preference monotonicity, stationarity, and maximization of utility rate.

A fundamental assumption in rational choice theories is that preference orders should be consistent across time. Preference monotonicity and stationarity are directly related to this assumption. Monotonicity means that a prospect X1 that is preferred over another prospect X2 will also be assigned a higher utility than X2 as long as the utility function is monotonic. Monotonicity of time preference (Lancaster, 1963) holds that

$$X(t_1) \ge X(t_2), \text{ if, and only if, } t_2 \ge t_1 \tag{1}$$

This means that commodity X, available at timepoint  $t_1$ , will be preferred over X, available at timepoint  $t_2$ , only when  $t_2$  occurs later than  $t_1$ .

Stationarity is related to the axiom of monotonicity of time preference and posits that

If 
$$X(t) \sim Y(t+\tau)$$
 then  $X(s) \sim X(s+\tau)$  (2)

This means that if an agent is indifferent (~) between commodity X, delivered at timepoint *t*, and commodity Y, delivered at timepoint  $t+\tau$ , he would still be indifferent when X was delivered at timepoint *s* and Y at timepoint  $s+\tau$  (Strotz, 1955; Koopmans, 1960; Fishburn and Rubinstein, 1982). Indifference means that the utility of both options is identical, and thus the frequency of choosing A or B is about 50%. Thus, if both options were deferred by the same time interval, preference orders should be preserved. In other words, if you desire to receive \$10 in 5 days as much as receiving \$50 in 20 days, then you will still desire to receive \$10 in 15 days as much as receiving \$50 in 30 days, i.e., when both delays are prolonged by 10 days.

It was proposed that the discounting rate by which future commodities are delivered should be constant (Samuelson, 1937), for instance resulting in a linear or exponential discount function. Many theories, therefore, assumed exponential discounting (Lancaster, 1963; Fishburn and Rubinstein, 1982; Benzion et al., 1989; cf., Ainslie, 1975; Loewenstein, 1992; Fehr, 2002). Combining exponential discounting with stationarity yields (Lancaster, 1963):

$$(A,t) \sim A e^{-k(t-t_0)} \tag{3}$$

This expression states that a reward with the amount A, delivered at timepoint t, is equally valuable (~) as a reward amount A at  $t_0$  (i.e., now), exponentially discounted for the interval t- $t_0$ , with  $t_0$  referring to the present timepoint, and k being an individually different discount value. In other words, the utility of a future outcome can be expressed as an exponential function of the same outcome realized today.

## Violation of Stationarity

As outlined above, stationarity predicts that the ranking of preferences between several future outcomes should be preserved when the choice outcomes are deferred by the same time interval. This has been investigated in an empirical study where human subjects chose between pairs of monetary rewards available after different delays (Green et al., 1994). Subjects preferred a small, short-delayed over a large, long-delayed reward, but their preference reversed away from the small towards the large reward when the delays to both rewards were advanced by the same time interval. Notably, the prolongation of the delays resulted in a preference reversal even though the difference in the delays remained identical (Green et al., 1994). This finding therefore represents a violation of stationarity. Numerous other studies with human subjects (Ainslie, 1975; Logue, 1988; Benzion et al., 1989; Loewenstein, 1992; Kirby and Herrnstein, 1995; Green et al., 1997; Frederick, Loewenstein and O'Donoghue, 2002; McClure et al., 2004; Rohde, 2005), pigeons (Chung and Herrnstein, 1967; Rachlin and Green, 1972; Ainslie, 1974; Green et al., 1981) and rats (Ito and Asaki, 1982; Bennett, 2002) replicated and confirmed these results (cf., Kalenscher and Pennartz, in preparation). Thus, human and non-human animals systematically violate the crucial assumption of inter-temporal consistency of choice. Note that many studies in the animal literature do not defer both choice outcomes equally, but only one outcome is increasingly delayed, whereas the delay to the other outcome remains constant (see e.g. figure 4A). Preference reversals in those cases do not challenge stationarity, as changes in valuations would be predicted for the increasingly delayed outcome, but not the constant outcome.

The fact that human and animal subjects prefer the small, short-term reward over the large, delayed reward when the receipt to the small reward is near, but not when it is in the relatively far future, suggests that short-term rewards are discounted more steeply than long-term rewards. Such asymmetric discounting poses a strong challenge for the postulation of exponential discounting (Lancaster, 1963). Accordingly, as theoretically suggested by Ainslie (1975), and later empirically shown by Mazur (Mazur, 1984, 1987, 1988; Grossbard and Mazur, 1986) and others (Rachlin et al., 1991; Myerson and Green, 1995; Green and Myerson, 1996; Rohde, 2005; Jones and Rachlin, 2006; Laibson, 1997), discounting curves can be better approximated with hyperbolic than exponential or other constant discount functions, as outlined in figure 3. Why do humans and other animals systematically violate such crucial laws in economics?

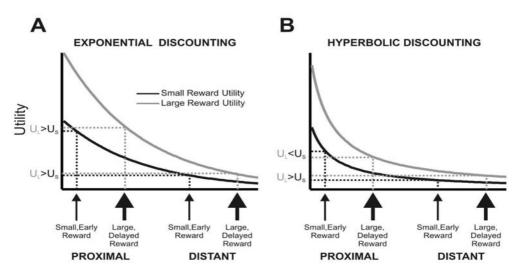


Figure 3. Exponential vs. Hyperbolic discounting of future events. (a) exponential Utility curve of a large, delayed (grey line) and small, short-term reward (black line). With Exponential discounting, stationarity holds because the utility of the large reward (ul) is always higher than the utility of the small reward (us). This is true when both rewards are temporally proximal, or when they are deferred by the same time interval into the future (distant rewards). (b) hyperbolic discounting can explain preference reversals and the violation of stationarity. Due to the steeper decay for short delays, the utility of the small, short-term reward is higher than the large, delayed reward for temporally proximal rewards, but the utility order reverses when both rewards are deferred into the future.

#### Why do we Discount the Future?

An assumption underlying most economic theories is utility maximization. This assumption is shared by many ecological theories of choice, namely the postulation that evolution favors choice mechanisms that maximize fitness levels, and minimize fitness losses. Applied to inter-temporal decision making, this means that the decision maker should act so as to maximize the utility rate, or in ecological terms, the energy intake rate per time unit (Stephens and Krebs, 1986). Rate maximization can explain why humans and animals sometimes prefer a less attractive, but temporally proximal outcome over a more attractive, but temporally remote outcome.

For example, we consider an inter-temporal choice task in which an animal has to choose between a small, always immediate reward and a large reward that is initially also delivered immediately, but that is delayed further as the experimental session progresses. Let's further assume that the large reward is 1.5 times as big as the small reward, and that the next choice opportunity always follows immediately after the animal has consumed its previous reward. The rate maximization hypothesis would predict that the animal should begin the session by preferring the large reward. However, the delay preceding the large reward gets longer and longer over the course of the session. So, at some point, the waiting time for the large reward, and thus the time until the next reward can be realized, will be more than twice as long as the delay preceding the small reward. Naturally, it would make sense now to prefer the small reward, as the hungry animal would be able to consume two small rewards in the same time that it would have to wait for only one large reward. Because two small rewards represent a larger food quantity than one large reward, the animal would maximize its energy intake per time unit by shifting its preference to the small reward once the delay preceding the large reward gets too long. In formal terms, optimal foraging theory assumes that organisms maximize, at least on the long run, the ratio of food intake and the time needed to obtain or consume the food, as described by the following quantity (Stephens and Krebs, 1986):

$$\max \frac{\sum_{i=1}^{\infty} G_i}{\sum_{i=1}^{\infty} t_i}$$
(4)

where  $G_i$  represents the net energy gain obtained from consuming the *i*th food item (here basically corresponding to its amount), and  $t_i$  represents the time between food item *i* and the previous food item *i*-1.

#### Animals do not (always) maximize intake rate

The above example implies that the next choice opportunity follows immediately after receipt of the reward. In many studies (Rachlin and Green, 1972; Ainslie, 1974; Grossbard and Mazur, 1986; Mazur, 1988; Evenden and Ryan, 1996; Cardinal et al., 2000; Isles et al., 2003, 2004; Winstanley et al., 2004, 2006; Kalenscher et al., 2005; Hwang et al., 2006; Louie and Glimcher, 2006), however, the inter-trial interval between reward and next choice was adjusted so that the total trial length was identical in all trials and independent of delay length and other factors. In such a scenario, the rate maximization hypothesis predicts that subjects should always choose the large reward, independent of the delay between response and reward, because only then would the animals maximize the total energy intake per trial, or per experimental session respectively. However, neither pigeons (Rachlin and Green, 1972; Ainslie, 1974; Grossbard and Mazur, 1986; Mazur, 1988; Kalenscher et al., 2005), nor rats (Evenden and Ryan, 1996; Cardinal et al., 2000; Winstanley et al., 2004, 2006; Roesch et al., 2006), mice (Isles et al., 2003, 2004), or monkeys (Hwang et al., 2006; Louie and Glimcher, 2006) show the predicted perseverance on the large reward alternative, but instead reverse their preference to the small, immediate reward once the large reward delay exceeds an individual threshold limit. This shows that the animals' choices depended on the waiting time preceding the rewards, but not on the ratios of reward amount and duration between the rewards, as would be predicted from rate maximization.

In fact, rate maximization models predict that amount and/or delay variations shouldn't play any role in the animals' decisions, because the choices should be only and exclusively directed towards maximizing the rate on the long-term. If, for example, an animal chooses between a fixed medium-term reward and or a variable-delay reward with either short or long delays (variable interval schedule), animals should always choose the option yielding the higher average reward rate. If the average reward rate is identical, animal should be indifferent between both options. However, contrary to this prediction, they usually prefer variable-interval over fixed schedules (Kacelnik and Bateson, 1996), indicating that delay variance does influence an animal's reward preference in addition to other factors, such as reward rate. This variance-proneness is interesting as animals are usually variance-averse if reward magnitude, and not delay, is variable, as explained above. Proneness to delay variance can be explained with hyperbolic discounting (see below).

In summary, animals do not make their choices according to the predictions of rate maximization models. They seem to employ rather short-sighted, waiting-time sensitive choice heuristics, and have a preference for delay variability.

# Preference for delay variability

Hyperbolic discounting, as outlined in figure 3B, can explain the preference for variable over fixed interval schedules. Since, due to the hyperbolic decay, the utility of short-term rewards is disproportionally higher than the utility of medium-term or delayed rewards, but the difference in utility of medium-term and delayed rewards is negligible, the average expected utility of short-term and delayed rewards (variable interval schedules) will exceed the expected utility of fixed medium-term rewards. Hence, animals should prefer variable over fixed delays.

An alternative hypothesis, scalar expectancy theory (SET), can account for both variance aversion when reward magnitude is variable, and variance proneness when reward delay is variable. SET refers to the subjective time and magnitude representation which is normal around the actual means, but as a consequence of Weber's law, has a constant coefficient of variation (ratio of standard deviation to mean). Thus, the combination of an early and a late distribution results in a positively skewed integral, which explains preference for variation in delay (Reboreda and Kacelnik, 1991). Evidence for or against either SET or the hyperbolic discounting account is equivocal (Kacelnik an Bateson, 1997; Bateson and Kacelnik, 1998) and needs to be further tested in the future.

#### Ecological models of inter-temporal decisions: Ecological rationality

In addition to the unclear support, neither of these accounts can explain why animals developed delay sensitivity in the first place: Why does evolution favor choice heuristics that produce suboptimal results in many cases by over-emphasising the delay to the next reward (e.g., through hyperbolic discounting), and ignoring the long-term relevance of time/amount sequences? Obviously, animals may equate delay with collection risk, as outlined in greater detail below. If delays are mentally treated as risks, a risk-averse animal will naturally avoid long delays. However, this doesn't provide an acceptable answer because the question remains why evolution has favoured suboptimal decision rules, be they related to risk avoidance or delay aversion. The first answer that comes to mind is that short-sighted rules

have higher fitness values than long-sighted rules because the animals' constitutions do not allow them to tolerate too long waiting periods. For example, animals with a high metabolism cannot afford to wait too long for a large amount of food, or, in other words, what is the use of high quality, high amount of food if the animal has starved to death while waiting for it? Thus, short-sighted rules may have a certain evolutionary advantage over long-sighted rules.

This would certainly hold if the waiting times were close to the animals' starvation thresholds. However, mice, rats, pigeons, monkeys or other animals used in inter-temporal choice experiments shift their preference away from the economically more advantageous reward even when the waiting time to the larger reward exceeds less than a few seconds, and not hours or days (cf., McDiarmid & Rilling, 1969). Certainly, all those animals would be able to survive longer waiting periods than just a few seconds without food, but nevertheless, they prefer the short-term option over the long-delayed option, even if the long-delayed reward is a multiple of the short-term reward. Such extremely myopic decision patterns are difficult to explain with a fitness advantage of faster available food items. Why does evolution favor such extremely myopic choice heuristics?

Bounded rationality or ecological models, such as the ecological rationality hypothesis (Stephens et al., 2004) claim that choice heuristics that fail to produce maximum fitness in artificial experimental settings do, in fact, perform well in more ecologically valid contexts. For example, Stephens and colleagues (2004) argued that a more ecologically valid choice context entails decisions about limited food resources. A typical decision would consist of whether to entirely exploit all food resources in a given food patch, or leave the food patch early before having consumed all resources, and search for a new patch. The difference between the patch situation and the standard inter-temporal choice task is that, in a standard inter-temporal choice task, an animal has a binary choice between a large, delayed or a small, immediate reward, whereas in the patch situation, it chooses whether to continue to stay in a given patch, or to leave and search for a new patch. Figure 4 illustrates an inter-temporal choice situation (4B).

Put in more formalized terms, an animal consuming a reward in a food patch (the initial food amount,  $A_{Init}$ ) has to decide whether to stay in the current patch and wait for further rewards until the patch is completely depleted, or whether to leave early and initiate a new travel time to the next patch. If it decides to leave, it has to travel for time  $t_{Travel}$  until the next patch is encountered where it receives a new initial reward amount  $A_{Init}$ . If it decides to stay, it obtains additional food rewards of amount  $A_{Stay}$  delivered after a certain waiting time  $t_{Stay}$  until the patch is depleted. It then has to leave the patch as well, and initiate a further travelling time  $t_{Travel}$  until it encounters a new patch and obtains reward amount  $A_{Init}$  in the new patch (see figure 4B).

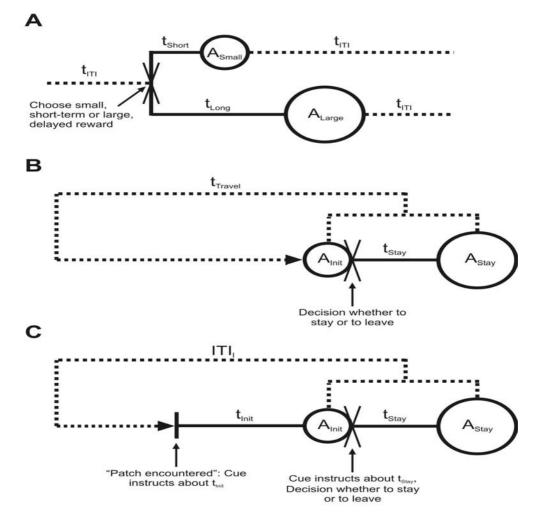


Figure 4. Inter-temporal choice and patch situation. (a) displays a schematic drawing of a standard inter-temporal choice task used in animal research, often also referred to as a 'Self-control' task. Following an inter-trial interval (titi), the animal makes a binary choice between a large and a small reward amount (alarge and asmall), delivered after a long or a short time delay (tlong and tshort). Reward consumption is followed by another iti until the next choice opportunity. The length of the iti, titi, is adjusted to compensate for differences in delay length and choice, so that every trial is of identical duration. (b) patch situation. An animal travels until it encounters a food patch, where it consumes an initial reward amount, ainit. After consumption, the animal has to decide whether to leave the patch and initiate travel time ttravel, until it finds the next patch, or whether to stay in the patch, wait time tstay for an additional food amount, astay, and then leave the patch and initiate the travel time travel to the next food patch. (c) patch situation as used in the experiment by stephens and Anderson (2001). The situation is equivalent to figure 1b, but includes an additional initial waiting time, tinit, preceding the initial reward ainit. In all panels, the cross of thin lines indicates the timepoint where the animal makes its decision.

Staying is more time-consuming  $(t_{Stay}^{Total} = t_{Stay} + t_{Travel})$ , but yields higher reward amounts  $(A_{Stay}^{Total} = A_{Stay} + A_{Init})$ , leaving is less time-consuming  $(t_{Leave}^{Total} = t_{Travel})$  because the animal doesn't need to wait for the additional reward in the old patch, but it also misses out

on that additional reward, and thus receives lower reward quantities ( $A_{Leave}^{Total} = A_{Init}$ ). A farsighted decision rule based on rate maximization would predict that the animal prefers to leave if it gained more rewards per time unit in the leave alternative than in the stay alternative, and it would stay in case it gained more rewards per time in the stay alternative than in the leave alternative. Assuming that travelling to a new patch takes always the same time, and that the yield in a new patch is always identical, an animal would leave if

$$\frac{A_{Init}}{t_{Travel}} > \frac{(A_{Add} + A_{Init})}{(t_{Stay} + t_{Travel})}$$
(5)

and it would predict to stay in the opposite case.

A short-sighted, waiting-time sensitive rule, as observed in most experimental settings on inter-temporal decision-making, would predict that an animal considers only the delay until the next reward in its decision. That is, the rule would predict that an animal prefers to leave if

$$\frac{A_{lnit}}{t_{Travel}} > \frac{A_{Add}}{t_{Stay}}$$
(6)

and it would predict to stay in the opposite case. Because travel time,  $t_{Travel}$ , and initial reward amount,  $A_{Init}$ , are identical in the leave and the stay case, the long-term, rate-maximising rule in expression (5) is algebraically equivalent to the short-term, waiting-time-sensitive, impulsive rule in expression (6). Thus, in the above described patch situation, a short-sighted impulsive choice rule (expression 6) would approximate long-term rate maximization, consistent with predictions from optimal foraging theory.

If this theoretical line of argument was true, then the same short-sighted choice heuristic should produce rate maximization in the patch-situation, but not in the standard intertemporal choice task. To test this prediction, Stephens and Anderson (2001) trained blue jays in a 'self-control' situation and a patch situation. The 'self-control' situation was essentially equivalent to the inter-temporal choice task sketched in figure 4A: Blue jays chose between a small, immediate or a large, delayed reward,  $A_{Small}$  and  $A_{Large}$ , delivered after  $t_{Short}$  and  $t_{Long}$ , by hopping on a perch on the left or right side in a training box. After reward delivery and consumption, they had to leave the perch, and an inter-trial interval (ITI, equivalent to travel time in figure 4B) was initiated, after which they could make their next choice. Instead of a binary choice between large and small rewards, the patch situation (figure 4C) consists of a sequence of choices between smaller rewards potentially summing up to a large reward: a cue instructed the animals about the delay  $t_{lnit}$  to an initial reward of  $A_{lnit}$  amount. By hopping on a randomly activated perch in the box, the initial reward was delivered after the indicated delay. Afterwards, the animals could choose whether to stay and wait for the additional reward by remaining on the perch (a second cue indicated the delay length  $t_{Stay}$  to the additional reward of amount  $A_{Stay}$ .), or whether to leave the perch, miss out on the additional reward, and initiate the ITI, and a new waiting time  $t_{Init}$  to the next small reward. Hence, the patch task resembled the situation illustrated in figure 4B, with the exception that there was an additional waiting time  $t_{Init}$  preceding the initial food reward  $A_{Init}$ . Moreover, in the patch situation, the travel time needed to leave the initial patch and move to the next one corresponds to the ITI of the standard inter-temporal choice task. Note that defined this way, the ITI becomes an integral part of the decision in the patch situation but not in the standard choice situation. This is because a short-term decision rule taking into account only the delay to the next reward would predict that, in the 'self-control' situation, animals consider the delays to the large and the small reward ( $t_{Short}$  versus  $t_{Long}$ ), and that, in the patch situation, they consider the delays to the different rewards in the stay or leave cases ( $t_{Stay}$  vs.  $t_{ITI} + t_{Init}$ ).

It is now possible to choose the task parameters so that the 'self-control' and the patch situation are economically equivalent: The sum of the two rewards in the patch situation equals the large reward amount in the 'self-control' situation, and the sum of the delays to the first and second rewards in the patch situation is equivalent to the delay preceding the large reward in the 'self-control' situation. This means that animals would receive the same amount of food within the same time when they chose to stay in the patch situation or when they chose the large, delayed reward in the 'self-control' situation

Because of this economical equivalence, the animals should always show identical preferences across both situations if their decision rule was purely economical, e.g., farsighted. If, on the other hand, the animals indeed used a short-sighted decision rule, then they should show inconsistent choices under most circumstances, and should maximize their intake rate in the patch, but not in the 'self-control' situation. Such inconsistencies would arise because, first, the ITI is an integral part of the decision in the patch situation (remember

that in the leave case, the total waiting time for the next reward would be  $t_{Init}^{Leave} + ITI$ ), but not in the 'self-control' situation (the delay between choice and reward-delivery is not dependent on the ITI), and, second, because the large reward in the patch situation consists of a series of smaller rewards, and not of a one-shot delivery of one single large reward as in the 'self-control' situation.

In their study, Stephens and Anderson (2001) systematically varied the delays, amounts and ITIs, and tested the preference patterns of their blue jays. As predicted, they reported inconsistencies in choice between the two situations. In particular, they found that the blue jays were overall more ready to maximize their intake rate in the patch situation than in the 'self-control' situation. This supports the main notion of the ecological rationality hypothesis that one and the same short-sighted decision rule results in rate maximization in one, but not the other choice situation.

In summary, theoretical considerations and empirical evidence suggests that evolution may have favored the development of short-sighted choice heuristics because such rules produce long-term rate maximization in many natural situations, but not necessarily in artificial laboratory settings.

#### Ecological models of inter-temporal decisions: Feeding ecology

Although all animals have in common that they discount the future, the rate by which future events are discounted differs dramatically between species. Mice, for example, seem to tolerate waiting times up to only a few seconds (Isles et al., 2003, 2004), capuchin monkeys wait for several minutes (Ramseyer et al., 2006), and humans can wait for months or even years for a relatively attractive reward (Green et al., 1994, 1997). The ecological rationality

hypothesis can explain why evolution has favored the development of impulsive decision rules, but it cannot readily account for these large inter-individual and inter-species differences in delay-tolerance.

Another theory, the feeding ecology hypothesis, aims to explain those differences. It departs from the comparison of discount rates between very similar monkey species: Cottontop tamarins and common marmosets. These two new-world monkeys are similar in terms of social behavior, mating system, life span, life history, home range size, parental care, body size and weight, brain size and weight, and other factors (cf., Stevens et al., 2005a). However, despite those similarities, the animals show very different choice behavior when tested in an adjusting delay procedure. In an adjusting delay procedure (Mazur, 1987), animals choose between a small, short-term reward and a large, delayed reward. After large reward choices, the delay to its receipt is increased in the next trials, after small reward choices, the delay preceding the large reward is decreased. This procedure allows the measurement of indifference points, i.e., the delay length at which the large, delayed reward has equal value to the small, short-term reward. Stevens and colleagues (2005a) found that the marmosets waited considerably longer for a large reward than tamarins. However, when tested in a spatial version of the adjusting feature task, in which travelling distance, but not delay, to a large reward was varied, the pattern reversed: Whereas the tamarins preferred the large reward independent of the travel distance to the large reward, the marmosets preference for the large reward continuously decreased with increasing travel distance (Stevens et al., 2005b). Taken together, space and time affected the monkeys' decisions differentially: compared to tamarins, marmosets were more patient when waiting for a delayed reward, but discounted spatially distant rewards steeper than tamarins.

How come that the two monkeys have evolved so different discounting patterns? Because of their striking similarity in many aspects, differences in metabolism, physical condition, starvation threshold or the like can be ruled out as possible explanations. Stevens et al. (2005a, 2005b) point out that one of the main differences between the two New World monkeys is their diet: Although both species eat fruits, marmosets additionally feed on plant exudates, such as gum and sap, and tamarins feed on insects.

Feeding ecology plays a major role in shaping cognitive and neural functions (e.g., Clayton & Krebs, 1995; Basil et al., 1996; Emery & Clayton, 2001). Accordingly, the differences in foraging behavior between the tamarins and marmosets may prove to be the key evolutionary pressure for the differential development of the temporal and spatial discount rates: Marmosets feed on localized, immobile food sources that do not require far-distance travels. But feeding on gum and sap requires to scratch the bark of the tree and then wait for the sap to exude. For the marmosets, it is therefore essential to be patient (in time), but not necessarily mobile (in space) in order to get most of the slowly exudating sap. Conversely, for the insectivore tamarins, it is crucial to be constantly alert, and react quickly in order not to miss any passing-by insects. Moreover, they feed on dispersed food sources and have to cover rather large territories to find insects. Therefore, in contrast to marmosets, tamarins must be quick and impulsive, and willing to travel relatively far distances to find enough food to survive.

In summary, given the individual differences in foraging behavior, it may be more advantageous for the marmosets to be patient in time, but impulsive in space, and for the tamarins to be impulsive in time, but patient in space. The individual differences in foraging ecology may therefore explain the differential evolution of temporal and spatial discounting. Commonalities and differences between risky and inter-temporal decisions

The preceding parts of this chapter have treated inter-temporal and risky decisions as separate. However, several authors argued that there might not be a real difference between delay and risk because each dimension can be expressed in terms of the other (Mischel, 1966; Stevenson, 1986; Rachlin et al., 1986, 1987, 1991; Mazur, 1989, 1995, 1997; Green and Myerson, 1996, 2004; Sozou, 1998; Hayden and Platt, 2007). A delayed reward might be less likely to occur (at least in natural situations) and therefore its expected value might be lower than that of earlier rewards. Moreover, as the state of the agent might change, the value of a later reward might also be more uncertain due to the unpredictability of the subject's own state, including its own survival. For example, our annual mortality risk is about 1% but was considerably higher in our evolutionary past. There is little use in waiting for a large, delayed reward if we will never experience the reward. Therefore a delayed reward may be equivalent to a risky reward and decision-makers may equate temporal distance with collection-risk (Kacelnik and Bateson, 1996; Sozou, 1998).

Theoretically, the proposal that delay and risk are processed similarly boils down to models that incorporate only either risk attitude or discounting as subjective weighting factors for utility but not both. For example, consider the following model (Mazur, 2007):

$$U = \Sigma P_i (A/I + KD_i) \tag{7}$$

Here, utility is a function of objective probability (P) but subjective time discounting (K). A denotes reward amount and D delay. Conversely, the following model comprises three subjective weighting factors determining utility (Kheramin et al., 2003, adapted):

$$U = \Sigma(1/(1 + Q/q_i) \times 1/(1 + Kd_i) \times 1/(1 + H\theta_i))$$
(8)

where q denotes reward amount, d reward delay and  $\theta$  odds against reward ( $\theta = [1/P_i] - 1$ ). The subjective weighting factors Q, K and H denote subjective sensitivity to reward magnitude, delay and probability, respectively. H>1 corresponds to risk aversion, H<1 to risk seeking, Note, that this model treats probability and delay similarly by using a hyperbolic form for both.

Conversely, instead of treating delayed rewards as uncertain, it has also been proposed that risky rewards may be treated as variably delayed rewards (Rachlin et al., 1986; Mazur, 1989; Hayden and Platt, 2007). For example, take a gamble between two options with equal expected values, but where one option yields a medium-sized, certain reward and the other one yields either a large or a small reward with a 50% chance each. Animals are usually not risk-neutral in those types of tasks (Kacelnik and Bateson, 1996, 1997; Bateson and Kacelnik, 1998; McCoy and Platt, 2005b; Hayden and Platt, 2007). Macaques, for example, generally prefer the risky over the certain option (McCoy and Platt, 2005b; Hayden and Platt, 2007). To explain this risk-proneness, it has been argued that, if an animal consistently sticks with the risky option offering a 50% chance of a large pay-off, they will almost ce rtainly receive the large reward eventually: If not now, then on a future trial. Therefore, the risky option gives a practically guaranteed, though potentially delayed large pay-off. Thus, because probabilistic rewards may be treated as large and delayed rewards, they may recruit similar cognitive mechanisms (Rachlin et al., 1986; Mazur, 1989; Hayden and Platt, 2007).

The empirical evidence supporting commonalities of delay and risk sensitivity is equivocal (Mazur, 1989; Rachlin et al., 1986, 1991; Green and Myerson, 1996, 2004; Estle et al., 2006; Hayden and Platt, 2007). Commonalities are entertained by the occurrence of similar preference reversals when the delay or the probability of reward is increased for both options. Thus, violations of the independence axiom (Allais paradox) with probability are similar to violations of the stationarity axiom with delay (both described above). In other words, the utility of sooner and more probable rewards increases more than that of later and less probable rewards as reward immediacy and probability increase (reviewed in Green and Myerson, 2004). Accordingly, both probability and delay are amenable to hyperbolic discounting functions.

Conversely, there is considerable evidence that risk and delay are processed differentially. For example, humans discount smaller delayed rewards more steeply than larger delayed rewards but discount smaller probabilistic rewards less steeply than larger probabilistic rewards (Du et al., 2002; Estle et al., 2006). Inflation affects decisions involving delayed but not risky monetary rewards (Ostaszewski et al., 1998). Drug addiction appears to affect delay discounting more than risk processing (Reynolds et al., 2004) whereas problem gambling might have the opposite effect (Holt et al., 2003). Even culture appears to influence probability and delay differentially, with Japanese graduate students discounting probabilistic rewards more steeply and delayed rewards less steeply than Chinese students (Du et al., 2002). Further reinforcing the notion that delay differs from risk, earlier rewards may be preferred for several reasons over later rewards over and above to the later reward being riskier or the future subjective state more uncertain (Kacelnik and Bateson, 1996):

Earlier rewards can be put to use and earn compound interest (corresponding to offspring's offspring) before later rewards arrive.

Waiting for a delayed reward may prevent an agent from pursuing other courses of action. This might diminish the value of delayed rewards.

It might be easier to learn about action-reward and stimulus-reward contingencies with earlier rewards because at their arrival the mental representation of the causally relevant antecedents has decayed to a lesser degree compared with later rewards.

With long delays to the later reward and fixed intertrial duration, choosing earlier rewards can maximize the energy intake per unit time even if the later reward is larger than the earlier reward (as explained above on the chapter on rate maximization).

# CONCLUSION

In the preceding sub-chapters, we have outlined a selection of different attempts from different disciplines to explain decisions under risk and inter-temporal choices. Broadly speaking, normative approaches, such as eut or optimal foraging (rate maximization), focus on how an animal should behave in order to meet formulated choice criteria, such as optimal decision-making, utility maximization or consistency of choice. Descriptive and empirical approaches, on the other hand, challenge many of the predictions and implications of normative models, and show that the normative analysis of decision making may not always be consistent with the empirical reality of choosing and acting. For example, humans do not always choose according to the predictions of eut, and, contrary to the prediction of optimal

foraging theory, animals frequently fail to maximize their intake rate. Ecological models, such as the budget rule or the ecological rationality hypothesis, deal with the question why evolution favored the development of choice patterns that often violate the predictions of the normative approaches. In particular, they provide an analysis within an ecologically valid framework of the sense and non-sense of the way animals (and humans) make their decisions.

Last but not least, psychological approaches attempt to identify the choice-mediating Cognitive mechanisms, such as whether animals employ far-sighted vs. Myopic choice Heuristics.

We hope to have illustrated that these different approaches have strong limitations when isolated from each other. In particular, we believe that normative models have high explanatory power, but are of questionable validity if not substantiated with empirical results.

Empirical studies are of potentially higher validity per se, but it is difficult, if not impossible In many cases, to generalize single experimental results to a common framework of choice. They are therefore often of little use to answer the question how we generally make decisions

When presented outside a theoretical context. Moreover, although empirical studies can be used to identify the short-comings of theoretical models, such de-construction is only useful When followed by the formulation of a better theoretical model. In conclusion, neither normative, nor empirical, nor psychological approaches alone can produce useful results, but only the combination of all approaches yields fruitful outcomes.

#### ACKNOWLEDGMENTS

We greatly appreciate the comments of Martin Kocher on an earlier version of the manuscript. TK is supported by a grant from senternovem (bsik-03053).

# REFERENCES

Ainslie G (1974) Impulse control in pigeons. J exp anal behav 21:485-489.

- Ainslie G (1975) Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychol bull* 82:463-496.
- Allais M (1953) Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica* 21:503-546.
- Basil Ja, Kamil Ac, Balda Rp, Fite Kv (1996) Differences in hippocampal volume among food storing corvids. *Brain behav evol* 47:156-164.
- Bateson M, Kacelnik A (1998) Risk-sensitive foraging: decision-making in variable environments. In: *cognitive ecology* (dukas r, ed), pp 297-341. Chicago, Illinois: University of chicago press.
- Bennett Sm (2002) Preference reversal and the estimation of indifference points using a fastadjusting delay procedure with rats. Dissertation, University of florida.
- Benzion U, Rapoport A, Yagil J (1989) Discount rates inferred from decisions: an experimental study. *Management sci* 35:270-284.

- Bernoulli D (1954) Exposition of a new theory on the measurement of risk. *Econometrica* 22:23-36.
- Caraco T, Blanckenhorn WU, Gregory GM, Newman JA, Recer GM, Zwicker SM (1990) Risk-sensitivity: ambient temperature affects foraging choice. *Anim behav* 39:338-345.
- Cardinal RN, Robbins TW, Everitt BJ (2000) The effects of d-amphetamine, chlordiazepoxide, alpha-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology* (berl) 152:362-375.
- Charnov EL (1976) Optimal foraging: the marginal value theorem. *Theor popul biol* 9:129-136.
- Chung SH, Herrnstein RJ (1967) Choice and delay of reinforcement. J exp anal Behav 10:67-74.
- Clayton NS, Krebs JR (1995) Memory in food-storing birds: from behaviour to brain. *Curr* opin neurobiol 5:149-154.
- Du W, Green L, Myerson J (2002) Cross-cultural comparisons of discounting delayed and probabilistic rewards. *Psychol rec* 52:479-492.
- Ellsberg D (1961) Risk, ambiguity, and the savage axioms. *Qj econ* 75:643-669.
- Emery NJ, clayton NS (2001) effects of experience and social context on prospective caching strategies by scrub jays. *Nature* 414:443-446.
- Estle SJ, green L, Myerson J, Holt DD (2006) Differential effects of amount on temporal and probability discounting of gains and losses. *Mem cognit* 34:914-928.
- Evenden JL (1999) Varieties of impulsivity. Psychopharmacology (berl) 146:348-361.
- Evenden JL, Ryan CN (1996) the pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology* (berl) 128:161-170.
- Fehr R (2002) The economics of impatience. Nature 415:269-272.
- Fishburn PC (1982) Nontransitive measurable utility. J math psychol 26:31-67.
- Fishburn PC, rubinstein A (1982) time preference. Internat econ rev 23:677-694.
- Frederick S, Loewenstein G, O'donoghue T (2002) Time discounting and time preference: a critical review. *J econ lit* 40:351-401.
- Friedman M, savage LJ (1948) The utility analysis of choices involving risk. *J polit Econ* 56:279-304.
- Glimcher PW (2002) Decisions, decisions, decisions: choosing a biological science of choice. *Neuron* 36:323-332.
- Glimcher PW, Rustichini a (2004) Neuroeconomics: the consilience of brain and decision. *Science* 306:447-452.
- Green L, Fisher EB jr, Perlow S, Sherman 1 (1981) Preference reversal and self- control: choice as a function of reward amount and delay. *Behav anal lett* 1:43-51.
- Green l, fristoe n, myerson j (1994) temporal discounting and preference reversals In choice between delayed outcomes. Psychonom bull rev 1:383-389.
- Green L, Myerson J (1996) exponential versus hyperbolic discounting of delayed outcomes: risk and waiting time. *Amer zool* 36:496-505.
- Green L, Myerson J (2004) a discounting framework for choice with delayed and probabilistic rewards. *Psychol bull* 130:769-792.
- Green L, Myerson J, McFadden E (1997) Rate of temporal discounting decreases with amount of reward. *Mem cognit* 25:715-723.

- Grossbard CL, Mazur JE (1986) A comparison of delays and ratio requirements in selfcontrol choice. *J exp anal behav* 45:305-315.
- Hamilton W D (1964) The genetical evolution of social behaviour. J theor biol 7:1-52.
- Hayden BY, Platt ML (2007) Temporal discounting predicts risk sensitivity in rhesus macaques. *Curr biol* 17:49-53.
- Holt DD, Green L, Myerson J (2003) Is discounting impulsive? Evidence from temporal and probability discounting in gambling and nongambling college students. *Behav processes* 64:355-367.
- Hwang J, Kim S, Lee D (2006) Neuronal signals related to delayed reward and its discounted value in the macaque dorsolateral prefrontal cortex. *Annual meeting of the Society for neuroscience, program no.* 71.7.
- Isles AR, Humby T, Walters E, Wilkinson LS (2004) Common genetic effects on variation in impulsivity and activity in mice. *J neurosci* 24:6733-6740.
- Isles AR, Humby T, Wilkinson LS (2003) Measuring impulsivity in mice using a novel operant delayed reinforcement task: effects of behavioural manipulations and d-Amphetamine. *Psychopharmacology (berl)* 170:376-382.
- Ito M, Asaki K (1982) Choice behavior of rats in a concurrent-chains schedule: amount and delay of reinforcement. *J exp anal behav* 37:383-392.
- Jones B, Rachlin H. (2006) Social discounting. Psychol sci 17, 283-286.
- Kacelnik a (1997) Normative and descriptive models of decision making: time discounting and risk sensitivity. *Ciba found symp 208:51-67; discussion 67-70.*
- Kacelnik A, Bateson M (1996) Risky theories the effects of variance on foraging decisions. *Amer zool* 36:402-434.
- Kacelnik Aa, Bateson M (1997) Risk-sensitivity: crossroads for theories of decision- making. *Trends cogn sci* 1:304-309.
- Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47:263-291.
- Kalenscher T, Ohmann T, Gunturkun O (2006) The neuroscience of impulsive and selfcontrolled decisions. *Int j psychophysiol* 62:203-211.
- Kalenscher T, Windmann S, Diekamp B, Rose J, Güntürkün O, Colombo M (2005) Single units in the pigeon brain integrate reward amount and time-to-reward in an Impulsive choice task. *Curr biol* 15:594-602.
- Kalenscher T, Pennartz CMA (2007) A bird in the hand is worth two in the future. *The Neuroeconomics of inter-temporal decision making*. In preparation.
- Kheramin S, Body S, Ho M, Velazquez-Martinez DN, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2003) Role of the orbital prefrontal cortex in choice between delayed and uncertain reinforcers: a quantitative analysis. *Behav processes* 64:239-250.
- Kirby KN, Herrnstein RJ (1995) Preference reversals due to myopic discounting of delayed reward. *Psychol sci* 6:83-89.
- Koopmans TC (1960) Stationary ordinal utility and impatience. Econometrica 28:287309.
- Laibson DI (1997) Golden eggs and hyperbolic discounting. Q j econ 42: 861-871.
- Lancaster K (1963) An axiomatic theory of consumer time preference. *Internat econ Rev* 4:221-231.
- Loewenstein G (1992) The fall and rise of psychological explanations in the economics of intertemporal choice. In: *choice over time*, pp. 3-34. Eds G. Loewenstein, J. Elster. Russell sage foundation: New York.

- Logue AW (1988) Research on self-control: an integrating framework. *Behav brain Sci* 11:665-709.
- Loomes G, Starmer C, sugden R (1991) Observing violations of transitivity by experimental methods. *Econometrica* 59:425-439.
- Loomes G, Sugden R, (1982) Regret theory: an alternative theory of rational choice under uncertainty. *Econ j* 92:805-824.
- Louie K, Glimcher PW (2006) Temporal discounting activity in monkey parietal neurons during intertemporal choice. *Annual meeting of the society for neuroscience*, Program no. 605.5.
- Machina MJ (1982) "Expected utility" analysis without the independence axiom. *Econometrica* 50:277-323.
- Mankiw, NG (2002) Macroeconomics. 5th Edition, Worth publishers: New York
- Mazur JE (1984) Tests of an equivalence rule for fixed and variable reinforcer delays. *J exp* psychol: anim behav process 10:426-436.
- Mazur JE (1988) Estimation of indifference points with an adjusting-delay procedure. *J exp* anal behav 49:37-47.
- Mazur JE (1989) Theories of probabilistic reinforcement. J exp anal behav 51:87-99.
- Mazur JE (1995) Conditioned reinforcement and choice with delayed and uncertain primary reinforcers. *J exp anal behav* 63:139-150.
- Mazur JE (1997) Choice, delay, probability, and conditioned reinforcement. *Anim Learn behav* 25:131-147.
- Mazur JE (2007) Species differences between rats and pigeons in choices with probabilistic and delayed reinforcers. *Behav processes* 75:220-224.
- Mazur JE, Commons ML, Mazur JE, Nevin JA, Rachlin H (1987) An adjusting procedure for studying delayed reinforcement. In *Quantitavie Analyses of Behaviour: Foraging*. pp 55-73. Hillsdale, NJ: Erlbaum.
- McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004) Separate neural systems value immediate and delayed monetary rewards. *Science* 306:503-507.
- McCoy AN, Platt ML (2005a) Expectations and outcomes: decision-making in the primate brain. *J comp physiol a neuroethol sens neural behav physiol* 191:201-211.
- McCoy AN, Platt ML (2005b) Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat neurosci* 8:1220-1227.
- McDiarmid CG, Rilling ME (1965) reinforcement delay and reinforcement rate as determinants of schedule preference. *Psychon sci* 2:195-196.
- McNamara JM, Houston AI (1986) The common currency for behavioral decisions. *Am nat* 127:358-378.
- Mischel W (1966). Theory and research on the antecedents of self-imposed delay of reward. In ba maher (ed.), *progress in experimental personality research*: vol. 3 (pp. 85-132). New york: academic press.
- Moore FR, Simm PA (1986) Risk-sensitive foraging by a migratory bird (dendroica coronata). *Experientia* 42:1054-1056.
- Myerson J, Green L, (1995) discounting of delayed rewards: models of individual choice. J exp anal behav 64, 263-276.
- Ostaszewski P, Green L, Myerson J (1998) Effects of inflation on the subjective value of delayed and probabilistic rewards. *Psychon bull rev* 5:324-333.
- Powell K (2003) Economy of the mind. Plos biol 1:e77.

- Rachlin h, castrogiovanni a, cross d (1987) probability and delay in commitment. J Exp anal behav 48:347-353.
- Rachlin H, Green L (1972) Commitment, choice and self control. J exp anal behav 17:15-22.
- Rachlin H, Logue AW, Gibbon J, Frankel M (1986) Cognition and behavior in studies of choice. *Psychol rev* 93:33-45.
- Rachlin H, Raineri A, Cross D (1991) Subjective probability and delay. *J exp anal Behav* 55:233-244.
- Ramseyer A, Pele M, Dufour V, Chauvin C, Thierry B (2006) Accepting loss: the temporal limits of reciprocity in brown capuchin monkeys. *Proc biol sci* 273:179-184.
- Reboreda JC, Kacelnik A (1991) Risk sensitivity in starlings: variability in food amount and food delay. *Behav ecol* 2:301-308.
- Reynolds B, De Wit H, Richards J (2002) Delay of gratification and delay discounting in rats. *Behav processes* 59:157-168.
- Reynolds B, Richards JB, Horn K, Karraker K (2004) Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behav Processes*, 65:35-42.
- Rohde K (2005) The hyperbolic factor: a measure of decreasing impatience. *Research Memoranda 044*. Maastricht research school of economics of technology and Organization: Maastricht.
- Samuelson PA (1937) A note on measurement of utility. Rev econ stud 4:155-161.
- Shafir S (1994) Intransitivity of preferences in honey bees: support for 'comparative' evaluation of foraging options. *Anim behav* 48:55-67.
- Sibly RM, McFarland DJ (1976) On the fitness of behaviour sequences. *Am nat* 110:601-617.
- Sozou PD (1998) On hyperbolic discounting and uncertain hazard rates. *Proc r soc Lond b* 265:2015-2020.
- Stephens DW (1981) The logic of risk-sensitive foraging preferences. *Anim behav* 29:628-629.
- Stephens DW, Anderson D (2001) The adaptive value of preference for immediacy: when shortsighted rules have farsighted consequences. *Behav ecol* 12:330-339.
- Stephens DW, Kerr B, Fernandez-Juricic E (2004) Impulsiveness without discounting: the ecological rationality hypothesis. *Proc r soc lond b biol sci* 271:2459-2465.
- Stephens DW, Krebs JR (1986) Foraging theory. *Monographs in behavior and ecology*. Princeton. Princeton University press: New Jersey.
- Stevens JR, Hallinan EV, Hauser M (2004) The ecology and evolution of patience in two new world monkeys. *Biol lett* 1:223-226.
- Stevens JR, Rosati AG, Ross KR, Hauser MD (2005) Will travel for food: spatial discounting in two new world monkeys. *Curr biol* 15:1855-1860.
- Stevenson MK (1986) A discounting model for decisions with delayed positive and negative outcomes. *J exp psychol: general* 115:131-154.
- Strotz RH (1955) Myopia and inconsistency in dynamic utility maximization. *Rev Econ stud* 23:165-180.
- Tobler PN, Fletcher PC, Bullmore ET, Schultz W (2007a) Learning-related human brain activations reflecting individual finances. *Neuron* 54:167-175.
- Tobler PN, O'Doherty JP, Dolan RJ, Schultz W (2007b) Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J neurophysiol* 97:1621-1632.

Tversky A (1969) Intransitivity of preferences. Psych rev 76: 31-48.

- Tversky A, Kahneman D (1986) Rational choice and the framing of decisions. *J business* 59:s251-s278.
- Tversky A, Kahneman D (1992) Advances in prospect theory: cumulative representation of uncertainty. *J risk uncert* 5:297-323.
- Von Neumann J, Morgenstern O (1944) *Theory of games and economic Behavior*. Princeton. Princeton University Press: New Jersey.
- Waite TA (2001) Intransitive preferences in hoarding gray jays (perisoreus Canadensis). *Behav ecol sociobiol* 50:116-121.
- Winstanley CA, Theobald DE Cardinal RN, Robbins TW (2004) Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J neurosci* 24:4718-4722.
- Winstanley CA, Theobald DE, Dalley JW, Cardinal RN, Robbins TW (2006) Double dissociation between serotonergic and dopaminergic modulation of medial prefrontal and orbitofrontal cortex during a test of impulsive choice. *Cereb cortex* 16:106-114.
- Wolf M, Van Doorn GS, Leimar O, Weissing FJ (2007) Life-history trade-offs Favour the evolution of animal personalities. *Nature* 447:581-584.

Chapter 28

# **DECIDING WITH WHOM TO MATE: DO FEMALE FINCHES FOLLOW FASHION?**\*

# John P. Swaddle and Earl Clelland

Institute for Integrative Bird Behavior Studies, Biology Department, College of William and Mary, Williamsburg, VA 23185, USA

# ABSTRACT

Choosing an appropriate mate is one of the most important decisions that any animal has to make. The traditional view in non-human systems is that animals are largely slaves to their genes and an individual's mate choice is handed down from their parents. However, in recent years it has become clear that many animals show active decision making in who to mate with and that females may copy mate preferences from other females in the population. In other words, females' mating decisions are affected by the current fashion in their population. Here, we explore whether "mate choice copying" occurs in a model monogamous mating system—the zebra finch. Females were given the opportunity to observe another female courting a particular type of male (we manipulated male appearance by placing small colored leg bands on each bird). In preference tests, our focal females significantly shifted their mate preferences towards the type of male that they had observed as being courted by other females. Therefore, female finches do seem to copy mate preferences, implying that there is social inheritance of information that fundamentally affects mating decisions. This is one of the first demonstrations of mate choice copying in any monogamous system and implies that many other birds may also use social information to affect their mating decisions. We need to rethink evolutionary models of mate choice and sexual selection incorporating this form of social decision making process.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

## INTRODUCTION

Although selecting a mate is one of the most important decisions in an animal's life, there is surprisingly little discussion and exploration of decision making in traditional studies of mate choice (e.g., Andersson 1994; Shuster & Wade 2003; Andersson & Simmons 2006). Evolutionary studies of mate choice are dominated by gene-centered explanations of amongindividual variation in mate preferences (Brooks & Endler 2001; Kokko et al. 2002; Kokko et al. 2003; Shuster & Wade 2003; Andersson & Simmons 2006), stating that mate choice is largely determined by genetically inherited factors. However, it is becoming increasingly clear that mate choice and the benefits of choosing an appropriate mate are more plastic than commonly thought, varying within-individuals as well as among-individuals (Patricelli et al. 2002; Rodriguez & Greenfield 2003; Welch 2003; Greenfield & Rodriguez 2004; Lynch et al. 2005; Lynch et al. 2006). Plasticity is the manifestation of how environmental factors increase variation (in both additive and epistatic effects) above and beyond any genetic contribution to variation in mate preferences. Even this approach to mate choice tends to ignore actual decision making processes and largely represents plasticity as either environmentally determined noise in mate preference or a correlate of changes in life stage and physiology, rather than as the result of definable cognitive process that alter mate choice.

Recently, behavioral ecologists have started to merge the individual level processes of cognitive psychology with the population level processes of evolutionary biology to explain how mate choice is affected by decision making. One example of this is the study of mate choice copying (Pruett-Jones 1992; Dugatkin 1996a; Schlupp & Ryan 1997; Brooks 1998, 1999; Freeberg et al. 1999; Galef & White 2000; Westneat et al. 2000; White & Galef 2000; Swaddle et al. 2005; Uehara et al. 2005). In mate choice copying, the choosing individual (classically females) copies the mate preferences they observe in the population and expresses this copied preference through mate choice. In other words, mate preferences are inherited through social cues that are filtered through cognitive processes (White & Galef 2000; Swaddle et al. 2005). Females view what kind of male is successful and make a decision to mate with that type of male.

Mate choice copying has been demonstrated in a handful of polygynous or lekking species, but has only recently been explored in a monogamous species (Doucet et al. 2004; Swaddle et al. 2005). It makes sense for polygynous females to copy a mate preference as it would be in a female's interest to find quickly the few high quality males in the population. Mate choice copying promotes rapid acquisition of information that relates to mate quality. In contrast to polygynous mating systems, many males will be mated in a monogamous mating system and mate quality will be more evenly distributed across the population. Therefore, mate choice copying may not skew mating success substantially in a monogamous species and, hence, not be strongly selected for through sexual selection. However, mate choice copying could still be selected for in monogamous species where there is the opportunity to observe other mated individuals, such as in colonial breeders, and/or where the costs of developing an independent mate choice are high (Pruett-Jones 1992; Stohr 1998; White & Galef 1999). In other words, mate choice copying by monogamous females may be a cheap way of getting reliable information about male quality.

Even though we are currently unsure about the evolutionary origins or consequences of mate choice copying (Kirkpatrick & Dugatkin 1994; Laland 1994; Agrawal 2001), it is clear

that animals must engage in various elements of decision making to copy mate preferences. For example, the age and sexual experience of individuals affects how females copy mate preferences (Dugatkin & Godin 1993; Ophir & Galef 2004; Amlacher & Dugatkin 2005). In addition, the information that females learn about males through mate choice copying can be generalized to new males (White & Galef 2000; Godin et al. 2005; Swaddle et al. 2005). In other words, if a hypothetical female sees males with red bills as being chosen by other females, then the copying female will be more likely to favor any male with a red bill, not just those particular individual males she saw as chosen. Hence, in some species, copying females are able to internalize copied information and make mate choice decisions dependent on the age, aggression, and sexual experience of the demonstrating female or the copied males. These are clearly cognitively complex decision making processes that push mate choice well beyond the confines of genetically inherited mate preferences.

Previously, we have shown that the monogamous zebra finch can copy mate preferences and that females can generalize information about preferred males to affect future mate preferences (Swaddle et al. 2005). This was the first convincing evidence of mate choice copying in a monogamous species (Brown & Fawcett 2005). Here, we report an experimental study further investigating how female zebra finches make decisions about mate preference and how mate choice copying is affected by social cues.

Our previous evidence for mate choice copying by female zebra finches relied on test (i.e., observer) females observing demonstrator females actually mating and starting to build nests with males for a two week period (Swaddle et al. 2005). Test females copied the preference for types of males they saw as being mated with other females. In the natural ecology of this species, archetypal copying females are probably unmated for shorter periods than two weeks (Zann 1996). Also, in other monogamous species, it is more likely that a potentially copying female would observe courtship between demonstrator females and males rather than prolonged periods of actual mating (i.e, copulations and nest building). Therefore, we investigated whether observing courtship for short bouts was sufficient to elicit mate choice copying in female zebra finches. In addition we investigated whether a known preexisting preference for physical symmetry, manipulated by placing colored plastic leg bands on the males' legs in symmetric and asymmetric arrangements (Swaddle & Cuthill 1994a; Swaddle & Cuthill 1994b; Swaddle 1996), could be eroded by mate choice copying. In general, it is unclear how the decision making associated with mate choice copying can override or accentuate pre-existing (e.g., genetically inherited or sexually imprinted) mate preferences (Dugatkin & Godin 1992; Dugatkin 1996b).

Specifically, we examined how female zebra finches' preferences for symmetrically and asymmetrically leg banded males changed from before to after exposure to courting females and males. We predicted that females would show a general preference for symmetrically banded males before the observation period, consistent with previous studies (Swaddle & Cuthill 1994b; Swaddle 1996). We also predicted that females would increase their preference, from before to after the observation period, for males wearing the band patterns that they observed being courted by other females. In other words, we predicted that observation of conspecific courtship would be sufficient to change mate choice decisions in female zebra finches.

# METHODS

#### Experimental Subjects and General Housing Conditions

We used 24 virgin adult male zebra finches, 15 virgin adult test (observer) females, and eight virgin adult demonstrator females in this study. Birds were randomly selected from our outbred zebra finch colony and were either one or two generations from wild caught stock. Males and females were housed in visual but not acoustic isolation from each other in same sex group cages at approximately 20 °C. The males were housed in groups of three while the observer and demonstrator females were housed in groups of four. The birds were housed in wire cages (approximately 60 x 30 x 40 cm) and provided nutritionally complete seed and water *ad libitum*. The birds were kept on a 14:10 light:dark photoperiod under full spectrum lighting to maintain their readiness to breed (Zann 1996). None of the birds had prior experience with other individuals in the study.

The experiment was separated into three phases. First, we assessed test females' preferences for males wearing symmetric and asymmetric arrangements of red and yellow plastic leg bands. Then test females observed demonstrator females display apparent preferences for new males wearing particular arrangements of these same leg bands. Finally, we tested whether test females' altered their mate preferences in favor of males wearing the leg band arrangements that demonstrator females preferred. In other words, we tested whether test females from the demonstrator females.

#### Pre-Observation Mate Preference Trials

We assessed test females' preference for males wearing three arrangements of red and yellow plastic leg bands, in a three chamber preference apparatus (Figure 1). There were three arrangements of plastic leg bands: right asymmetric, left asymmetric, and symmetric. (a) In the right asymmetric arrangement the leg bands were positioned so that males wore three units of red color and one of yellow on their right leg, while wearing three units of yellow and one of red on their left leg (Figure 2a). (b) This band arrangement was mirror reflected for the left asymmetric treatment group (Figure 2b). (c) In the symmetric leg band arrangement each male wore two units of red and two of yellow on each leg, with the red part of the bands in the center of the arrangement (Figure 2c). It is important to note that there was the same amount of red and yellow color in each band treatment group, hence reducing the effect that particular colors would have on mate preferences.

To commence a mate preference trial, a test female was placed in the preference apparatus for two hours to acclimate to the cage (Figure 1). After the acclimation period, one cage of three males was randomly selected and banded according to each of the three leg band arrangements (i.e., one wore the right asymmetric treatment, one wore the left asymmetric treatment, and one wore the symmetrically arranged bands). These three males were randomly assigned to display cages in the preference apparatus to minimize positional bias across the series of preference trials. An opaque curtain that temporarily separated the display cages from the female part of the chamber was removed so that the female could observe the display males. The female's cage was arranged so that she could view only one male at a time (Figure 1). There were opaque dividers between male display cages so that males could not visually interact with each other. During preference trials, all birds had *ad libitum* access to seed and water.

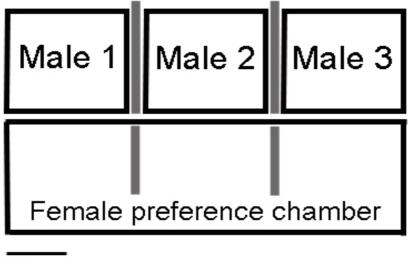




Figure 1. Plan view of the mate preference apparatus. Each of the stimulus males were placed in a small cage, separated by opaque barriers (gray bars). The test female was placed in a long chamber so she could observe each of the stimulus males. Each observation compartment was separated with an opaque barrier so that a female could see only one male at a time. There were abundant perches throughout all the cages. We used the proportion of time females spent displaying in the compartment immediately in front of each male as an index of female preference. All birds had *ad libitum* access to food and water throughout trials.

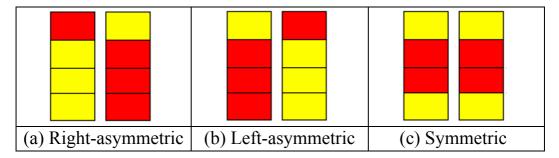


Figure 2. Cartoon representations of the three colored plastic leg band arrangements used for this study: (a) the right-asymmetric; (b) left-asymmetric; and (c) symmetric band treatments. These small plastic leg bands were easily removed and replaced on birds between trials.

After a further 10 minute acclimation period, we videotaped (with a Sony digital video camera) all interactions among the birds for a 1-hour preference trial. We analyzed the tapes to record the amount of time a female spent performing ritualized display behaviors (short hops) in front of each male (Zann 1996; Swaddle et al. 2005). Quantification of this behavior is known to reflect actual mate choice in larger aviary cages and in the wild (Burley 1988; Swaddle & Cuthill 1994b; Swaddle 1996; Zann 1996). We used the relative amount of time a

female spent displaying in front of each male as a measure of her mate preference for each leg band treatment. Each female experienced one mate preference trial and then was returned to her housing cage.

### **Observation Trials**

Observation (mate choice copying) trials were conducted in a modified preference apparatus. We placed three female demonstrator cages between the three male display cages and the larger test female observation cage (Figure 3). The demonstrator female could visually interact with only the single male she was placed in front of; she could not move into any other cages of the observation apparatus. This arrangement of cages was intended to simulate the demonstrator female displaying a preference for a particular male (i.e., the male she was placed in front of) over the other males. The test female was free to observe this female and all the males; hence, she could gain information about which male was apparently preferred over other males.

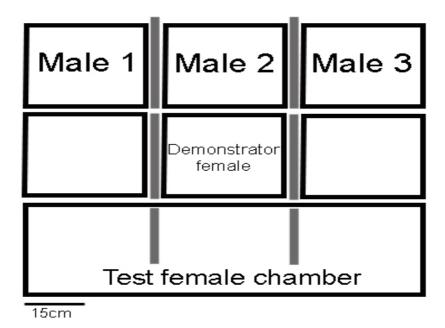


Figure 3. Plan view of the test observation chamber. Each of the stimulus males were placed in a small cage, separated by opaque barriers (gray bars), as in Figure 1. A demonstrator female was placed in one of three small cages in front of each of the stimulus males. A demonstrator female could see only one stimulus male. The test (observing) female was placed in a long chamber so she could see each of the stimulus males (one at a time) and also see the demonstrator female courting one of these stimulus males. The position of the demonstrator female changed between trials (see Methods section for more details). Here we have illustrated an example where the demonstrator female is placed in the middle cage. All birds had *ad libitum* access to food and water throughout trials.

The 15 test females were randomly allocated to two groups: one that was reinforced to prefer the right-asymmetric males in observation trials (N = 7), and one that observed left-asymmetric males as being preferred (N = 8). To begin an observation trial, we randomly

selected a cage of three males to serve as stimuli for the observations and banded them as before. Next we placed a randomly selected demonstrator female in the appropriate demonstrator cage (according to which type of male should appear to be preferred). Then we introduced the test female, allowed for a 10 minute acclimation period and videotaped the test female's activity for a 1-hour observation trial. These videotapes verified that test females viewed all of the males during each observation trial. Every test female experienced ten 1-hour observation trials, only nearly consecutive days. For each observation trial, females observed different males than they had experienced in the pre-observation mate preference trials. Following each observation trial, all birds were returned to their housing cages.

#### Post-Copying Mate Preference Test

A new set of 24 males (additional to the original 24) was used in the post-copying preference test trials so that the test females' preferences were not confounded with familiarity with particular males. We followed the same procedure as for the pre-observation mate preference trials, with display males being randomly assigned to each of the three leg band treatments (right-asymmetric, left-asymmetric, and symmetric; Figure 2) and randomly assigned to cage positions in the preference apparatus (Figure 1). Again, we analyzed the videotapes to discern test females' preferences for males wearing these leg band treatments.

#### Statistical Analyses

All proportional preference data were arc-sine square-root transformed to improve normality. We tested for differences in the pre-observation preferences among leg band treatments using a one-way ANOVA. We compared pre-observation to post-observation leg band preferences with a paired *t*-test to determine whether test females' preferences shifted toward the band arrangement that they were reinforced to copy. All statistical tests were performed with SPSS v.13 and employed two-tailed tests of significance.

## RESULTS

In the pre-observation mate preference trials, females consistently preferred the symmetrically banded males over the asymmetric males ( $F_{2,42} = 3.61$ , P = 0.036; Figure 4). This is consistent with previous data concerning general symmetry preferences among female zebra finches (Swaddle & Cuthill 1994a; Swaddle & Cuthill 1994b; Bennett et al. 1996; Swaddle 1996; Waas & Wordsworth 1999).

Test females significantly changed their mate preference from pre- to post-observation trials, with test females shifting band preferences toward the arrangement that was courted by demonstrator females in the observation trials ( $t_{14} = 2.48$ , P = 0.026; Figure 5). In other words, test females who observed an apparent preference for right-asymmetric banded males shifted their preference toward this type of male, and females who observed an apparent preference for that type of male.

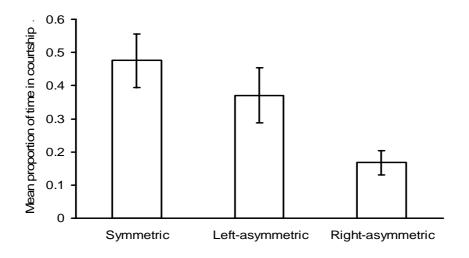


Figure 4. Mean ( $\pm$  s.e.m.) preference for leg band arrangements in the pre-observation mate preference trials. Before observation trials, females had a general preference for males wearing the symmetric band arrangement

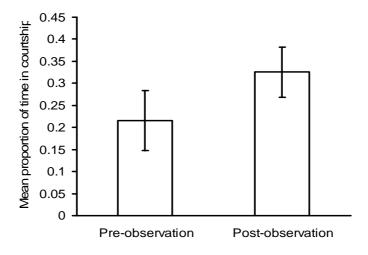


Figure 5. Mean ( $\pm$  s.e.m.) preference for males wearing the reinforced (i.e., courted by another female in the observation trials) leg band arrangement in the pre- and post-observation trials. Females significantly increased their preference for novel males wearing the reinforced band arrangement (whether that was the left- or the right-asymmetric band treatment) indicating that they copied apparent mate preferences displayed in the observation trials.

# CONCLUSION

Our study indicates that observations of short bouts of stereotypical courtship are sufficient to stimulate mate choice copying in female zebra finches. This is important as unmated female zebra finches are very likely to observe other courting females over several days in natural conditions. These birds breed in small colonies where fledglings mature quickly (in a matter of months) and quickly join the breeding population (Zann 1996). A young female, maturing into her first breeding attempt, will consistently be surrounded by older breeding females from which she could copy a mate preference. Therefore, our results indicate that social information could affect mate choice decision making in wild birds. At the very least, we have shown that female zebra finches' mate preference decisions are altered by information about who other females court in the local population. Females shift their preference toward the phenotype of other courted males—they follow the current fashion in mate preferences. It is also relevant that observing courtship appears sufficient to elicit mate choice copying in other species, such as quail *Coturnix coturnix* (White & Galef 1999, 2000) and the guppy *Poecilia reticulata* (Dugatkin 1996b; Amlacher & Dugatkin 2005). However, our study is the first to show that courtship is a cue which guides mate choice decisions in a monogamous species.

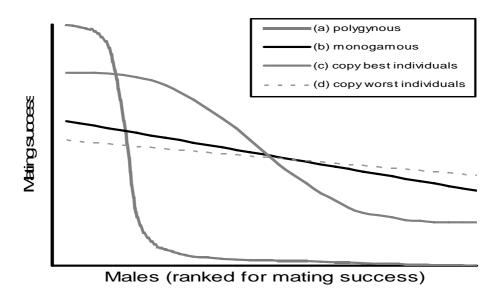


Figure 6. Hypothetical mating skew in (a) polygynous and (b) monogamous mating systems. In the polygynous situation a small number of males get most of the matings. In the monogamous mating system most males will get a mate in any one breeding season; hence, the slope of mating skew is much more shallow than in the polygynous situation. The slope of the mating skew curve indicates the strength of sexual selection acting on males. In the monogamous situation, if mate choice copying shifts mating toward the already successful males, the mating skew and strength of sexual selection can increase, as indicated by line (c) on the graph. However, if mate choice copying favors the least successful males, the mating skew could become flatter and sexual selection weaker, as indicated by line (d) on the graph. Therefore, this simple model can help visualize the outcome of decision making on the strength of sexual selection in a monogamous mating system where mate choice copying occurs.

As test females shifted their preference toward new males who wore the same band patterns as the courted males in the observation period, we provide further evidence that female zebra finches can generalize copied preferences to new males (Swaddle et al. 2005). In a species that mates for life (i.e., a male-female pair-bond is commonly only broken by death or emigration of either partner) this is an important finding, as mating can then be skewed

toward unpaired (i.e., available) males who happen to share phenotypic characters with mated (or courted) males. Potentially, this could skew mating preferences toward particular male phenotypes, therefore increasing the strength of sexual selection for those particular traits (see Figure 6). An unanswered question is whether copied preferences for male traits can also skew extra-pair mating decisions (i.e., copulations and fertilizations outside the pair-bond). Although wild zebra finches show relatively low levels of extra-pair paternity (Birkhead et al. 1990), it is still possible that social information could mediate this small degree of fitness variation—which could have an effect on evolutionary processes. We postulate that social observations of courtship and/or copulations among other conspecifics may be a significant factor in affecting extra-pair mating decisions in other species. As yet, this question is unexplored from both empirical and theoretical perspectives.

Consistent with previous studies (Swaddle & Cuthill 1994b; Swaddle 1996), female finches preferred symmetrically banded males (Figure 4). Importantly, the pre-observation symmetry preference gave us the opportunity to examine how social information can affect a pre-existing mate preference. At least in terms of leg band symmetry preferences, copying of preferences indicated by biased courtship significantly eroded a pre-existing symmetry preference. Therefore, mate choice copying could change mate choice decisions sufficiently to alter evolutionary selection pressures; thereby increasing the evolutionary significance of this form of decision making.

Our study, in general, sheds light on the evolutionary significance of the decision making associated with mate choice copying. As stated above, our results indicate that social information could skew mating preferences consistently toward particular male phenotypes. However, the pattern and strength of this change in mating skew depends on cognitive processes. If female finches decide to copy the most successful (i.e., fittest) phenotype, then the slope of mating skew could easily increase (Figure 6). An increase in mating skew would result in stronger sexual selection pressures for that male phenotype. However, if mate preferences are copied at random, or the least successful phenotypes are copied, then the slope (and sexual selection pressure) could be reduced in some scenarios (Figure 6). As, by definition, the most successful males are the ones that are most likely to provide the courtship cues that could be copied, we hypothesize that the form of copying process we have indicated here has the potential to increase the strength of sexual selection, even in a monogamous species.

However, if the copied preferences are ephemeral and are quickly forgotten, or are easily replaced with other copied preferences, then the evolutionary implications are diminished. Copied preferences have to last over at least one genetic generation to have an evolutionary effect, but last for longer for noticeable directional change. Hence, copied preferences must also likely be stored in some form of longer term memory, further emphasizing the importance of exploring the cognitive processes involved in making a mate choice decision. Outside of one study reporting that copied preference can last for 24 hours in female sailfin mollies *Poecilia latipinna* (Witte & Massmann 2003), there is very little known about the erosion and stability of copied mate preferences. Our own preliminary data from a follow-up study of female zebra finch preferences indicates that copied preferences can last for over a month and, hence, can last between consecutive nesting attempts (J. P. Swaddle, unpublished data).

At this stage in our explorations it is not clear whether mate choice copying is adaptive in zebra finches. It is possible that what we have documented in this chapter and our previous

study (Swaddle et al. 2005) is a by-product of other evolved cognitive and decision making processes. For example, zebra finches show sexual imprinting at a young age (Oetting & Bischof 1996; Bischof & Rollenhagen 1999; ten Cate & Vos 1999), in other words the mate preferences and sexual displays they exhibit at maturity are affected by the social (parental) environment in which chicks are raised. Perhaps these early age processes result in the ability of mature adults to acquire information about mate preferences, but that mate choice copying itself has not been directly selected for. Mate choice copying could be a by-product of selection for sexual imprinting.

Despite how mate choice copying evolved it is clear that these small birds use socially available courtship cues to develop their mate preferences. They appear to make decisions about which males to pay attention to and, further, they generalize the appearance of these apparently preferred males and apply that information to assessing new, unpaired males. Therefore, the experimental results we report here are an important step toward showing that mate choice copying and complex decision making play important roles in establishing mate preferences in a monogamous species. To extrapolate beyond the zebra finch, it may be that this form of decision making is much more prevalent than we commonly think. The zebra finch is a classic monogamous species, with low rates of extra-pair paternity (Birkhead et al. 1990; Burley et al. 1996). In other words, this species is both socially and genetically monogamous. Mate choice copying is not expected to be common in such situations and, hence, may be even more common in other socially monogamous species that show higher levels of extra-pair paternity (genetic polygamy). This would mean that mate choice copying and social decision making could occur in many monogamous species that have the opportunity to observe courtship among other individuals in the population. We predict that complex decision making and the social inheritance of mate preferences will be discovered in many other animals, including many socially monogamous species.

#### ACKNOWLEDGMENTS

We thank Paul Heideman and Dan Cristol for comments on an earlier draft of this chapter. Funding was provided by the National Science Foundation (IOB-0133795 and EF-0436318) awards to JPS.

#### REFERENCES

- Agrawal, A. A. (2001). The evolutionary consequences of mate copying on male traits. *Behavioral Ecology and Sociobiology*, *51*, 33-40.
- Amlacher, J., & Dugatkin, L. A. (2005). Preference for older over younger models during mate-choice copying in young guppies. *Ethology Ecology and Evolution*, 17, 161-169.
- Andersson, M. (1994). Sexual Selection. Princeton, NJ: Princeton University Press.
- Andersson, M., & Simmons, L. W. (2006). Sexual selection and mate choice. *Trends in Ecology and Evolution*, 21, 296-302.
- Bennett, A. T. D., Cuthill, I. C., Partridge, J. C., & Maier, E. J. (1996). Ultraviolet vision and mate choice in zebra finches. *Nature*, 380, 433-435.

- Birkhead, T. R., Burke, T., Zann, R. A., Hunter, F. M., & Krupa, A. P. (1990). Extrapair paternity and intraspecific brood parasitism in wild zebra finches, *Taeniopygia guttata*, revealed by DNA fingerprinting. *Behavioral Ecology and Sociobiology*, *27*, 315-324.
- Bischof, H.-J., & Rollenhagen, A. (1999). Behavioural and neurophysiological aspects of sexual imprinting in zebra finches. *Behavioural Brain Research*, 98, 267-276.
- Brooks, R. (1998). The importance of mate copying and cultural inheritance of mating preferences. *Trends in Ecology and Evolution*, 13, 45-46.
- Brooks, R. (1999). Mate choice copying in guppies: Females avoid the place where they saw courtship. *Behaviour*, 136, 411-421.
- Brooks, R., & Endler, J. A. (2001). Female guppies agree to differ: Phenotypic and genetic variation in mate-choice behavior and the consequences for sexual selection. *Evolution*, 55, 1644-1655.
- Brown, G. R., & Fawcett, T. W. (2005). Sexual selection: Copycat mating in birds. Current Biology, 15, R626-R628.
- Burley, N. (1988). Wild zebra finches have band-colour preferences. *Animal Behaviour*, 36, 1235-1237.
- Burley, N. T., Parker, P. G., & Lundy, K. (1996). Sexual selection and extrapair fertilization in a socially monogamous passerine, the zebra finch (*Taeniopygia guttata*). *Behavioral Ecology*, 7, 218-226.
- Doucet, S. M., Yezerinac, S. M., & Montgomerie, R. (2004). Do female zebra finches (*Taeniopygia guttata*) copy each other's mate preferences? *Canadian Journal of Zoology*, 82, 1-7.
- Dugatkin, L. A. (1996a). Copying and mate choice. In: Heyes, C. M., & Galef, B. G., Jr. (Eds.), Social Learning in Animals: The Roots of Culture (pp. 85-105). New York, NY: Academic Press.
- Dugatkin, L. A. (1996b). The interface between culturally-based preferences and genetic preferences: female mate choice in *Poecilia reticulata*. *Proceedings of the National Academy of Sciences, USA*, 93, 2770-2773.
- Dugatkin, L. A., & Godin, J. G. J. (1992). Reversal of female mate choice by copying in the guppy (*Poecilia reticulata*). *Proceedings of the Royal Society of London B*, 249, 179-184.
- Dugatkin, L. A., & Godin, J. G. J. (1993). Female mate choice copying in the guppy (*Poecilia reticulata*): age dependent effects. *Behavioral Ecology*, 4, 289-292.
- Freeberg, T. M., Duncan, S. D., Kast, T. L., & Enstrom, D. A. (1999). Cultural influences on female mate choice: an experimental test in cowbirds, *Molothrus ater. Animal Behaviour*, 57, 421-426.
- Galef, B. G., & White, D. J. (2000). Evidence of social effects on mate choice in vertebrates. Behavioural Processes, 51, 167-175.
- Godin, J. G. J., Herdman, E. J. E., & Dugatkin, L. A. (2005). Social influences on female mate choice in the guppy, Poecilia reticulata: generalized and repeatable trait-copying behaviour. *Animal Behaviour*, 69, 999-1005.
- Greenfield, M. D., & Rodriguez, R. L. (2004). Genotype-environment interaction and the reliability of mating signals. *Animal Behaviour*, 68, 1461-1468.
- Kirkpatrick, M., & Dugatkin, L. A. (1994). Sexual selection and the evolutionary effects of copying mate choice. *Behavioral Ecology and Sociobiology*, 34, 443-449.
- Kokko, H., Brooks, R., Jennions, M. D., & Morley, J. (2003). The evolution of mate choice and mating biases. *Proceedings of the Royal Society of London B*, 270, 653-664.

- Kokko, H., Brooks, R., McNamara, J. M., & Houston, A. I. (2002). The sexual selection continuum. *Proceedings of the Royal Society of London B*, 269, 1331-1340.
- Laland, K. N. (1994). Sexual selection with a culturally transmitted mating preference. *Theoretical Population Biology*, 45, 1-15.
- Lynch, K. S., Crews, D., Ryan, M. J., & Wilczynski, W. (2006). Hormonal state influences aspects of female mate choice in the Tungara Frog (Physalaemus pustulosus). *Hormones* and Behavior, 49, 450-457.
- Lynch, K. S., Rand, A. S., Ryan, M. J., & Wilczynski, W. (2005). Plasticity in female mate choice associated with changing reproductive states. *Animal Behaviour*, 69, 689-699.
- Oetting, S., & Bischof, H.-J. (1996). Sexual imprinting in female zebra finches: changes in preferences as an effect of adult experience. *Behaviour*, *133*, 387-397.
- Ophir, A. G., & Galef, B. G., Jr. (2004). Sexual experience can affect use of public information in mate choice. *Animal Behaviour*, 68, 1221-1227.
- Patricelli, G. L., Uy, A. C., Walsh, G., & Borgia, G. (2002). Sexual selection: Male displays adjusted to female's response. *Nature*, 415, 279-280.
- Pruett-Jones, S. (1992). Independent versus nonindependent mate choice do females copy each other. *American Naturalist*, 140, 1000-1009.
- Rodriguez, R. L., & Greenfield, M. D. (2003). Genetic variance and phenotypic plasticity in a component of female mate choice in an ultrasonic moth. *Evolution*, *57*, 1304-1313.
- Schlupp, I., & Ryan, M. J. (1997). Male sailfin mollies (*Poecilia latipinna*) copy the mate choice of other males. *Behavioral Ecology*, 8, 104-107.
- Shuster, S. M., & Wade, M. J. (2003). *Mating Systems and Strategies*. Princeton, NJ: Princeton University Press.
- Stohr, S. (1998). Evolution of mate-choice copying: a dynamic model. *Animal Behaviour*, 55, 893-903.
- Swaddle, J. P. (1996). Reproductive success and symmetry in zebra finches. *Animal Behaviour*, 51, 203-210.
- Swaddle, J. P., Cathey, M. G., Correll, M., & Hodkinson, B. P. (2005). Socially transmitted mate preferences in a monogamous bird: a non-genetic mechanism of sexual selection. *Proceedings of the Royal Society of London B*, 272, 1053-1058.
- Swaddle, J. P., & Cuthill, I. C. (1994a). Female zebra finches prefer males with symmetric chest plumage. *Proceedings of the Royal Society of London B*, 258, 267-271.
- Swaddle, J. P., & Cuthill, I. C. (1994b). Preference for symmetric males by female zebra finches. *Nature*, 367, 165-166.
- ten Cate, C., & Vos, D. R. (1999). Sexual imprinting and evolutionary processes in birds: a reassessment. *Advances in the Study of Behavior*, 28, 1-31.
- Uehara, T., Yokomizo, H., & Iwasa, Y. (2005). Mate-choice copying as Bayesian decision making. *American Naturalist*, 165, 403-410.
- Waas, J. R., & Wordsworth, A. F. (1999). Female zebra finches prefer symmetrically banded males, but only during interactive mate choice tests. *Animal Behaviour*, 57, 1113-1119.
- Welch, A. M. (2003). Genetic benefits of a female mating preference in gray tree frogs are context-dependent. *Evolution*, 57, 883-893.
- Westneat, D. F., Walters, A., McCarthy, T. M., Hatch, M. I., & Hein, W. K. (2000). Alternative mechanisms of nonindependent mate choice. *Animal Behaviour*, 59, 467-476.
- White, D. J., & Galef, B. G., Jr. (1999). Mate choice copying and conspecific cueing in Japanese quail, *Coturnix coturnix japonica*. Animal Behaviour, 57, 465-473.

- White, D. J., & Galef, B. G., Jr. (2000). 'Culture' in quail: social influences on mate choices of female *Coturnix japonica*. *Animal Behaviour*, *59*, 975-979.
- Witte, K., & Massmann, R. (2003). Female saiflin mollies, *Poecilia latipinna*, remember males and copy the choice of others after 1 day. *Animal Behaviour*, 65, 1151-1159.
- Zann, R. A. (1996). *The Zebra Finch: A Synthesis of Field and Laboratory Studies*. Oxford, UK: Oxford University Press.

Chapter 29

# **RELATIVITY OF FINANCIAL PREFERENCES: HOW CHOICE OPTIONS INFLUENCE INVESTMENT DECISION MAKING<sup>\*</sup>**

# *Ivo Vlaev*<sup>\*</sup> *and Nick Chater*

Department of Psychology, University College London 26 Bedford Way, London, WC1H 0AP United Kingdom

# ABSTRACT

We report a study in which methodologies from psychophysics are adapted to investigate context effects on financial decision making related to retirement savings and risky investment. The aim was to determine how the range of the options offered as possible saving rates and levels of investment risk influences decisions about these variables. The respondents were presented with either a full range of choice options or a limited subset of the feasible options. The study was conducted on a sample of working people, and we controlled whether the participants can financially afford in their real life the decisions taken in the test. The results showed that choices of saving and risk are affected by the position of each option in the range of presented options. Various measures of risk aversion did not account for the risk taken in each condition. Only the simplest and most direct risk preference measure was a significant predictor of the responses within all contexts (conditions), although the actual choices were still very much influenced by each context. Thus, the results reported here suggest that judgments and choices are relative, rather than absolute, which corroborates, in a more applied and realistic setting, previous related work with abstract gambles and hypothetical risky investments.

Keywords: context effects, decision making, judgment, investment risk, saving.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Economics, Business and Finance*, edited by Klaus P. Hofmann, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>\*</sup> Correspondence concerning this article should be addressed to Ivo Vlaev, Department of Psychology, University College London, London, WC1H 0AP, United Kingdom. E- mail: i.vlaev@ucl.ac.uk.

This article presents a study that investigated how far, simple, and practically relevant, modifications in the decision making context can affect the way people make financial decisions related to retirement savings and investment. Specifically, we aim to see to what extent a new and powerful context effect, prospect relativity (Stewart, Chater, Stott, and Reimers, 2003), is likely to be important in practical settings. The applied objective of the experiment presented here is to test whether contextual factors can be used to stimulate financial consumers to save more for retirement and, also, to encourage them to take higher investment risks in relation to their retirement savings products in order to obtain the longterm benefits from higher expected returns in the stock market. The relevance of this objective is particularly significant given the major demographic shift that is occurring across the industrialized and post-industrialized world. Rising life-expectancy implies that growing, aging population will live longer after retirement and they would need substantial income to support their lifestyles. Without growing younger working population, our aging society requires working people to save more for retirement and also to opt for bigger returns on their investments by utilising the opportunities offered by the capital markets. The interests of individuals, the financial services industry, and governments, are all likely to be served both by a higher overall level of saving and by increased uptake of long-term investments in the capital markets (rather than risk-free fixed-interest products). Therefore, such an applied objective is in the consumer's interest and governments around the world are also encouraging such an initiative, because previous research suggests that people save too little and do not take enough financial risk (e.g., for analysis of this issue in UK see Oliver, Wyman and Company, 2001).

In this article, we present a study that investigates the effects of the framing and presentation of financial information when asking people to express their preferences in relation to different retirement savings and investment scenarios. The experimental design and method are based on the recent discovery of a substantial dependence of human preferences on the set of options they are presented with. This phenomenon was termed *prospect relativity* and indicates a lack of stable underlying preference function (Stewart, Chater, Stott, and Reimers, 2003). This finding, which we describe in more detail below, is a striking illustration of the view that preferences are constructed, on the fly, rather being a stable basis for decision making (Slovic, 1995).

A theoretically important question is how far these effects transfer from abstract lowstakes gambles, to major financial decisions that could significantly affect long-term wellbeing. If we observe such transfer to realistic financial decisions, then this would have crucial practical significance for marketing, sales, and provision of advice in the financial services industry. Our goal was to investigate how far individual variation in risk preferences is stable across different realistic decision contexts.

# **THEORETICAL BACKGROUND**

Various forms of context dependence of human decision making have been studied extensively in the past. These effects are inconsistent with the normative theory of choice in economics, and therefore, they can be viewed as irrational (although see Sher and McKenzie, 2006, for discussion of a sense in which many context effects may have a rational basis).

Traditionally, psychologists and economists have tried to empirically assess the assumptions of the standard theory of choice in economics, revealing mounting evidence that human behaviour diverges from the predictions of the theory, especially in the context of decision making under risk (e.g., Kagel and Roth, 1995; Kahneman and Tversky, 2000; and also Camerer, 1995, for a review). Thus, a large literature in the psychology of decision making has demonstrated that expected utility theory does not fit with observed behaviour in a large variety of ways (e.g., Kahneman, Slovic, and Tversky, 1982). Some of these disparities between observed behaviour and the normative strictures may be due to lack of familiarity, practice or understanding of the problem in hand; however, behaviour that appears inconsistent with rational choice theory is very difficult to eliminate (e.g., Shafir and LeBoeuf, 2002). Thus, there has been considerable interest, in both psychology and economics, in building models that provide more accurate descriptions of observed decision making behaviour. These psychological theories have typically involved variations of the expected utility, to bring its predictions in line with people's observed behavior (e.g., prospect theory [Kahneman and Tversky, 1979], regret theory [Loomes and Sugden, 1982], rankdependent utility theory [Quiggin, 1982]). The result involves building a bridge from expected utility theory (the normative theory) to psychology (the descriptive theory), where the core elements of the normative approach are maintained.

Rather than starting from a normative economic theory, and attempting to make modifications that render it descriptively acceptable, an alternative research strategy is to start from assumptions about elementary cognitive processes, and attempt to construct an account that can address the economic problem of choice. That is, our attempt is to bridge from psychology to economics (instead from economics to psychology as it used to be the case) with the ultimate aim of showing that the resulting model predicts decision-making behaviour (see, e.g., Gigerenzer and Todd, 1999 and Payne, Bettman and Johnson, 1992, for examples of this strategy).

In search of such realistic psychological foundations for descriptive decision theory, Stewart et al. (2003) considered whether context effects observed in psychophysics might transfer to risky decision making. Specifically, Stewart et al. (2003) found that the set of options from which an option was selected almost completely determined the choice. They demonstrated this effect in selection of a risky prospect and in a certainty equivalent estimation task (the amount of money for certain that is worth the same to the person as a single chance to play the prospect). Similarly, the selection of a preferred option from a set of prospects was strongly influenced by the prospects available. More recently, Stewart, Chater, and Brown (2006) developed a model of risky choice, *decision by sampling*, which assumes relative judgments and provides one explanation for the results in Stewart et al. (2003).

The cognitive claim that Stewart et al. (2003) argue for, is that people are unable to represent absolute magnitudes, whether psychophysical or abstract (including a choice option's attributes like utilities, payoffs, and probabilities). That is, when people represent a magnitude, they can only do so on the basis of whether it is larger or smaller than other magnitudes retrieved from memory or observed in the environment. If people cannot represent the absolute value of magnitudes on any cardinal scale, and the subjectively judged utility of an option is determined by its relationship to comparison options, then judgments will be strongly affected, or even determined, by the context.

The core evidence for this claim comes from the study of the perception of the intensity of basic psychophysical magnitudes such as the brightness of a light or the loudness of a sound. Much traditional research in psychophysics has assumed the existence of some cardinal internal scale of intensities, onto which physical stimulation must somehow be mapped; and there has been consequent debate concerning the nature of this mapping (e.g., whether it is logarithmic, as argued by Fechner, 1966; or a power law, as argued by Stevens, 1957). But more recent theory (reviewed and analysed in Laming, 1997) suggests a different point of view---that the very idea of an internal scale is incoherent. In particular, Laming (1997) has shown that empirical data in line with Stevens' power law (relating psychophysical variables and free numerical judgments) can arise without assuming any representation of absolute information. In addition, Stewart, Brown, and Chater (2005) demonstrated that a wide range of data from the psychophysical task of absolute magnitude identification can be captured by a model which has no absolute scales, and relies entirely on local comparisons between recent stimuli.

A particularly telling example supporting this viewpoint is an elegant experiment conducted by Garner (1954), who asked participants to judge whether tones were more or less than half as loud as a 90 dB reference loudness. Participants' judgments were entirely determined by the range of tones played to them. There were three groups of participants who received tones in three different ranges respectively. Participants who heard tones in the range 55-65 dB had a half-loudness point (i.e., where their judgments were "more than half as loud" 50% of the time and "less than half as loud" 50% of the time), of about 60 dB. Another group, who received tones in the range 65-75 dB had a half-loudness point of about 70 dB. A final group, who heard tones in the range 75-85 dB, had a half-loudness point of about 80 dB. Garner's experiment indicates, therefore, that people have no idea of the absolute intensity of the sound or what it means for one sound to be half as intense as another. Instead, it seems that people adjust their responses depending on the presented sound intensities from which they are asked to choose.

Other examples of similar context effects abound in psychophysics. Thus, empirical investigations in absolute identification (e.g., Garner, 1953; Holland and Lockhead, 1968; Lockhead, 1984; Luce, Nosofsky, Green, and Smith, 1982; Ward and Lockhead, 1970), magnitude estimation (e.g., Jesteadt, Luce, and Green, 1977), relative intensity judgment (Lockhead and King, 1983), and matching tasks (Stevens, 1975), have shown that perceptual judgments of stimuli varying along a single psychological continuum are strongly influenced by the preceding material. In summary, context effects, like those found by Garner (1954), are consistent with participants making perceptual judgments on the basis of relative magnitude information, rather than absolute magnitude information (see also Laming, 1984, 1997).

Here we have digressed briefly into psychophysics, in order to make it explicit how we apply the resulting conclusion to a decision-making context. Note, though, that the parallel between the two domains, psychophysics and economic decision making, is relatively close. After all, just as perceptual theorists traditionally assumed that people have internal scales for the representation of loudness and brightness, so a traditional psychological or economic picture of an agent assumes that the agent must have internal scales for the representation of the utility of various outcomes; for representing, perhaps distorted (e.g., Kahneman and Tvesky, 1979) of the probability that they will occur; and so on. Without some type of scale for utility or probability, the model of the economic actor (or any decision-making agent in general) would look very different.

One prediction based on such model is that the attributes of the previously or currently seen risky prospects influence the decisions in the current prospect. Stewart et al. (2003)

argued for the existence of what they call "prospect relativity": that the perceived value of a risky prospect (e.g., "p chance of x") is relative to other prospects with which it is presented. Note that Stewart et al. studied peoples' perception of utilities in individual decision making tasks in gambling situations. The prediction, based on the psychophysical studies described above, is that the option set (i.e., the context) will affect peoples' choices because there is no fixed internal scale according to which people make their judgements of the values of certain options. Recall that the results demonstrated a powerful context effect in judging the value of different risky prospects - the set of options offered as potential certainty equivalents for simple prospects was shown to have a large effect on the certainty equivalents selected. To illustrate this result, for example, when during judging the value of a 50% chance of winning  $\pounds 200$  people have options of  $\pounds 40$ ,  $\pounds 50$ ,  $\pounds 60$ , and  $\pounds 70$ , the most popular choice is  $\pounds 60$  and then second choice is £50. When people have options of £90, £100, £110, £120 pounds, the most popular choice is  $\pm 100$ , and then second choice is  $\pm 110$ . So the set of alternatives affected valuation by a factor of (almost) 2. This effect was replicated despite monetary incentives designed to encourage participants to deliver accurate and truthful certainty equivalents. In another experiment, the set from which a simple prospect was selected was also shown to have a large effect on the prospect that was chosen. Vlaev and Chater (2006) discovered similar results in a very different context, where people play the strategic games based on Prisoner's Dilemma, indicating the generality of this effect. Finally, Vlaev, Chater, and Stewart (2007) report three studies, in which methodologies from psychophysics (which were similar to the methods used by Stewart et al., 2003) were adapted to investigate context effects on individual financial decision making under risk. The aim was to determine how the range and the rank of the options offered as saving amounts and levels of investment risk influence people's decisions about these variables. In the range manipulation, participants were presented with either a full range of choice options or a limited subset, while in the rank manipulation they were presented with a skewed set of feasible options. The results showed that choices are affected by the position of each option in the range and the rank of presented options, which suggests that judgments and choices are relative.

Effects of the type presented above suggest that people's expressed (or revealed) risk preferences are not absolute, but are, to some degree at least, relative to the range of available options (see Stewart, Chater, and Brown, 2006, for model of risky choice that assumes relative judgments only). A plausible account of the context effects caused by the range of options is the range-frequency theory proposed by Parducci (1965, 1995). Parducci found that the neutral point of the judgment scale did not correspond to the mean of the contextual events, contrary to popular at that time adaptation level theory (Helson, 1964), but rather to a compromise between the midpoint (defined by the range) and median (depending on the skew) of the distribution of contextual events. For example, satisfaction judgements are different between two distributions of life events, which have different skew and identical means. The *range principle* reflects tendency to judge an event relative to its position within the range of stimuli on the specified dimension of judgment, while the *frequency principle* reflects a tendency to judge an event relative to its rank within the immediate context. The subjective value given to an attribute is a function of its position within the overall range of attributes, and its rank. Thus, this model implies that attributes are judged purely in relation to one another and their subjective value is independent of their absolute value. Rangefrequency theory has already been used to account for context effects in decision making under risk. Birnbaum (1992), Stewart et al. (2003), Vlaev and Chater (2006), and Vlaev, Chater, and Stewart (2007) found their data to be consistent with the theory.

In this article, we present a study that aimed to test the practical relevance of the prospect relativity principle to more realistic financial decision making scenarios. The results we present here suggest that, when people make financial decisions, the attractiveness of the choice options significantly depends on the other available options.

# PROSPECT RELATIVITY PRINCIPLE AND REALISTIC FINANCIAL DECISION SCENARIOS

The goal of this experimental study was to make a provisional estimate of the degree to which the kinds of effects that are revealed by Stewart et al. (2003) could also be applicable to real financial decisions. Our study builds on, and is similar to, the study by Vlaev, Chater, and Stewart (2007). The domain used here, and also by Vlaev, Chater, and Stewart, was saving and investment for retirement, because it is an issue having serious social relevance at the moment. We currently live in a financial environment, in which aging population and younger consumers are increasingly expected to take command of their own pension and investment decisions. Therefore, the following two key issues arise: How does the range of options people choose between affect the level of pension investment they choose? How does the range of options from which people choose affect the level of risk they accept with that investment?

This experiment followed logic similar to the decision experiments reported by Stewart et al. (2003) and Vlaev, Chater, and Stewart (2007), which were in turn inspired by Garner's (1954) loudness judgment experiment. In various questions, the participants in our study were asked to select among a predefined set of values related to five variables: (a) the desired percentage of the annual income that will be saved for retirement, (b) the investment risk expressed as the percentage of the saving that will be invested in risky assets, (c) retirement age, (d) expected retirement income, and (e) possible variability of the retirement income.

There was a control condition called a *free choice* condition, in which the participants had to freely decide the value of each one of these variables selecting from the full range of options. In two *context* conditions, participants were asked to select these values from subranges of the set of options offered by the experimenter in the free choice condition. Thus, there were three between-participant conditions in the experiment presented here, i.e., with separate groups for the free choice, low range, and high range conditions. In the free choice condition, all options were presented. In the two other conditions, the choice of prospects was limited to either the first or second half of the prospects available in the free choice condition, so that the participant in the high range condition were presented with a range of values the lowest of which coincides with the highest option in the low context condition. In the free choice condition for saving, the options were presented in monetary terms and varied from 2% to 22% of the hypothetical salary (£25,000) increasing with 2% between the options; so there were eleven options to choose among, while the low range condition spanned from 2% to 12% and the high range condition was from 12% to 22%. The same design was applied for the other risk variable. The choice option values in the free choice condition for investment risk varied from 0% to 100% and were increasing with 10% between the options. For retirement age the values varied from 48 to 68 increasing with 2 years. Table 1 presents the values for savings, risk, and retirement age, in the three conditions. For the retirement income and its variability, the values were different for every question depending on the combination of saved amount, investment risk, and retirement age.

Free Choice			Low Range			High Range		
Save	Risk	Retire	Save	Risk	Retire	Save	Risk	Retire
500	0	48	500	0	48			
1,000	10	50	1,000	10	50			
1,500	20	52	1,500	20	52			
2,000	30	54	2,000	30	54			
2,500	40	56	2,500	40	56			
3,000	50	58	3,000	50	58	3,000	50	58
3,500	60	60				3,500	60	60
4,000	70	62				4,000	70	62
4,500	80	64				4,500	80	64
5,000	90	66				5,000	90	66
5,500	100	68				5,500	100	68

Table 1. Figures for saved amount (£), investment risk (%), and retirement age in thethree conditions of the experiment

Note that for risk, it is natural to assume that people anchor the possible percentage of their retirement savings that can be invested in risky assets to be between on 0% and 100% (i.e., people cannot have negative savings, or invest more than 100% of their savings). For the range of possible savings values, it is equally easy to imagine that the lower bound is fixed at 0%, while the upper bound depends on many factors like for example legal requirements, cost of living, etc. We fixed the savings range across all conditions at values 0% and 22% (the range for the full range condition) because 22% approximates the upper bound for the retirement savings rate in UK due to legal and tax restrictions. The team of professional actuaries who monitored our research project suggested 22% as the practically relevant upper bound.

Our prediction was that if a participant is not influenced by the set of options, then his or her choice of each value in the high and low range conditions should be independent of the other values in the set and the chosen values should be the nearest to his or her free choice. The first prediction was that if people are influenced by the context (i.e., the other available options), then the mean saving and risk selected in the both high range and low range conditions should be different from the free choice condition. This is because we assume that if people's true preferences are represented by the results in the free choice condition, and their choices are not influenced by the context in the high or low range condition. Thus, for example, if peoples' preferences are naturally amongst the lower options in the full range, then there should not be any significant difference between the results in the low range and the full range; the only exception should be the highest value in the low range, which ought, if there are no context effects, to include all the people who would choose that option, or all higher options, in the full range condition. Conversely, if all participants that truly prefer options in between the higher options in the full range, then there should be no significant difference between the average results in the high range and the full range; the only exception should be the lowest value of the high range, which should, in the absence of context effects, include all those people who would choose that options or a lower one, in the full range. However, if *both* the low range and the high range conditions are significantly different from the full range condition, in terms of the low range being lower than the full range *and* the high range being higher than the full range, then we can conclude that the context had significantly influenced the choices in the sub-range conditions (i.e., this result should be due to the effects of the choice set in the high range and low range conditions). In other words, the choice set is affecting the responses and inducing people to select higher/lower options than they would have selected if they were in the full range condition.

We also conducted the following statistical comparison between the high range and low range conditions in order to test furthermore whether the context (range of options) had a significant effect on the choices. We compared the lowest option in the high range, against the sum of the all the items in the low range, except the highest item (i.e., every option in the low range, which was missing in the high range). In other words, we compared whether the lowest option in the high range condition was significantly lower than the proportion of times the options below it were selected in the low range condition. Conversely, we compared the highest option in the low range, against the sum of the options in the high range, except the lowest option. In either direction (and we tested both directions), if the latter is bigger, we have prospect relativity effect (or at least, a rational choice model will fail to predict this result) and we can conclude that this result should be due to the effects of the choice set in the high range and low range condition respectively. The logic behind this analysis is that people who do not select the highest item in the low range condition should definitely select the lowest item in the high range condition. So the former ought to be less numerous than the latter. This is because their true preferences should be within the options lower than the highest item in the low range condition. In other words, the reason we are missing the highest item in the low/high range, is that we do not know what proportion of the people selecting this option want to select options above/below it, but they do not have the opportunity in the low/high range. Alternatively, if participants' responses are solely determined by the set of options presented to them, then the distribution of responses across options should be identical for both the low and the high range conditions.

In addition, the design of the experiment presented in this article had four new features relative to the work reported by Vlaev, Chater, and Stewart (2007). We designed these new characteristics in order to increase our study's relevance to real-world financial advice. These four new design features were:

#### 1) Representative Sample

The study was conducted on a sample of working people, rather than university students. In other words, we used sample of participants who are more realistic to be consumers of financial advice (e.g., people who are already working, have a family, and need to save for retirement pension provision) than the student population used in Vlaev, Chater, and Stewart (2007).

## 2) Realistic Financial Assumptions

The future financial outcomes (e.g., expected annuity values) were calculated using very plausible financial assumptions (like inflation, risk free rate, risk premium rate, etc.). We undertook this work on consumer understanding of risk, both from a mathematical and from a psychological standpoint, with the help of the Actuarial Profession's Personal Financial Planning Committee in United Kingdom (who actively participated in creating the test materials and the descriptions of the risky assets). The age of the participants was also taken into account in calculating the time horizon of future returns.

## 3) Financial Affordability Questionnaire

In the light of the important discussion raised at our meetings with professional actuaries and personal financial advisers, we took account of financial affordability, as a constraint on people's choice of pension. We created a financial affordability test, which categorised people according to their individual financial circumstances. The financial affordability test was designed to help and encourage the participants to think through the practical viability of the financial options that they choose. An additional purpose of the financial affordability test was to make the experimental situation appear as a very realistic example of a financial advisory process. This was achieved by asking the participants concrete questions about their real life financial circumstances and problems. Thus, by explicitly focusing respondents' attention on their real life struggles at the beginning of the experimental session, we expected them to provide more adequate and valid responses to our saving and investment questions.

The financial affordability questionnaire is presented in Appendix A and it has the following main features:

- a) Question 2 asks people to estimate to what extent their current income (question 1) is sufficient to cover their expenditures, and then to judge in percentage term whether, and by how much, this income is sufficient or insufficient to cover these expenses (e.g., "I would be happy to earn around 20% more than my current salary"). In addition, the participants were asked to indicate how much they are able to save at the moment (question 3).
- b) In question 4, the participants were provided with a list of various types of spending and they had to answer how much of their current income is spend on each of these expenditures. In general, there were two types of expenditure examples discretionary (e.g., leisure activities) and essential ones (e.g., food and rent) and we asked the respondents to give estimates of their expenditure across these categories.
- c) Question 6 asks people whether they can give up some of their discretionary spending in order to increase their savings. Here the focus again was on the amount and type of current saving and discretionary spending; and hence the degree to which people can readily reallocate money towards a pension. Here we also aimed to test how important and essential some of these discretionary expenditures are (e.g., some people might be unwilling to give up some types of social life, hobbies, sport activities, etc.). Participants were also informed that at the and of the experimental session if their average preferred savings rate is above their current savings as

indicated in question 3, then they had to readjust some of their expenditures in question 4 so that to be able to provide the additional capital that is the lacking difference between their real savings and the saving levels selected in the second part of the experiment (in addition, question 5 separately asked what is the maximum amount that they would consider saving each year).

#### 4) Risk Preference Tests

We collected information about participants' risk attitudes in order to investigate to what extent their decisions were influenced by their general risk preferences. Thus, we expected to test whether people can be manipulated to take more risk than their genuine risk preferences are, or whether they intuitively know how much investment risk to take. Such results could also inform us whether it is worth trying to stimulate people to invest in a way that best matches their risk and time preferences (for example, to increase their investment risk exposure if they are particularly risk seeking), or, if people's decisions are easily manipulated by the context, whether to offer them financial products only on the bases of their individual goals and social and financial circumstances (e.g., to offer them relatively risky investments that will accomplish the desired retirement income, independently of their risk preferences). We used five different measures of risk aversion (presented in Appendix C) in order to measure whether the choices of investment risk in the high and low range conditions are due to natural risk preferences instead of context effects. These measures represented typical selfreport hypothetical measures (as used in the literature) in the form of simple direct questions and hypothetical gambles

Questions 1-4 (Direct Risk, Direct Concern, Relative Risk, Relative Concern). These questions are rather simple and direct measures. We used these both as a base-line, and also because of existing results showing that simple self-report measures of risk preferences could be more powerful predictors of portfolio allocation than sophisticated measures based on economic theory (Kapteyn and Teppa, 2002). Two of these questions measured risk attitudes with the basic questions "How much risk are you prepared to take?" (Direct Risk) or "How much are you concerned about your financial future?" (Direct Concern) and the participants had to answer on a scale from 1 (not at all) to 5 (very much) to what extend they agree with these statements. There were also two questions about how people perceive their level of risk aversion in relation to other people – "Are you more or less willing to take risks than the average person?" (Relative Risk) or "Are you more or less concerned about your financial future than the average person?" (Relative Concern) and the participants had to answer on the following scale: 1 - much less, 2 - less, 3 - the same as the average, <math>4 - more, and 5 - much more.

*Question 5 (Income Gamble).* Question 5 is a well-known test by Barsky, Juster, Kimball, and Shapiro (1997), who constructed a measure of risk aversion by asking respondents about their willingness to gamble on lifetime income. By contrast, experiments in the existing literature ask people to gamble over spending or consumption and typically involve stakes that have little impact on lifetime resources. However, a gamble whose outcome is too small to be meaningfully related to consumption should not require a risk premium, on normative grounds, and therefore such gamble is not a good measure of economic risk preference. So the principal requirement for a question aimed at measuring risk aversion according to Barsky

et al. is that it must involve gambles over lifetime income. In addition, after pre-testing, Barsky et al. concluded that survey respondents would better understand income than consumption lotteries. The three questions in this test, in the first paragraph and then in (a) and (b), separate the respondents into four distinct risk preference categories, depending on the combinations of their answers (see Question 5 in Appendix C): (1) reject the risk to cut the (family) income by one-third in the first question and also reject the risk in (b) to cut the income by one-fifth (20%); (2) reject the risk for one-third income cut in the first question but accept the possibility for one-fifth cut in (b); (3) accept the possibility for one-third income cut in the first question but reject the one-half cut risk in (a); and (4) accept both possibilities for one-third income cut in the first question and one-half cut in (a). These four categories can be ranked by the level of risk-seeking without having to assume a particular functional form for the utility function. Barsky et al. (1997) provide four numerical indices of relative increasing risk-seeking corresponding to each category respectively: 0.11, 0.36, 0.68, and 1.61. In the original study by Barsky et al., their measure was significantly correlated with various demographic factors, and it was positively related to risky behaviors, including smoking, drinking, failing to have insurance, and holding stocks rather than treasury bills.

## Метнор

Participants. We sent the materials (the financial affordability questionnaire and the savings and investment questionnaire) by post to a population of working individuals, as this population is typically not able to attend laboratory sessions. We sent out the questionnaire to 600 people, and received 64 completed questionnaires. These respondents were typical of the demographics in the geographical area and selected from a big subject pool of people who expressed desire to participate. Participants who completed their questionnaires were paid £10 for their participation (received as a check after they have returned the answer sheet). There were 24 men with average age 36.5 and 40 women with average age 37. The Low Range Condition had 20 participants: 7 men (av. age 37) and 13 women (av. age 38); the Free Choice Condition had 21 participants: 9 men (av. age 33) and 12 women (av. age 36); and the High Range Condition had 23 participants: 8 men (av. age 40) and 15 women (av. age 36).

Design. The questions in the prospect relativity test were formulated as long-term saving/investment decision tasks related to retirement income provision. The participants had to make decisions about five key variables. These variables were the saved proportion of the current income, the risk of the investment expressed as the proportion invested in risky assets,<sup>1</sup> the retirement age, the desired income after retirement, and the preferred variability of this income (we explained that such variability is due to favourable and respectively unfavourable economics conditions).

The experimental materials were designed as 10 independent hypothetical questions, in which we varied each of the five key variables. However five of the questions focused only

<sup>&</sup>lt;sup>1</sup> There are, of course, various types of risky assets, including a wide variety of bonds and equities; but in reality these various investment vehicles differ mainly in their risk-return characteristics. Therefore, we simply described the characteristics of these two assets – the High Risk Asset and the Low Risk Asset, rather than labelling them explicitly as bonds and equities, although while setting the basis so that it is not out of line with typical assumptions made about actual assets. This aimed to avoid some of the potential challenges that might otherwise result and which could draw attention away from the results.

on savings while the other five questions focused on risk, and some questions showed how changing savings or risk would affect another variable or set of variables. For example, how changing the investment risk can affect the projected retirement income and its variability - with higher risk offering higher expected income on average, but also wider spread of the possible values.<sup>2</sup>

As an example, Figure 1 presents a question in the free choice condition, in which the participants were asked to choose their preferred level of investment risk by selecting one of the rows in the table (note that in this format the key choice variable is in the first column of the table while the other columns are showing the effects on the other variables like the minimum, average, and maximum retirement income<sup>3</sup>). This future distribution (risk) of the investment in risky assets is calculated as follows. Assuming a variable annual interest rate with mean  $\mu$  and standard deviation  $\sigma$ , the expected return on an *n*-year investment is also log normally distributed with mean  $\mu^n$  and standard deviation  $\sigma = \mu^{2n}(((\sigma^2/\mu^2)+1)^{n-1}))$ . We also assumed that an annuity that provides 1/14th of the lump sum saved each year is purchased, which is a typical figure used in the UK financial services industry.

In Appendix B, there is a detailed description of each question and its purpose (the questions are grouped by the key variable that participants are asked to select – savings or investment risk). The ten questions were presented in different order in the various conditions. We also counterbalanced the order of saving and risk questions by dividing the participants into two groups: one that first answered the saving questions and then the risk questions, and second group that answered the risk questions before the saving ones.

Note that since we used mature population of participants who vary in age, this might create problems with using the same test materials for all participants, because older people would have so save for fewer years (compared with younger people) in order to get the same retirement income. Thus for example if the materials (saving and risk choice options) are created for people with average age of around 25 yrs, and if we give the same materials to somebody who is 50, then of course the older respondent would be willing to save the highest possible amounts for the remaining 10-15 years until retirement, while a younger person has to commit to this higher saving (and lower consumption) rate for 25-30 years. In order to avoid this problem, we decided to create test materials for three different age groups, namely, 30, 40, and 50. We sent the identical test materials to respondents who are plus or minus 5

<sup>&</sup>lt;sup>2</sup> In order to derive plausible figures for the various economic variables we implemented a simple econometric model into a spreadsheets Monte Carlo simulator that calculates the likely impact of changes in each variable on the other four variables. For example, this model can derive what retirement income can be expected from certain savings, investment risk, and retirement age, or what are the possible potential investment options that could lead to the preferred retirement income. The sort of basis the professional actuaries suggested was 2.5% for Inflation, 1.5% real return on Low Risk asset, 4.5% real return on High Risk asset, and 15% annual volatility. Note also that all figures are in pounds and the participants knew this. It is important to stress that all figures shown were in today's money terms (i.e. after taking out the effects of inflation). This is important when comparing figures for different retirement ages.

<sup>&</sup>lt;sup>3</sup> Most of the questions showed the expected retirement income and its variability like in the example above. The possible variability of the retirement income was explained by referring to the 95% and respectively 5% confidence intervals of the income variability, i.e. maximum and minimum possible values of the income, for which there is 5% chance to be more than the higher or less than the lower value respectively. On each row of the table these two values were placed on the both sides of the average expected retirement income. The confidence intervals were expressed also in verbal terms using the words very likely. For example, the participants were informed that it is very likely (95% chance) that their income will be below the higher value and above the lower value, and that these two values change depending on the proportion of the investment in equities.

years around each age group (e.g., the financial options calculated for somebody who is 30, were also sent to all respondents between 25 and 35 yrs old). Thus the projected retirement income was calculated for three time horizons: after 35, 25, and 15 years of investment respectively.

Assume that you will retire at 65 and decided to save 11 percent of your current salary (£2750) in order to provide for your retirement income. The following options offer different ranges of retirement income (in pounds) depending on the percentage of your savings allocated to shares (in the stock market) and you can see the effects on the expected average retirement income and its variability (minimum and maximum). Note that you are very likely (have 95% chance) to be between the minimum and maximum figures indicated in the table below. Please select one of the following options.

Invest	Minimum	Average	Maximum
0 %	16,000	16,000	16,000
10 %	17,000	19,000	22,000
20 %	17,000	21,000	23,000
30 %	17,000	23,000	29,000
40 %	16,000	26,000	35,000
50 %	15,000	29,000	42,000
60 %	14,000	33,000	51,000
70 %	11,000	37,000	62,000
80 %	7,000	41,000	76,000
90 %	2,000	47,000	92,000
100 %	0	53,000	112,000

Figure 1. A question in the free choice condition, in which the participants were asked to choose their preferred level of investment risk by selecting one of the rows in the table bellow. In this format the key variable is in the first column of the table bellow while the other columns are showing the effects on the other variables like the minimum, average, and maximum retirement income.

The high range condition was derived by deleting the lower five rows of the table for each question in the control condition and the low range condition was derived by deleting the higher five rows in the tables in the free choice condition (i.e., the same was done for each question). Therefore, in the free choice condition, the participants had to choose among eleven possible answer options for each questions while in the high and low range conditions there were only six available answer options. Note that in this design the participant had to choose among predetermined option values in all conditions. This design was similar to the design used in Experiment 4 reported by Stewart et al. (2003), where in the free choice condition the participants had to choose among predetermined choice among predetermined choice among predetermined choice among predetermined choice options that were either the higher halve or the lower half of the list of options offered in the free choice condition. Vlaev, Chater, and Stewart (2007) also used very similar design.

*Procedure.* Participants were sent a booklet containing the financial affordability questionnaire, the ten saving and risk questions, and the five questions measuring risk aversion. They received written instruction explaining that the purpose of the experiment is to answer series of questions about savings and investment related to retirement income provision, and that there were no right and wrong answers and that they are free to choose

whatever most suits their preferences. It was explained that the choice options are predetermined because these are the outcomes that can be realistically accomplished according to a standard economic model and that the task is to choose the option that is nearest to the participant's preferences. The participants were also informed that if they find them unsatisfactory then they can indicate values outside these ranges.

The questions and the answer options were presented as in the example question in Figure 1, which showed the projected retirement income for 30 years old age group (i.e., after 35 years of investment). The participants chose one of the figures in the first column of the table (which were either savings or investment risk values) and they were provided with a separate answer sheet on which to write their answers. Participants were informed that their answers do not need to be consistent between the questions, and that they can freely change their preferences on each question and choose different savings and risk values.

Another issue that we had to deal with was how to account for people's existing savings because we wanted to make our session as realistic as possible. If we give our questions to somebody who already has got a good pension scheme, then she might choose very low saving amounts and investment risk just because she does not need to save much more. On other side, if we tell them to imagine that our scheme is offering them to start anew, then our calculation will have to include also their accumulated savings up to date. This would also require some sophisticated software to be used online with every individual (and which is probably used by the real financial advisors). Our solution to this problem was to write in the instruction that most people in UK are underprovided and that we research what kind of pension top-up product people might find attractive (in addition to the social security scheme), and therefore this is an extra to what they already have.

## RESULTS

Financial Affordability Questionnaire. Table 2 presents the results from the Financial Affordability Questionnaire. We checked whether the participants were in a position to afford the saving levels selected in the experiment. None of the participants had selected inappropriate saving rates in relation to their income and expenditures. Only one person decided to give up half of her essential spending and also discretionary spending in order to provide the additional capital that is required to cover the difference between her real savings rate and the maximum savings rate that she had indicated in the main test. Otherwise, 54% of the participants indicated that they can give up some discretionary spending in order to increase their current savings rate if it is bellow their preferred maximum amount (indicated in question 5). This result indicates that all respondents took their task seriously and carefully selected the saving options in the main test so that their choices reflected their real financial circumstances. We also hoped that asking people to provide this financial information would encourage them to consider their saving decisions carefully and as a result to give answers that are reasonably close to what they would choose for real (recall that the purpose of this questionnaire was mainly to prompt the respondents to give realistic answers to the prospect relativity test).

Question		Mean	Std Deviation	Category		
Annual income		£19,235.5	15,492.6			
Spend less than you ear	n by	£3,601.1	3,032.7			
Spend exactly the amount you earn		£18,658.8	8,729.8	General		
Spend more than you ea	rn by	£1,593.8	1,136.4	—		
Current Saving		£1,829.3	2,363.2			
Food		£2,147.3	1,905.0			
Rent / Mortgage		£3,177.0	1,990.6			
Utilities (electricity, gas	, water, etc.)	£659.9	832.4	—		
Car		£1,301.5	1,616.8	—		
Other transport (train, b	usses)	£343.6	686.3			
Debt repayment		£1,021.2	1,234.8	Essential		
Communications (telepl	none, etc.)	£424.1	300.5	expenditure		
Childcare and Schooling		£283.5	747.0			
Health	-	£76.3	109.7			
Repairs and Maintenand	e	£439.6	542.6			
Other (like health and li		£425.0	479.6			
Holiday	. ,	£327.0	392.3			
Entertainment (e.g., cinema)		£234.0	356.5			
Sport		£177.5	222.1	Discretionary		
Hobbies		£554.7	516.4	expenditure		
Meals and Drinks		£663.7	1,351.6			
Other		£1,1249.0	5,816.9			
Total Expenditure		£3,177.0	1,990.6			
Desired Saving		£3,606.9	3,589.4			
Household Income		£30,328.4	£34,202.2			
Give up discretionary	Yes	54.2%				
spending to save						
Employment	Part-time	31.7%				
	Full-time	68.3%				
Education	School	5.08%		<ul> <li>Demographics</li> </ul>		
	College	28.8%				
University		66.1%				
Time spend managing Not at all		13.3%				
finances	Occasionally	28.3%				
	Regularly	35.0%		_		
	Often	15.0%				
	Very often	8.3%				

## Table 2. Results from the Financial Affordability Questionnaire

*Prospect Relativity Test.* Note that although the questions related to saving and risk asked the participants to trade-off different variables (e.g., savings versus retirement income in one question, and savings versus risk in another question), we used the weighted average of the answers of each participant across all five questions related to saving and all five questions related to risk, in order to derive the mean values for saving and risk in each condition; and

these averaged results are presented here. This was done because the results showed qualitatively the same pattern and there were no significant differences across the five questions for saving and risk respectively.

There are two sets of responses to consider. First, we present the results on the savings rate. The proportion of times each saving option was chosen in the free choice, low range, and high range conditions is plotted in Figure 2. The presented results were averaged over all participants (which was also done for all statistical tests presented here). The error bars represent the standard error of the mean, which is also presented in all other figures in this article. The mean savings in the high range condition (£3,540) was significantly higher than the full range condition (£2,285), t(38) = 4.94, p < .0001, while the mean savings in the low range condition (£1,660) was significantly lower than the full range condition (£2,343), t(38) = 2.05, p = .0469. Thus, since both the low range and the high range conditions are significantly different from the full range condition in the predicted directions, we can conclude that the range of offered saving options has strongly affected the mean of the selected values in each group.

Now we consider the direct test of whether these data are compatible with stable absolute preferences, by comparing the restricted (high and low range) conditions (using the logic outlined earlier). The proportion of times the lowest saving option in the high range condition (the £3000 option) was selected was .42 and was significantly lower than .77 which is the proportion of times the options below it were selected in the low range condition, t(38) = 3.34, p = .0019. This result indicates that the context has affected choices in the high range condition.

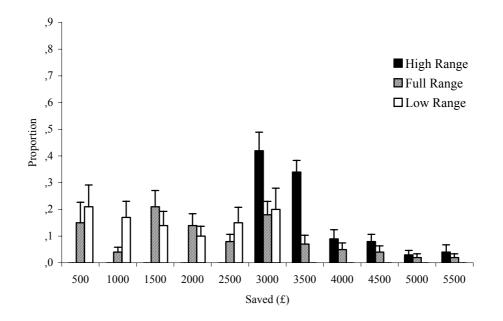


Figure 2. Proportion of times each saving option was chosen in the low range, full range and high range conditions. (Error bars are standard error of the mean).

The proportion of times the highest option in the low range condition (again the £3000 option) was selected was .20 and this value was significantly lower than 0.58, which was the sum of the options in the high range condition, except the lowest option, t(38) = 3.60, p = .0009. This result also means that the hypothesis that participants' choices were unaffected by context should be rejected; the pattern of prospect relativity (Stewart et al., 2003) is evident.

Now we turn to the second set of data, concerning the level of risk people were willing to take with their saving. The proportion of times each investment risk option was chosen in the free choice, low range, and high range conditions is plotted in Figure 3. The average investment risk was significantly higher in the high range condition (56.1%) compared to the free choice condition (31.8%), t(38) = 8.52, p < .0001, and also the average investment risk in the low range condition (22.9%) was significantly lower that in the free choice condition (31.8%), t(38) = 2.35, p = .0239. This result also indicates significant context effects on the mean risk preferred in each condition because both the low range and the high range conditions are significantly different from the full range condition in the expected directions (higher in the high range condition and lower in the low range condition).

Turning to the comparison of the high and low range conditions, the proportion of times the lowest option in the high range condition (the 50% option) was selected was .58 and this value was significantly lower than the proportion of times the options below it were selected in the low range condition, which was .85, t(38) = 3.11, p = .0035. The proportion of times the highest option in the low range condition (again the 50% option) was selected was .11 and this result was significantly lower than .41, which was the sum of the options in the high range condition, except the lowest option, t(38) = 3.60, p = .0009. Again, the results are incompatible with the assumption that people have stable absolute preferences among the choices options, which is a further illustration of prospect relativity.

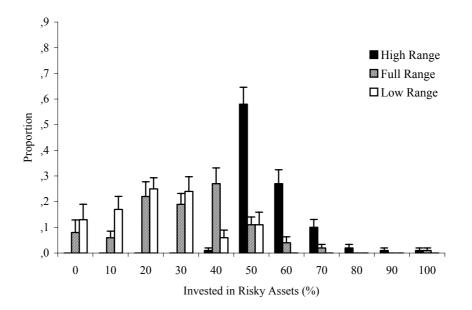


Figure 3. Proportion of times each investment risk option was chosen in the low range, full range, and high range conditions. (Error bars are standard error of the mean).

In summary, the results for saving and risk clearly demonstrate that the choices were strongly influenced by the set of offered choice options.<sup>4</sup> Thus, we replicated our previous findings (Vlaev, Chater, and Stewart, 2007) with realistic financial assumptions and population that does need to make these decisions in real-life. Note, however, that even though the choices were significantly affected by the context in the high range condition, the skewed results in this condition clearly show that there is a tendency towards certain most preferred values for savings and risk. This result suggests that people's preferences are not completely malleable by the context.

Thus, we have demonstrated that prospect relativity arises when people are faced with familiar situations with which they are likely to have some exposure and practice like saving, consumption, pension plans, and investment in the capital markets (at least the media provides enough information on the last issue). It seems also that people might, nonetheless, have developed some more stable preferences or anchors for savings and investment risk, although their responses are still very malleable to context effects. This could be viewed as evidence for some stable absolute scale for assessing money or risk; or, alternatively, as arising because people sample salient comparison 'anchors' from memory (e.g., income, expenditure, current savings---the type of information elicited in the first part of the experiment) against which the choices options are compared (Stewart, Chater, and Brown, 2006).

Risk Preferences. Table 3 presents the results from the five questions measuring respondents' risk preferences and the mean levels of investment risk in each condition. All five risk-aversion measures indicate that the respondents typically perceived themselves to be moderately risk averse. Note that risk-averse preferences are implied by values that are: a) lower than 3.0 for Direct Risk; b) higher than 3.0 for Direct Concern; c) lower than 3.0 for Relative Risk; d) higher than 3.0 for Relative Concern; and e) lower than 0.69 for the Income Gamble, which is the mean between the four indices of relative risk-seeking.

The significance value of the F test in the Analysis of Variance (ANOVA) test shown in Table 3 indicates that the average scores were significantly different between the three conditions (there was a main effect of the factor Condition) only for the investment risk chosen in the main test, F(2, 61) = 65.98, p < .0001. The investment risk levels, in line with our predictions, were lowest in the low range condition and highest in the high range condition. The average scores for the five risk preference tests were very similar (and for most measures the score in the low range condition was even higher than the score in the high range condition). In summary, these results demonstrate that the self-reported subjective risk preferences did not change as the context changed, while the investment risk choices did change accordingly.

However, a rather interesting result in presented in Table 4, which shows the correlations between the risk preference measures and the investment risk in each condition. We used the Spearman's correlation coefficient because some of the measures were quantitative variables (the Investment Risk and the Income Gamble) and some were variables with ordered categories (the questions 1-4). Table 4 shows that in the Low Range condition, there was a

<sup>&</sup>lt;sup>4</sup> Additional analysis also established that the context effects are relatively similar for people from different income ranges. In other words, people who can least afford it are not more or less likely to be influenced by the range of options offered to them. We also did not find any gender differences in terms of saving and investment risk preferences, and context malleability on these two dimensions (i.e., women were not more context sensitive than men). All these analyses are not reported here, but these additional results are available on request.

strong correlation between the Direct Risk measure and Investment Risk, r = .64, p = .0022. In the Full Range condition, the Investment Risk correlated significantly again with Direct Risk, r = .68, p = .0015, but also with Relative Risk, r = .48, p = .0345. In the High Range condition, the Investment Risk correlated significantly only with Direct Risk, r = .50, p = .0264.

		ANOVA			
Risk Measure	Low Range	Full Range	High Range	p	
Investment Risk	22.9%	31.8%	56.1%	.0000	
	(12.1)	(11.7)	(5.00)		
Direct Risk	2.40	2.47	2.20	.5864	
	(1.14)	(0.70)	(0.62)		
Direct Concern	3.30	3.47	3.15	.6863	
	(1.38)	(1.12)	(0.93)		
Relative Risk	2.65	2.75	2.47	.6126	
	(0.93)	(0.91)	(0.77)		
Relative Concern	2.85	3.05	2.84	.7270	
	(1.04)	(0.83)	(0.90)		
Income Gamble	0.70	0.67	0.53	.5921	
	(0.62)	(0.53)	(0.50)		

# Table 3. Means risk levels chosen for each risk preference measure and for Investment Risk in each condition

Standard deviations within parentheses.

Investment Risk is calculated as the mean proportion (%) of the savings invested in the Risky Asset. p is the significance value of the F test in the ANOVA testing the hypothesis that average scores are equal across conditions

In summary, only Direct Risk was significantly associated with risky choice in all three conditions. In other words, within each context, people who selected the options with higher risk also indicated that they were more risk seeking, and vice versa. This result suggests that while people's choices are dependent on the context, their subjective risk-aversion is a stable trait. The predictive power of the simple question measuring Direct Risk could be explained if we assume that people are more or less aware about their risk preferences, but their risk perception is determined by the context. The fact that only the simplest Direct Risk measure was a significant predictor, suggests that people use some very crude heuristics (e.g., "How much risk I am prepared to take?") to select choice options, which are perceived as relatively safe or risky only in comparison to the other available options (i.e., in the current context). Thus, the significant predictive power of the Direct Risk measure implies that people define their preferences in relation to the available set of choice options, which again corroborates our claim that judgments are made relative to the available reference points in the current environment.

Condition	Risk Measures	Investment Risk	Direct Risk	Direct Concern	Relative Risk	Relative Concern	Income Gamble
Ŧ	Investment Risk Direct Risk	- .64**	_				
Low	Direct Concern	.38	.27	_			
Range $(N - 20)$	Relative Risk	.17	.29	10	_		
(N = 20)	Relative Concern	02	02	.39	37	_	
	Income Gamble	.34	.16	.08	.64**	39	_
	Investment Risk	_					
	Direct Risk	.68**	-				
Full Range	Direct Concern	.29	.33	_			
(N = 20)	Relative Risk	.48*	.75**	.08	_		
	Relative Concern	.40	.32	.58**	.38	_	
	Income Gamble	.10	.39	13	.43	10	_
	Investment Risk	_					
High Range (N = 20)	Direct Risk	.50*	_				
	Direct Concern	.07	.08	_			
	Relative Risk	.37	.60**	.19	_		
	Relative Concern	.25	.17	.75**	12	_	
* Comolotion is	Income Gamble	.00	09	41	44	11	_

Table 4. Spearman's rho correlations between the investment risk and the five risk	
aversion measures in the three conditions of the experiment	

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

Investment Risk is calculated as the mean proportion (%) of the savings invested in the Risky Asset

One implications of this result is that directly asking people about their risk preferences may be as useful as apparently more sophisticated risk-diagnostics in helping people to choose financial products. Moreover, our results concerning prospect relativity, and the literature on framing effects more generally, suggest that attempting to associate individuals with an 'economic' risk preference (e.g., the curvature of the utility or value functions) is likely to be ineffective, because results will depend substantially on the framing of the question, rather than reflecting an underlying attribute of the consumer. Nonetheless, simple direct risk measures may still be used to help in the design of the financial products offered by financial advisers, by making the range of offered investment options to vary depending on the risk profile of the consumer. For example, in order to prevent a risk-averse client to make an investment which is too risky for him/her, one could offer a relatively safe range of investment options, and thus utilise the powerful effect of the context in order to accomplish a better match with individual risk preferences.

# **GENERAL DISCUSSION AND CONCLUSIONS**

Our results demonstrate that when people make financial decisions the attractiveness of the choice options significantly depends on the other available options. In particular, the set of options offered as potential savings and risk options was shown to have a large effect on the selected options. In general, it seems that the context provided by items that are considered simultaneously does affect decisions about saving and investment risk. These results could be considered as an example of the prospect relativity principle, which suggests that risky prospects are judged relative to accompanying prospects.

We believe that our new result reflects what would be likely to occur were these choices being made for real, e.g., in a session with a sales person or a financial advisor. The Financial Affordability questionnaire was designed to enforce the participants to make their decisions in light of their real financial circumstances. We believe that this type of test of context effects could not be done in more realistic conditions, because no regulating or legislative authority would allow people's choices of (real) investment products and saving rates to be manipulated in such a drastic way. Retirement investment and saving choices can fundamentally affect a person's quality of life (during her entire life span) and so ethical considerations preclude direct experimentation in the context of a real selling process.

We also tested whether people naturally tend to make decisions matching their individual risk preferences and whether they can be manipulated to pass this level by manipulating the context (which in reality might be necessary in order to accomplish some financial objectives). The results showed that very simple and direct risk preference measures are significant predictors of the responses within a particular context, although these responses differ between different contexts (i.e., the actual choices were very much influenced by the context). One conclusion is that the context has much more powerful effect on risky choices than the underlying risk preferences, but another possible conclusion is that risk preferences are consistently defined in relation to the particular context.

### **Theoretical Accounts**

Range-frequency theory (Parducci, 1965, 1995) is consistent with the result in our experiment, which showed that preferences for saving and risk are very much determined by the range of offered choice options (in particular, preferences for the £3,000 saving option and 50% risk option were different in the high and low context conditions). The success of range-frequency theory (Parducci, 1965, 1995) in accounting for our financial prospective relativity results (and also for choice in gambling, financial, and game theoretic contexts as discussed at the beginning) suggests that the mental representation of utility is analogous to the representation of any other magnitude information, and in particular like the representation of simple perceptual stimuli, as discussed earlier. Nonetheless, as we have already noted, behaviour in our task is by no means entirely driven by context (i.e., choice behaviour is not completely insensitive to whether the choices are from the low, high, or full range). In range-frequency theory, this might be captured, for example, by allowing some prior knowledge about income, expenditure, or current savings, to determine the range of considered options.

The predictive power of the simple question measuring Direct Risk suggests that people are more or less aware about their risk preferences, but their risk perception is determined by the context. Weber (1997) argues the perceived risk is different from risk preferences and that people's perceptions of risk may be different from any theoretical risk measures. Weber also demonstrates that people act on the basis of the perceived risk and that they could have stable responses to perceived risk within particular domains (like health, finance, environment, etc.). In a similar vein, Weber and Milliman (1997) provide support for the hypothesis that factors that change and affect choice also affect risk perception and that inherent risk preference may thus be a constant for a given individual. Furthermore, Weber and Milliman suggest that risk perception may arise from a number of different concerns like for example the chance of injury or loss, magnitude and/or probabilities of losses, aspiration levels/disaster levels, controllability, gain/loss balance, different domains (situational differences), and so on.

In light of this evidence, our results presented here suggest that the context, or the set of choice options, is likely to affect how risky the available options are perceived to be. However, whilst such relative comparisons will allow people to evaluate which options are more risky than others, and even by how much, it does not provide information on how risky the overall set is; all of the options in the set may all be relatively less risky, relatively very risky, or span the entire range of risk (from 0% to 100% as in our study). Thus higher options within the available range would be seen as more risky while the lower options as less risky. Once this conceptualisation of the risky options has taken place, then people would choose according to their genuine risk preferences.

A small number of experiments have investigated the effects of context (defined as the set of available options) on decision making under risk in a way analogous to the effects described here. For example, the set of options available as potential certainty equivalents has been shown to affect the choice of certainty equivalent for risky prospects (gambles). In making a certainty equivalent judgment, participants suggest, or select from a set of options, the amount of money for certain that is worth the same to them as a single chance to play the prospect. Birnbaum (1992) demonstrated that skewing the distribution of options offered as certainty equivalents for simple prospects, whilst holding the maximum and minimum constant, influenced the selection of a certainty equivalent. When the options were positively skewed (i.e., most values were small) prospects were under-valued compared to when the options were negatively skewed (i.e., most values were large).

Benartzi and Thaler (1998, 2001) have found evidence of another effect of the choice set by studying how people allocate their retirement funds across various investment vehicles. In particular, they find evidence for a diversification bias, which they call the 1/n heuristic. The idea is that when an employee is offered *n* funds to choose from in her retirement plan, she divides the money approximately evenly among the funds offered. Use of this heuristic, or others only slightly more sophisticated, implies that the asset allocation an investor chooses will depend strongly on the array of funds offered in the retirement plan. Thus, in a plan that offered one stock fund and one bond fund, the average allocation would be 50% stocks, but if another stock fund were added, the allocation to stocks would jump to two thirds. Read and Loewenstein (1995) also reported that people tend to diversify equally between the set of available options.

These findings illustrate that investors have ill-formed preferences about their investments, which is consistent with Stewart et al.'s (2003) claims. Benartzi and Thaler (2002) asked individuals to choose among investment programs that offer different ranges of

retirement income (for instance, a certain amount of \$900 per month versus a 50-50 chance to earn either \$1,100 per month or \$800 per month). When they presented individuals with three choices ranging from low risk to high risk, they found a significant tendency to pick the middle choice. For instance, people viewing choices A, B, and C, will often find B more attractive than C. However, those viewing choices B, C, and D, will often argue that C is more attractive than B.

In summary, in all these experiments, the manipulated context was related to certain properties of the distribution (like range and rank) of the magnitude attributes of the choice options (like for example, the risk of a financial prospect); and the main purpose was to investigate whether decisions are affected by such manipulations of the simultaneously and sequentially presented options. In these cases, range-frequency theory (Parducci, 1965, 1995) is consistent with the results and could serve as a plausible model explaining the data. Note also that all these experiments were also based on an range-frequency type of account, which assumes that when people judge the attractiveness of, and thus their preferences for, a risky option, they do this by comparing this option with the other available options, instead of matching it with some stable internal scale for (absolute) judgment. Thus, these findings and the evidence that we have reported and reviewed here, present another challenge to the standard rational choice theory.

#### Practical Applications

The results presented here also show that we can increase savings and risky investment by manipulating the range of the choice options. Therefore we accomplished our practical goal to find a way to encourage people to save as much as possible, which is important because current saving rates are much less than the necessary level (see the report by Oliver, Wyman and Company, 2001, which details the UK savings gap), and at the same time also to stimulate them to invest at a higher risk in the capital markets. The rational behind the second aim is that by investing at a higher risk people would experience the least possible decrease in their current consumption because higher market risk would bring higher expected returns and therefore would require less income portion to be saved. In order to accomplish these goals, we used manipulations of the context in which the choice options are presented. We also investigated whether risk preferences affect these financial decisions and the test indicated that these characteristics should not be ignored when financial advice is provided.

The practical relevance of such results can be utilised by using such context manipulation methods during real financial advise, because financial advisers can take a normative stance and encourage people to behave in a direction that is expected to maximise their expected welfare. This also means that our assumption is that people are in principle unable to independently and autonomously make optimal decisions about their financial future, which is what the existing empirical evidence demonstrates as well (e.g., Benartzi and Thaler, 2002). Therefore, the presented results are also a direct test of whether the various documented context effects can be used (in combination) in order to produce certain desirable social objectives. Our results also serve as a good example of how psychological phenomena and decision-making theories could be applied to solve real-world problems.

# **AUTHOR'S NOTE**

Ivo Vlaev and Nick Chater (Department of Psychology, University College London, London, UK).

This work was supported by a grant from the Institute of Actuaries (London) to the Institute for Applied Cognitive Science, Department of Psychology, University of Warwick. We thank members of the actuarial profession Alan Goodman, Martin Hewitt, Ian Woods, and John Taylor for their comments and suggestions. Nick Chater was partially supported by European Commission grant RTN-HPRN-CT-1999-00065, the Leverhulme Trust, and the Human Frontiers Science Program.

# APPENDIX A Financial Affordability Questionnaire

The following questions ask you about various facts and preferences related to your personal finances. We also expect you to provide absolute numbers on your income and expenditure (in pounds). The purpose of this test is to investigate to what extent your real financial circumstances, at the moment, affect your choices in the saving and investment experiment that follows. It is essential to be as accurate and honest as possible. We greatly appreciate your cooperation and we guarantee that the information that you provide will remain strictly confidential.

Please answer the following questions:

- 1. What is your annual income:
- 2. Which of the following statements reflect your financial circumstances (circle the appropriate one and provide the appropriate figures):
  - a) You spend less than you earn; state by how much:
  - b) You spend exactly the amount that you earn:
  - c) You spend more than you earn (for example by borrowing or living on credit); state by how much: \_\_\_\_\_

In order to answer these questions try to estimate to what extent your current annual income is sufficient to cover your necessities, and in particular try to figure out *by how much* your income is sufficient or insufficient to cover your annual expenses (for example, you could say that you spent around £2000 more than your current salary in order to cover your necessities).

- 3. Try to estimate how much you are able to save at the moment. Please write down here your average annual savings: \_\_\_\_\_
- 4. Here we provide you with a list of various types of spending and you have to answer how much of your current annual income is spent on each of these expenditures. There are two types of expenditure examples – essential (e.g., food and rent) and discretionary (e.g., leisure activities), and you have to give estimates of your annual

spending across these categories (for example, you can say that you spend usually  $\pounds 200$  on food,  $\pounds 250$  on rent, and so on).

a) Essential expenditure

Food \_\_\_\_\_\_
Rent / Mortgage \_\_\_\_\_\_
Utilities (electricity, gas, heat, light, water) \_\_\_\_\_\_\_
Car \_\_\_\_\_\_
Other transport (train busses) \_\_\_\_\_\_
Debt repayment \_\_\_\_\_\_
Communications (telephone, etc.) \_\_\_\_\_\_
Childcare and Schooling \_\_\_\_\_\_
Health \_\_\_\_\_\_
Repairs and Maintenance \_\_\_\_\_\_
Other (e.g., health and life insurance, etc.) \_\_\_\_\_\_

#### b) Discretionary expenditure

- Holiday \_\_\_\_
- Entertainment (e.g., cinema)
- Sport \_\_\_\_
- Hobbies \_\_\_\_\_
- Meals and Drinks
- Other \_

TOTAL EXPENDITURE:

- 5. What is the maximum amount that you would like to save per year:
- 6. Can you give up some of your discretionary spending in order to increase your current savings rate if it is below your preferred maximum amount indicated in question 5?

YES / NO (circle the appropriate)

In order to answer this question you need to focus again on your discretionary spending and estimate the degree to which you can readily reallocate money towards a pension. Here we also aim to test how important and essential some of these discretionary expenditures are for you (e.g., some people might be unwilling to give up certain hobbies, sport activities, etc.). Note that in the following experiment we will ask you a series of questions about your preferred savings; if you answer values that are above your current savings (provided in question 3), then we will ask you to give up some of your essential or discretionary spending in order to provide the additional capital that is required to cover the difference between your real current savings rate and the savings rate that you have indicated in some of the test questions.

# APPENDIX **B**

Description of each question in the Prospect Relativity Experiment. The questions are grouped by the key variable that the participants were asked to select (savings or investment risk).

I. *Savings*. The were five questions asking people to choose between savings options formulated as percentage that is saved out of the hypothetical income of  $\pounds 25000$  per year.

- 1. Choose how much to save without information about other variables.
- 2. Choose how much to save and see expected retirement income.
- 3. Choose how much to save and trade it off with retiring at different age and see the expected retirement income.
- 4. Choose how much to save and see the retirement income and its minimum and maximum variability happening because assume that 50% of the savings are invested in the stock market.
- 5. Choose how much to save and take different levels of risk starting from low savings and investment risk and then increase both in parallel.

II. *Risk.* Next are the five questions asking people to choose levels of risk formulated as percentage of saving invested in risky assets:

- 1. Choose how much to invest without information about other variables.
- 2. Choose how much to invest and see expected retirement income and its variability.
- 3. Choose how much to invest and trade-off it with retiring at different age and see the expected retirement income and its variability.
- 4. Choose how much to invest and tradeoff it with amount to be saved (increasing investment corresponding to decreasing savings) and see the retirement income and its variability.
- 5. Choose between levels of variability of the retirement income. Variability reflects different investment strategies and is increasing with the income (higher variability corresponds to higher income).

# APPENDIX C Measures of Risk Aversion Used in the Study

(1) Please indicate here how much risk you are prepared to take on a scale from 1 (not at all – only sure outcomes) to 5 (very much):

Answer: \_\_\_\_\_

(2) How much are you concerned about your financial future? Indicate on a scale from 1 (not at all) to 5 (very much):

Answer: \_\_\_\_\_

(3) Are you more or less willing to take risks than the average person? Indicate using the following scale:

1 - much less

2 - less

3 - the same as the average

4 - more

5 - much more

Answer:

(4) Are you more or less concerned about your financial future than the average person? Indicate using the following scale:

1 - much less

2 - less

3 - the same as the average

4 - more

5 - much more

Answer:

(5) Suppose that you are the only income earner in the family, and you have a good job guaranteed to give you your current (family) income every year for life. You are given the opportunity to take a new and equally good job, with a 50–50 chance it will double your (family) income and a 50–50 chance that it will cut your (family) income by a third. Would you take the new job? Answer with YES or NO here:

Answer:

If your answer to this question is "yes," then answer only question (a) and if your answer is "no," then answer only question (b).

(a) Suppose the chances were 50–50 that it would double your (family) income, and 50–50 that it would cut it in half. Would you still take the new job? Answer with YES or NO here:

Answer:

(b) Suppose the chances were 50–50 that it would double your (family) income and 50– 50 that it would cut it by 20 percent. Would you then take the new job? Answer with YES or NO here:

Answer: \_\_\_\_\_

#### REFERENCES

- Barsky, R. B., Juster, E. T., Kimball, M. S., and Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *Quarterly Journal of Economics*, 112, 537-579.
- Benartzi, S., and Thaler, R. (1998). Illusory diversification and retirement savings. Unpublished manuscript, University of Chicago and UCLA.
- Benartzi, S., and Thaler, R. (2001), Naive diversification strategies in defined contribution saving plans, *American Economic Review*, 91, 79-98.
- Benartzi, S., and Thaler, R. (2002). How much is investor autonomy worth? *Journal of Finance*, *57*, 1593-1616.
- Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choicebased certainty equivalents. *Psychological Science*, 3, 310-314.
- Camerer, C. F. (1995). Individual decision making. In J. Kagel and A. E. Roth (Eds.), *Handbook of experimental economics* (pp. 587-703). Princeton, NJ: Princeton University Press.
- Fechner, G. T. (1966). *Elements of psychophysics* (Adler, H. E., Trans.). New York: Holt, Rinehart, and Winston. (Original work published 1860).
- Garner, W. R. (1953). An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology*, 46, 373-380.
- Garner, W. R. (1954). Context effects and the validity of loudness scales. *Journal of Experimental Psychology*, 48, 218-224.
- Gigerenzer, G., and Todd, P. (Eds.) (1999). Simple heuristics that make us smart. Oxford: Oxford University Press.
- Helson, H. (1964). Adaptation-level theory: An experimental and systematic approach to behavior. New York: Harper and Row.
- Holland, M. K., and Lockhead, G. R. (1968). Sequential effects in absolute judgments of loudness. *Perception and Psychophysics*, 3, 409-414.
- Jesteadt, W., Luce, R. D., and Green, D. M. (1977). Sequential effects of the judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 92-104.
- Kagel, J. H., and Roth, A. E. (1995). *The handbook of experimental economics*, Princeton University Press.
- Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, 47, 313-327.
- Kahneman, D., and Tversky, A. (Eds.) (2000). *Choices, values and frames.* New York: Cambridge University Press and the Russell Sage Foundation.
- Kahneman, D., Slovic, P., and Tversky, A. (1982). Judgement under uncertainty: Heuristics and biases. Cambridge University Press, Cambridge.
- Kapteyn, A., and Teppa, F. (2002). Subjective measures of risk aversion and portfolio choice. Discussion Paper 11, Tilburg University, Center for Economic Research.
- Laming, D. R. J. (1984). The relativity of 'absolute' judgements. *British Journal of Mathematical and Statistical Psychology*, 37, 152-183.
- Laming, D. R. J. (1997). The measurement of sensation. London: Oxford University Press.

- Lockhead, G. R. (1984). Sequential predictors of choice in psychophysical tasks. In: Kornblum, S., and Requin, J. (Eds.), *Preparatory states and processes* (27-47). Hillsdale, NJ: Erlbaum.
- Lockhead, G. R., and King, M. C. (1983). A memory model of sequential effects in scaling tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 461-473.
- Loewenstein, G. (2003). Preferences. Oxford University Press.
- Loomes, G., and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal*, *92*, 805-824.
- Luce, R. D., Nosofsky, R. M., Green, D. M., and Smith, A. F. (1982). The bow and sequential effects in absolute identification. *Perception and Psychophysics*, *32*, 397-408.
- Oliver, Wyman and Company (2001). *The future regulation of UK savings and investment*. Research report conducted by Oliver, Wyman and Company and commissioned by the Association of British Insurers.
- Parducci, A. (1965). Category judgment: A range-frequency theory. *Psychological Review*, 72, 407-418.
- Parducci, A. (1995). *Happiness, pleasure and judgment: The contextual theory and its applications*. Mahwah, NJ: Lawrence Erlbaum.
- Payne, J. W., Bettman, J. R., and Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology*, 43, 87–131.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organisation*, *3*, 323-343.
- Shafir, E., and LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology*, 53, 491-517.
- Sher, S., and McKenzie, C. R. M. (2006). Information leakage from logically equivalent frames. *Cognition*, 101, 467-494.
- Slovic, P. (1995). The construction of preference. American Psychologist, 50, 364-371.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64, 153-181.
- Stevens, S. S. (1975). Psychophysics. New York: Wiley.
- Stewart, N., Brown, G. D. A., and Chater, N. (2002). Sequence effects in categorization of simple perceptual stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 3–11.
- Stewart, N., Brown, G. D. A., and Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review*, 112, 881–911.
- Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53, 1-26.
- Stewart, N., Chater, N., Stott, H. P., and Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General*, 132, 23-46.
- Vlaev, I. and Chater, N. (2006). Game relativity: How context influences strategic decision making. Journal of Experimental Psychology: Learning, Memory and Cognition, 32, 131-149.
- Vlaev, I., Chater, N., and Stewart. N. (2007). Financial prospect relativity: context effects in financial decision-making under risk. *Journal of Behavioral Decision Making*.
- Ward, L. M., and Lockhead, G. R. (1970). Sequential effect and memory in category judgment. *Journal of Experimental Psychology*, 84, 27-34.

- Weber, E. (1997). The utility of measuring and modelling perceived risk. In A. A. J. Marley (Ed.). *Choice, decision and measurement: Essays in honour of R. Duncan Luce* (pp. 59-73). Mahwah, NJ: Erlbaum.
- Weber, E. U., and Milliman, R. A. (1997). Perceived risk attitudes: Relating risk perception to risky choice. *Management Science*, 43, 123-144.

Chapter 30

# THE PSYCHOLOGY OF DECISION-MAKING IN ECONOMICS: A REVIEW<sup>\*</sup>

# Jeffrey A. Livingston and Michael A. Quinn

Bentley College, Waltham, Massachusetts, USA

# ABSTRACT

Economics has always focused on how individuals make decisions. Traditionally, the discipline has viewed individuals as rational agents maximizing their own utility. However, economists have recently begun to incorporate research from the field of psychology in creating a richer view of decision making. This push is the result of challenges to the neoclassical model made by theoretical advances such as Kahneman's Nobel-winning prospect theory model and from the field of experimental economics. This growing field has revealed many aspects of human behavior that cannot be explained by traditional economic models. Some of these aspects of behavior include loss aversion (as explored by Kahneman); relative deprivation (the theory that individuals consider their relative position as compared to others when making decisions), motivations of altruism, fairness, and reciprocity; and the endowment effect (individuals tend to value goods more highly if they are already in possession of them). These innovations have impacted economists' views of issues such as consumption, worker-firm relations, labor supply, equities and real estate.

This paper reviews the impact of psychology on economic models of decision making. The major trends will be discussed, along with implications that these changes have for both economics and public policy.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Economics, Business and Finance*, edited by Klaus P. Hofmann, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

# I. INTRODUCTION

The foundation of modern economics is the neoclassical model of consumption and production, which views individuals as rational agents who seek to maximize their own expected utility. This model has been a powerful tool in accurately predicting how economic agents behave in most circumstances. However, in recent years economists have begun to recognize the model's limitations. Faced with evidence that individuals sometimes behave in ways that are inconsistent with neoclassical predictions, economists have developed a new field called behavioral economics that considers how psychological motivations impact the choices people make. Collaboration between psychologists and economists has resulted in the identification of many such behavioral "anomalies" that are now accepted as fact, and these studies have impacted a wide range of mainstream fields including macroeconomics, labor economics, financial markets and environmental economics.

The impact of psychology on economics tends to be viewed as a new phenomenon. However, many of the notable early writers in economics clearly believed in the importance of psychology as a basis for understanding individual motivations and behavior. For instance, in describing how individuals' human nature and beliefs influences their "rational decisions" David Hume writes "Reason is, and ought only to be, the slave of the passions, and can never pretend to any other office than to serve and obey them." (Hume, 1740). In 1919, Irving Fisher wrote of the interest rate "The basis of interest . . . lies in the preference for present over future goods: Neither the employers more than the employee likes to wait for the fruits of any enterprise . . . whoever does so is entitled to some reward . . . The essence of interest is impatience, the desire to obtain gratifications earlier than we can get them. It is a fundamental attribute of human nature." (Fisher, 1919).

In the 1950s and 1960s, economics moved towards the idea of a "rational" individual. This perspective proposes that individuals are "expected utility maximizers" and discounts emotional responses to situations. This move in economics was made by those trying to make economics into a "hard science". For instance, by the 1950s, economists such as Metzler describe the interest rate as "changes in the quantity of money accompanied by open-market operations do affect the rate of interest and thus the rate becomes a monetary phenomenon in the sense of being co-determined by the quantity of money outstanding and by changes in this quantity" (Metzler, 1951). This stands in stark contrast to Fisher's explanation of the same phenomenon a few decades earlier. It is this contradiction between those establishing the cold, rational individual that would move economics towards the physical sciences and the fact that economists study phenomena of human behavior that has caused problems. Behavior has been observed that simply cannot be explained by a model of "rational" individual behavior.

The resurgence of awareness of psychological motivations in economics in the 1990s is, however, quite different from that of previous economists. While early writers such as Fisher and Keynes saw a significant role for psychology in their theories, this role was based on their own insights and intuition. Recent work in behavioral economics is often based on experiments and data analysis. The methods of experimental economics employ rigorous methodologies to test for the different motivations behind individual behavior and use these insights to inform economic theory. The next section of the paper discusses the methodologies employed in behavioral economics and how they have lead to insights into individuals' motivations. After the discussion of these new methodologies, there is an examination of the basic results of behavioral economics that are no longer considered controversial. This is followed by applications of these results in other mainstream economics fields.

# **II. THE METHODOLOGY OF BEHAVIORAL ECONOMICS**

The primary tool of behavioral economists is experiments that are conducted in the laboratory. Most behavioral economics studies argue that one of the standard assumptions about individual behavior made in neoclassical theory is incorrect, leading to inaccurate predictions about how people behave in some particular situation. Experiments remain the most popular method because they allow the researcher to construct tests where different theories have clearly different predictions about how subjects will behave, so the experiment can cleanly test which theory more accurately predicts behavior.

"Ultimatum" game experiments are a nice example. In this well-known game, two subjects are endowed with some amount of money. One subject, frequently called the "proposer," is allowed to propose how this money should be shared with another subject. The second subject, frequently called the "responder," can accept or reject this offer. If the responder accepts the offer, then the subjects are paid according to the terms of the proposer's offer. If the responder rejects the offer, both subjects receive a payoff of 0. "Rational" responders would accept any offer that contains a positive payoff, since the alternative is a gain of 0, and "rational" proposers would accordingly offer the responder as little as possible, keeping the large majority of the funds for themselves. However, in many repetitions of this experiment, using many different designs, the results have been robust and they have not conformed to this prediction. Rather, subjects frequently reject proposals which they perceive to be unfair, showing a willingness to sacrifice their own monetary gains in order to punish people who treat them unfairly. Camerer and Thaler (1995) provide a nice discussion of how economists have dealt with these results, and Camerer and Loewenstein (2004) provide a clear discussion of why the experimental methodology is necessary to provide clean evidence of this aspect of human behavior:

"Suppose we observed this phenomenon in the field, in the form of failures of legal cases to settle before trial, costly divorce proceedings, and labor strikes. It would be difficult to tell whether rejection of offers was the result of reputation-building in repeated games, agency problems (between clients and lawyers), confusion, or an expression of distaste for being treated unfairly. In ultimatum game experiments, the first three of these explanations are ruled out because the experiments are played once anonymously, have no agents, and are simple enough to rule out confusion. Thus, the experimental data clearly establishes that subjects are expressing concern for fairness." (Camerer and Loewenstein, 2004).

Behavioral economists have only recently caught on to this method of studying human behavior which psychologists have long been familiar with. However, there are four striking differences between the methods used by scholars in the two disciplines. Hertwig and Ortmann (2001) provide a full review of these differences and their relative advantages and disadvantages; for a more thorough discussion, readers are encouraged to consult their article. The first is how the instructions for the experiment are presented to the subjects. Economists, who are always (perhaps too) keenly aware of how people react to any incentive, always follow a pre-written script to the letter. Subjects are given precise descriptions of the players involved, the set of actions that each player can possibly take, and the payoffs resulting from each possible sequence of actions. Subjects are then each assigned a role (for example, as a proposer or a responder in an ultimatum game) and follow the script for that role. Psychologists, meanwhile, frequently do not provide a script or assign roles, forcing participants to ad-lib. Second, economists use repeated trials, while psychologists typically instead conduct "snapshot" studies. Third, economists always use performance-based monetary incentives so that subjects make actual rather than hypothetical choices; psychologists often pay subjects a fee for showing up, but their payment is typically unaffected by how they behave in the experiment. Finally, economists consider it unacceptable to deceive subjects in any way, while psychologists frequently use deception in their experimental designs. Economists have elected to avoid deception on the grounds that it may pollute the behavior of subjects in future experiments. It would become impossible to rule out the possibility that subject behavior was guided by their suspicion of being mislead, rather than the monetary rewards that are assigned to each action.

The frequent use of experiments by behavioral economists leads to some confusion that behavioral economics and experimental economics are one in the same. This perception is often used as ammunition against the field – doubters present the important argument that behavioral anomalies found in experimental studies may not be prominent in real markets. The subjects of these experiments are typically university students who are very different from people who participate in actual markets. They are typically younger, poorer, and less experienced than participants in the business world, so it is easy for skeptics to dismiss laboratory studies on the grounds that once people gain more experience in the market, they will behave in ways that are more in concert with the neoclassical model of behavior.

In fact, experimental and behavioral economics are not synonyms: many experimental studies have been conducted that are not designed to challenge neoclassical behavioral assumptions, and behavioral studies are frequently conducted using non-experimental methods. Behavioral economists use all of the same tools used by economists in other fields – careful analysis of naturally occurring field data, theory, and most recently "field" experiments that are conducted in an actual market rather than in the laboratory, as well as traditional laboratory experiments. In fact, perhaps the most widely cited behavioral study is the theoretical paper by Kahneman and Tversky (1979), which introduced "prospect theory."

Field experiments in particular have quickly become very popular in recent years. The chief advantage of field experiments is of course the ability to study the behavior of actual market participants rather than students in the lab, while still maintaining some control of the experimental environment, making it possible to rule out many alternative explanations for results. The disadvantages are that control of the setting is no longer perfect, and field experiments tend to be enormously expensive to conduct. Examples of this work include Gneezy and Rustichini (2000), who studied the effect of introducing a small fine charged to parents who failed to pick up their children on time from a day-care center in Israel, and List (2003), who examined the "endowment" effect (the tendency for people to place a higher value on something if they are already in possession of it) among two sets of agents at a

sports memorabilia show – agents who had little experience in this marketplace, and highly experienced dealers who were very familiar with the market.

# **III. BASIC THEMES AND RESULTS**

Studies using all of these methodologies have over time come to a consensus on several different results, each of which has seen application in more mainstream fields. The most important of these results include:

# Risk Preferences, Loss Aversion, the Endowment Effect and Framing Effects

Perhaps the seminal paper in all of behavioral economics is "Prospect Theory: An Analysis of Decision Under Risk" by Daniel Kahneman and Amos Tversky, published in *Econometrica* in 1979. Kahneman was awarded the Nobel Prize in Economics in 2002 "for having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty" largely on the basis of this paper. The theory is motivated by hypothetical experiments that showed several situations where people make decisions that are inconsistent with the assumptions of expected utility theory. For example, they surveyed 72 people and asked them to state what choice they would make when confronted with two problems:

Problem 1: Choose between	
A: 2,500 with probability 0.33	B: 2,400 with certainty
2,400 with probability	0.66
0 with probability	0.01
Problem 2: Choose between	D. 2 400
A: 2,500 with probability 0.33 0 with probability 0.67	B: 2,400 with probability 0.34 0 with probability 0.66

In problem 1, 82 percent of the respondents selected option B, while in problem 2, 83 percent of the respondents selected option A. In oversimplified terms, expected utility theory states that preferences over several possible outcomes  $x_1, ..., x_n$  that occur with probabilities  $p_1, ..., p_n$  respectively can be represented by the following function:

 $U(x_1,p_1,...,x_n,p_n) = p_1u(x_1) + ... + p_nu(x_n)$ , where the utility function u(x) is concave, i.e. people are risk averse. In this case, the responses to problem 1 suggest that for most people,

u(2400) > 0.33u(2500) + 0.66u(2400),

or

0.34u(2400) > 0.33u(2500).

The responses to problem 2, however, suggest that for most people,

0.33u(2500) > 0.34u(2400),

the opposite inequality. Clearly, expected utility theory is missing something. Kahneman and Tversky's theory of preferences explains why people might make these choices, as well as several other choices which violate expected utility theory.

The theory has a number of important implications. The first is *loss aversion*: people dislike losing more than they like winning. In more technical terms, they lose more utility from a loss of x than they gain utility from an increase of the same amount of x. A related concept involves how people evaluate outcomes: they compare outcomes to a reference point, often caring more about whether a choice results in a gain or a loss, not about the resulting level of wealth that they end up with. Somewhat paradoxically, however, people tend to be risk averse when choosing among possible gains, as in problem 1 above, but risk loving when choosing among possible losses – they will often choose a gamble that has a chance to result in a big loss, but also has a chance to avoid losses altogether, over a certain loss of intermediate value.

A corollary of loss aversion is the *endowment effect*: people tend to place more value on items that they already have in their possession (i.e., that are already part of their endowment). Loss aversion explains this effect: people overvalue objects that they possess because the utility they lose from giving up what they already have is greater than the utility they gain from obtaining something that they did not have previously. Experimental studies that show this effect exists include Knetsch (1992), Kahneman, Knetsch and Thaler (1990), and Bateman et. al. (1997). For example, Knetsch endowed half of his subjects with a coffee mug, and half of his subjects with a pen. All subjects were then allowed to exchange their good for the other option if they so desired. If there is no endowment effect, we would expect half of the subjects to agree to the exchange. However, only 22 percent chose to trade for the other good, showing that most subjects valued the good they were initially randomly endowed with more highly. Recent evidence from field experiments suggests that as people gain market experienced collectors at a sports memorabilia show display an endowment effect, but dealers who are highly experienced in the market show no endowment effect at all.

Prospect theory also explains many *framing effects*, where people may make a different choice depending on how the options are presented. The most famous example is presented in Tversky and Kahneman (1981). In this hypothetical problem, respondents are told that 600 people are threatened by a disease, and they are asked to choose between two options. One set offers the options in a "positive frame":

Option A: save 200 lives for sure

- Option B: a one-third chance of saving all 600 lives, and a two-thirds chance of saving 0 lives
- Another set offers the options in a "negative frame":

Option C: 400 people die for sure

Option D: a two-thirds chance of all 600 dying, and a one-third chance of nobody dying

Despite the fact that the options result in the same outcomes with the same probabilities, most people chose option A over option B, but most chose option D over option C. When framed as a gain, people tend to prefer the safe option, but when framed as a loss, people tend to prefer the risky option.

#### Social Preferences: Fairness, Reciprocity, Altruism and Trust

The standard neoclassical model of behavior assumes that people are self-interested, and only care about maximizing their own wealth. Behavioral economists have assembled a large amount of evidence that there are exceptions to this rule. One exception concerns fairness and reciprocity: people often have a strong preference for being treated fairly. This preference is revealed by their exhibition in experimental studies of *negative reciprocity*: a willingness to sacrifice their own wealth in order to punish people who they believe have treated them unfairly. As Camerer and Loewenstein (2004) point out, the ultimatum game discussed in the introduction has been studied in over 20 countries, and the results have been consistent. Responders lower their own earnings in order to punish proposers by rejecting offers of less than one fifth of the available funds about half the time, and a large majority of proposers anticipate this rejection, offering between one third and one half of the available funds to the responder.

Studies have also shown that people exhibit *positive reciprocity*: a willingness to sacrifice their own wealth in order to reward people who they believe have been helpful or treated them fairly. This behavior has been observed in many different contexts. Fehr et. al. (1993) find that subjects cast in the role of a firm on average pay higher costs to increase product quality after buyers offer to buy at a higher, more fair price. Andreoni et. al. (2003) find that in an ultimatum game where responders have the option to reward as well as punish, the size of rewards increases with the amount offered to the responder. Charness (2004) finds that subjects placed in the role of workers are willing to exert more effort when employers initially offer higher wages.

Positive reciprocity is also observed in "trust" games, which have been studied by work such as Berg, Dickhaut and McCabe (1995) and Glaeser et. al. (2000). In a typical version of this game, a first mover is given ten dollars, and can choose to send any portion of it to a second mover. Each dollar that is sent is tripled by the experimenter, so if the first mover sends ten dollars, the second mover receives 30 dollars. The second mover can then return any amount to the first mover. Rational second movers would keep everything they receive for themselves, and anticipating this, rational first movers would send nothing. However, the studies show that first movers frequently send at least some money to the second mover, and second movers frequently reward first movers for trusting them by returning a larger percentage of what they receive when the amount they receive increases.

Even when there is no chance of punishment or reward, people exhibit altruism in "dictator" games, such as the ones studied in Hoffman et. al. (1994) and Eckel and Grossman (1996). In these games, one player, the "dictator," is given some amount of money, and is allowed to send some of it to a second player if they so choose, but the game ends there so there are no possible financial consequences to sending the second player nothing. Despite this fact, subjects frequently send substantial amounts to the second player. In the experiments of Hoffman et. al. (1994) the dictator was given \$10 and knew the second

player's identity. Out of 48 dictators, 14 sent \$3, 5 sent \$4, and 3 sent \$5. The rest sent less than \$3.

The motivation behind these behaviors is not entirely clear. Charness (2004) shows, for example, that the positive reciprocity he observes is due both to a desire to reward employers for fair treatment, as well as a desire to keep payoffs equal regardless of the actions of others. Several theories have been offered as explanations. Rabin (1993) presents a model where people gain psychological payoffs from kind treatment. Fehr and Schmidt (1999) suggest a model of inequality aversion where people find all inequality objectionable, but their distaste for inequality is stronger if they are the ones at a disadvantage. Finally, Charness and Rabin (2002) present a "quasi-maximin" model where people prefer the payoffs of the worst-off individual to be as high as possible, but also are concerned with the size of total payoffs for all individuals.

# **IV. APPLICATIONS OF PSYCHOLOGY IN ECONOMICS**

The recent advances in behavioral economics stemming from increased use of psychological motivations have many applications to economic decision making. This section will discuss work in the areas of money illusion, fair wages and unemployment, intertemporal substitution, rising consumption profiles over time, equity premiums, real estate markets and environmental economics.

## Money Illusion

Money illusion refers to individuals' decision making being influenced by nominal (rather than real) changes in variables. Nominal values of a variable are those actually observed, while "real" values take into account changing prices. Economic theory holds that individuals should only consider "real" values when making decisions, such as those regarding salaries and contracts. However, even early on, economists observed that individuals did not always base their decisions on real values. The idea that nominal accounting methods (i.e. inflation causing a change in nominal but not real values) can impact decision making has been cited by economists throughout the 20<sup>th</sup> century (Fisher, 1928; Patinkin, 1965; Fischer and Modigliani, 1986).

Early cases of money illusion observed by economists included downward rigidities in nominal wages. Individuals resist decreases in nominal wages more than they resist the same change in their real wage being caused by inflation. In 1936, Keynes noted

"Since there is imperfect mobility of labour, and wages do not tend to an exact equality of net advantage in different occupations, any individual or group of individuals, who consent to a reduction of money-wages relatively to others, will suffer a relative reduction in real wages, which is a sufficient justification for them to resist it. On the other hand it would be impracticable to resist every reduction in real wages, due to a change in the purchasing-power of money which affects all workers alike; and in fact reductions of real wages arising in this way are not, as a rule, resisted unless they proceed to an extreme degree." (Keynes, 1936). While Keynes and others noted this phenomenon, they did not have the psychological insights necessary to examine the motivations behind this behavior. While even economists in the early 20<sup>th</sup> century observed money illusion, they did not have the means to satisfactorily explain why it occurred. The psychological insight necessary for explaining instances of money illusion comes from framing effects.

Framing effects refers to individuals having different responses to the same situation if it is presented in a different manner to them (Tversky and Kahneman, 1981; Kahneman and Tversky, 1984). It has been observed that nominal wages rarely fall, even in times of recession (Bernanke and Carey, 1996; Bewley, 1998; Kahn, 1997). Solow's theory of efficiency wages proposes that individuals react to changes in their nominal wage, not just their real wage (Solow, 1979). If people observe their nominal wage falling then they may put forth less effort, lowering productivity and reducing any gain which the firm made in lowering wages. Solow proposes that an individual faced with a real wage falling by 2% will react better to a scenario in which nominal wages rise by 3% in a period of 5% than they will if nominal wages fall by 1% in a period of 1% inflation. The fall in real wages is 2% in either case but the individual perceives the situation differently and so has a different response regarding work effort. This is explained by framing effect but not by the "rational" economic view of the individual.<sup>1</sup>

A second observation of money illusion comes in the form of contracts. In particular, if individuals were concerned with "real" values then contracts for wages and loans should be indexed to inflation (so that the real value does not change). However, except under periods of high inflation, it is rare to see indexed contracts. The one exception to this tends to be union contracts which are often indexed. However, only 12% of the American labor force is unionized (CPS, 2007).

#### The Fair Wage-Effort Hypothesis and Unemployment

Psychological insights concerning with reference groups and relative deprivation has lead to advances in economic theories of worker productivity and unemployment. Akerlof and Yellen's (1990) work specifically deals with how worker effort is impacted by their feelings of "fair" treatment. This model stood in contrast to the prevailing orthodoxy in economics which held that worker productivity was impacted only by variables such as education, capital and technology. Traditional economics had no role for factors such as workers' emotional responses to the perception of their relative treatment. It should be noted that Akerlof and Yellen proposed their model as an expansion of current models. They also incorporated traditional economic variables into their model.

The basis of the fair wage-effort model is relative deprivation theory. Relative deprivation proposes that individuals view their situation as compared to a reference group (Runciman, 1966). In this application, individuals view their wages in reference to wages of their peers. Worker effort becomes a function of how people view their relative wages. If people believe they are relatively deprived with regards to wages then they will seek to punish the firm by reducing their work effort. Individuals will match their effort to their

<sup>&</sup>lt;sup>1</sup> For a more comprehensive and detailed survey of research on money illusion, readers are directed to Shafir, Diamond and Tversky (1997).

perceived treatment. Also, individuals who believe they are being paid relatively well may put forth additional effort.

This theory can help to explain why firms might pay above market clearing levels. On an economy-wide basis, this can result in involuntary unemployment. If hiring workers at market clearing levels would result in them being paid lower than their peers either at the firm or in the industry then these workers may put forth less effort. This lower effort increases the marginal cost of production to the point at which the firm actually could have lowered its costs by paying its workers more.

#### Intertemporal Substitution

One of the commonly held views in labor economics has been that people will choose to work more hours if they are offered a higher wage (Lucas and Rapping, 1969). The higher wage increases the opportunity cost of leisure time so people decide to work more hours (and consequently earn and consume more). Despite the popularity in mainstream economics of this proposition, the results from empirically testing it have been mixed (Altonji, 1986; Browning, Deaton and Irish, 1985; Laisney, Pohlmeier and Staat, 1992; Mankiw, Rotemberg and Summers, 1985; Mulligan, 1995). Thus, this view has recently been challenged by evidence suggesting that people do not make labor supply decisions in such a manner.

An alternative view is that people work to achieve income/consumption targets and will work whatever hours are necessary to achieve these goals. Camerer et al., (1997) tracked the hours worked by cab drivers on days with more fares (resulting in higher hourly wages) and those with less fares to see how their choice of hours worked was altered (cab drivers in New York have flexibility in choosing the number of hours worked).<sup>2</sup> The result which they found is that people did not work more on days with a higher per hour wage; rather, they actually worked more hours on days with a lower hourly wage. This is the opposite prediction of traditional economic theory. However, if drivers were targeting a daily income number then it is completely logical that they would work less on more lucrative days. Drivers also seemed to exhibit loss aversion as they would work considerably longer hours to avoid missing their target but would not work many more hours to increase gains above the target on good days. Daily targeting is also consistent with the idea that individuals are concerned with a possible lack of self-control (Shefrin and Thaler, 1988). Without daily targets, individuals may work less hours assuming that they would then make up the income on another day but then fail to do so.

#### Seniority and Rising Consumption Profiles

Economists have sought to explain why wages rise faster with seniority than does worker productivity (Lazear, 1979). Explanations that involve firm-specific training and productivity changes with experience have failed to account for the differences seen between junior and

<sup>&</sup>lt;sup>2</sup> Similar types of studies have been done with farmers (Berg, 1961; Orde-Brown, 1946) and with self-employed proprietors (Wales, 1973).

senior employees at firms. While traditional economics cannot explain this, it is completely consistent with a view of the individual that incorporates "demonstration effects".

Demonstration effects propose that individuals attempt to match the consumption of their peers. This results partially from a lack of self-control by individuals. Showing the influence of psychology on economics early on, this idea goes back in economics to the work of Duesenberry and his relative income hypothesis (Duesenberry, 1949). According to this idea, individuals attempt to match the consumption of their peers, regardless of whether or not they can afford this consumption. Duesenberry popularized the phrase "keeping up with the Joneses" in explaining this phenomenon.

If individuals react in such a manner then this is problematic for lifetime savings if wages are flat or decrease over time. In such cases, individuals would overconsume when young and would not have sufficient income later to save for retirement. Therefore, workers prefer a consumption profile that rises over time with seniority. Rising earnings profiles can be seen as a form of forced savings (Frank and Hutchens, 1993; Lowenstein and Sicherman, 1991; Neumark, 1995).<sup>3</sup> This is preferred even though productivity may not be rising as much over time, therefore workers are willing to accept lower wages when younger for the promise of higher wages later in life.

### The Equity Premium Puzzle

Economists have long wondered why people buy bonds. This is an apt question as the long-run return to bonds is much lower than that of stocks. There would be some difference expected in returns because of the additional risk of stocks (individuals are risk averse). However, this difference in returns is far higher than the difference predicted by theories of risk aversion. The unexplained difference in returns between stocks and bonds is called the equity premium. This puzzle was first coined by Mehra and Prescott (1985). This phenomenon was found in further studies using earlier stock and bond data (Siegel, 1992a,b; Weil, 1989) as well as in studies using international data (Campbell, 1996; Gielen, 1994; Hirose and Tso, 1995).

Recent work has sought to explain this puzzle using prospect theory (Benartzi and Thaler, 1995; Siegel and Thaler, 1997). As discussed earlier, prospect theory is based on loss aversion and mental accounting (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). These properties lead individuals to prefer returns that exhibit low variability, even in the short-run. Individuals want low variability because they are loss averse (Kahneman, Knetsch and Thaler, 1990; Tversky and Kahneman, 1991). The impact of this loss aversion on its own is not enough to explain the entire equity premium. However, the mental accounting aspect of prospect theory considers what the length of evaluation period individuals have (Kahneman and Tversky, 1984; Thaler, 1985). In particular, if individuals have a short length of evaluation then they will periodically review their portfolios and at any given point stocks may not be outperforming bonds. In fact, it is not uncommon for stocks to have losses in the short-term. So, the recurring short evaluation periods combined with loss

<sup>&</sup>lt;sup>3</sup> Other types of forced savings mechanisms such as Christmas clubs are discussed by Loewenstein and Thaler (1989), Loewenstein and Prelec (1992), Stigler (1966) and Thaler and Shefrin (1981).

aversion results in individuals still purchasing significant amounts of bonds despite the much lower returns.

#### Other Asset Markets: Real Estate

Given the scale of the real estate market, the behavior of buyers and sellers is of significant interest to economists. Economists have observed a negative relationship between selling prices and volume that could not be explained by "rational" economic behavior (Genesove and Mayer, 2001; Ortalo-Magne and Rady, 1998; Stein, 1995). However, prospect theory has offered a viable explanation for this result.

It has been observed that when prices fall in the housing market, many individuals will refuse to sell their home. The inventory of unsold properties increases as individuals become unwilling to sell their home "at a loss". Individuals will continue to offer their homes at selling prices well above the market rather than sell at market prices.

Prospect theory can explain this behavior through reference points and loss aversion. When individuals are selling their homes, they view the selling price relative to a reference point. This is often, but not necessarily, their original purchase price. This reference could also be a previous high point in the market. Individuals feel that if they sell the property for less than this reference point then they are selling "for a loss". Prospect theory also holds that people exhibit loss aversion so they feel losses more than they do gains (Tversky and Kahneman, 1991). Therefore, individuals become averse to sell their property for what they perceive to be even a small "loss."

#### **Environmental Economics**

A cornerstone of environmental economics is the contingent valuation (CV) method, which is used to estimate the value of all kinds of environmental services and benefits. These estimates are often used in an analysis of the costs and benefits of a policy that may cause environmental damage. The method involves conducting a survey to assess the monetary value people place on the environmental damage that would result from the policy. One major criticism of this method is related to the endowment effect. Depending on how the question is phrased, the reported valuation of the environmental benefits of a policy can be very different. Questions that ask how much you would be willing to pay (WTP) for a particular benefit typically elicit lower values than questions that ask how much you would be *willing to accept* (WTA) to give up the same benefit. Hammack and Brown (1974) found in a survey of duck hunters that they would be willing to accept the destruction of a wetland needed to support the duck population at its current level in return for \$1,044 on average, but they would be willing to pay only \$247 on average to maintain the wetland. In an experimental study, Knetsch and Sinden (1984) estimate that the difference between the amount subjects were willing to pay for a lottery ticket was \$1.28, but the price that subjects had to be paid before they were willing to give up the same ticket when already endowed with it was \$5.18. This wide divergence in value measurement based on the phrasing of the question has highlighted the importance of identifying whether the appropriate measure in each particular context is willingness to pay or willingness to accept, and in general has called the validity of the CV method into question.

# **V. CONCLUSION**

The neoclassical economic view of decision making has encountered several challenges in recent years. These challenges have been motivated by empirical observations of phenomenon that neoclassical models of decision making cannot explain. In particular, there has been an increasing use of psychological motivations in explaining individual behavior that does not fit the predictions of the neoclassical "rational agent".

These advances in behavioral economics have crossed many fields in economics. For example, economists' views on topics such as unemployment, wages, work hours, stock prices, real estate values and environmental valuations have all been impacted by the increased use of psychological motivations for human behavior. While these may seem to be disparate threads of research there are some common themes. Ideas such as prospect theory, fairness and reciprocity, and relative deprivation/consumption can be seen as running themes which impact several fields in economics. These themes represent an alternative view of decision making and therefore have a broad impact in economics, as individual decision making is at the core of economics. Therefore, advances in decision making such as prospect theory will no doubt continue to have a wide-ranging impact on the discipline of economics in the future.

## REFERENCES

- Akerlof, George and Janet Yellen. 1990. "The Fair Wage-Effort Hypothesis and Unemployment," *Quarterly Journal of Economics*, 105, 255-283.
- Altonji, Joseph G. 1986. "Intertemporal Substitution in Labor Supply: Evidence from Microdata," *Journal of Political Economy*, 94(3), S176-S215.
- Andreoni, James, William Harbaugh and Lise Vesterlund. 2003. "The Carrot or the Stick: Rewards, Punishments, and Cooperation," *The American Economic Review*, 93(3), 893-902.
- Bateman, Ian, Alistair Munro, Bruce Rhodes, Chris Stamer and Robert Sugden. 1997. "A Theory of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 112(2), 479-505.
- Benartzi, Shlomo and Richard H. Thaler. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110(1), 75-92.
- Berg, Elliot J. 1961. "Backward-Sloping Labor Supply Functions in Dual Economies-The Africa Case," *Quarterly Journal of Economics*, 75, 468-492.
- Berg, Joyce, John Dickhaut and Kevin McCabe. 1995. "Trust, Reciprocity and Social History," *Games and Economic Behavior*, 10, 122-142.
- Bernanke, Ben S. and Kevin Carey. 1996. "Nominal Wage Stickiness and Aggregate Supply in the Great Depression," *Quarterly Journal of Economics*, 111(3), 853-883.

- Bewley, Truman F. 1998. "Why Not Cut Pay?" *European Economic Review*, 42(3-5), 459-490.
- Browning, Martin., Angus Deaton and Margaret Irish. 1985. "A Profitable Approach to Labor Supply and Commodity Demands over the Life-Cycle," *Econometrica*, 53, 503-543.
- Camerer, Colin F., Linda Babcock, George Loewenstein and Richard H. Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time," *Quarterly Journal of Economics*, 112(2), 407-441.
- Camerer, Colin F. and George Loewenstein. 2004. "Behavioral Economics: Past, Present and Future," in Camerer, Loewenstein and Rabin, eds., *Advances in Behavioral Economics* (New York: Russell Sage Foundation).
- Camerer, Colin F. and Richard H. Thaler. 1995. "Anomalies: Dictators, Ultimatums, and Manners," *Journal of Economic Perspectives*, 9, 209-219.
- Campbell, John. 1996. "Consumption and the Stock Market: Interpreting International Experience," NBER Working Paper No.5619.
- Charness, Gary. 2004. "Attribution and Reciprocity in an Experimental Labor Market," *Journal of Labor Economics*, 22, 665-688.
- Charness, Gary and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 117, 817-869.
- Current Population Survey. 2007. Percentage of Workforce Unionized. Accessed online at http://www.bls.gov/cps/cpsaat40.pdf.
- Duesenberry, J.S. 1949. *Income, Savings and the Theory of Consumer Behavior*. Cambridge: Harvard University Press.
- Eckel, Catherine C. and Philip J. Grossman. 1996. "Altruism in Anonymous Dictator Games," *Games and Economic Behavior*, 16, 181-191.
- Fehr, Ernst, Georg Kirchsteiger and Arno Riedl. 1993. "Does Fairness Prevent Market Clearing? An Experimental Investigation," *Quarterly Journal of Economics*, 108(2), 437-459.
- Fehr, Ernst and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114(3), 817-867.
- Fisher, Irving, 1919. Elementary Principles of Economics. New York: Macmillan.
- Fischer, Stanley and Franco Modigliani. 1986. "Towards and Understanding of the Real Effects and Costs of Inflation," *Weltwirtschaftliches Archiv*, 114, 810-833.
- Frank, Robert H., and Robert M. Hutchens. 1993. "Wages, Seniority and the Demand for Rising Consumption Profiles," *Journal of Economic Behavior and Organization*, 21, 251-276.
- Genesove, David. and Christopher Mayer. 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 116 (4), 1233-1260.
- Gielen, Gregor. 1994. Konen Aktienkurse noch Steigen? Wiesbaden, Germany: Gabler.
- Glaeser, Edward L., David I. Laibson, Jose A. Scheinkman and Christine L. Soutter. 2000. "Measuring Trust," *Quarterly Journal of Economics*, 115(3), 811-846.
- Gneezy, Uri and Aldo Rustichini. 2000. "Pay Enough or Don't Pay at All," *Quarterly Journal* of Economics, 115(4), 791-810.
- Hammack, Judd and Gardner Mallard Brown. 1974. *Waterfowl and Wetlands: Toward Bioeconomic Analysis*. Baltimore: Johns Hopkins Press for Resources for the Future.

- Hertwig, Ralph and Andreas Ortmann. 2001. "Experimental Practices in Economics: A Methodological Challenge for Psychologists," *Behavioral and Brain Sciences*, 24(4), 383-403.
- Hirose, Hiroaki., and Yi Tso. 1995. "Japanese Market Returns," mimeograph, The Wharton School of the University Of Pennsylvania.
- Hoffman, Elizabeth, Kevin McCabe, Keith Shachat and Vernon Smith. 1994. "Preferences, Property Rights and Anonymity in Bargaining Games," *Games and Economic Behavior*, 7, 346-380.
- Hume, David. 1740. Treatise on Human Nature. New York: Oxford University Press, 2000.
- Kahn, Shulamit. 1997. "Evidence of Nominal Wage Stickiness from Microdata," *American Economic Review*, 87(5), 993-1008.
- Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler. 1990. "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 98, 1325-1348.
- Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263-291.
- Kahneman, Daniel and Amos Tversky. 1984. "Choices, Values and Frames," American Psychologist, 39, 341-350.
- Keynes, John M. 1936. *The General Theory of Employment, Interest and Money*. London: Macmillan; New York. Harcourt Brace.
- Knetsch, Jack L. 1992. "Preferences and Nonreversibility of Indifference Curves," *Journal of Economic Behavior and Organization*, 17(1), 131-139.
- Knetsch, Jack L. and J.A. Sinden. 1984. "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value," *Quarterly Journal of Economics*, 99(3), 507-521.
- Laisney, Francois., Winfried Pohlmeier and Matthias Staat. 1992. "Estimation of Labor Supply Functions Using Panel Data: A Survey," in Matyas and Sevestre, eds., *The Economics of Panel Data: Handbook of Theory and Applications* (Dordrecht, the Netherlands: Kluwer, 1992).
- Lazear, Edward. 1979. "Why is There Mandatory Retirement?" Journal of Political Economy, 87, 1261-1284.
- List, John A. 2003. "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics*, 118, 41-71.
- List, John A. 2004. "Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace," *Econometrica*, 72(2), 615-625.
- Loewenstein, George. 1988. "Frames of Mind in Intertemporal Choice," *Management Science*, 34, 200-214.
- Loewenstein, George and Nachum Sicherman. 1991. "Do Workers Prefer Increasing Wage Profiles," *Journal of Labor Economics*, 9, 67-84.
- Loewenstein, George and Drazen Prelec. 1992, "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Quarterly Journal of Economics*, 107(2), 573-597.
- Loewenstein, George and Richard H. Thaler. 1989. "Anomalies: Intertemporal Choice," *Journal of Economic Perspectives*, 3, 181-193.
- Lucas, Robert E., and Leonard A. Rapping. 1969. "Real Wages, Employment and Inflation," *Journal of Political Economy*, 77, 721-754.

- Mankiw, N.Gregory., Julio J. Rotemberg and Lawrence H. Summers. 1985. "Intertemporal Substitution in Macroeconomics," *Quarterly Journal of Economics*, 100, 225-251.
- Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle," Journal of Monetary Economics, 15, 145-162.
- Metzler, Lloyd A. 1951. "Wealth, Saving, and the Rate of Interest," *Journal of Political Economy*, 59, 93-117.
- Mulligan, Casey B. 1995. "The Intertemporal Substitution of Work What Does the Evidence Say?" University of Chicago, Population Research Center Discussion Paper Series #95-11.
- Neumark, David. 1995. "Are Rising Earnings Profiles a Forced-Savings Mechanism?" *Economic Journal*, 105(428), 95-106.
- Orde-Brown, G. 1946. Labour Conditions in East Africa. London: Colonial Office, H.M.S.O.
- Ortalo-Magne, Francois and Sven Rady. 1998. "Housing market fluctuations in a life cycle economy with credit constraints," Stanford University GSB Research paper 1501.
- Patinkin, Don. 1965. Money, Interest and Prices. New York: Harper and Row.
- Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics," American Economic Review, 83(5), 1281-1302.
- Runciman, Walter G. 1966. *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth-Century England*. University of California Press.
- Shafir, Eldar., Peter Diamond and Amos Tversky. 1997. "Money Illusion," Quarterly Journal of Economics, 112, 341-374.
- Shefrin, Hersh M. and Richard H. Thaler. 1988. "The Behavioral Life-Cycle Hypothesis," *Economic Inquiry*, 26, 609-643.
- Siegel, Jeremy J. 1992a. "The Equity Premium: Stock and Bond Returns Since 1802," *Financial Analyst Journal*, 48, 28-38.
- Siegel, Jeremy J. 1992b. "The Real Rate of Interest from 1800-1990: A Study of the U.S. and U.K.," *Journal of Monetary Economics*, 29, 227-252.
- Siegel, Jeremy J. and Richard H. Thaler. 1997. "Anomalies: The Equity Premium Puzzle" *Journal of Economic Perspectives*, 11(1), 191-200.
- Solow, Robert M. 1979. "Another Possible Source of Wage Stickiness," Journal of Macroeconomics, 1, 9-82.
- Stein, Jeremy C. 1995. "Prices and Trading Volume in the Housing Market: A Model with Down-Payment Constraints," *Quarterly Journal of Economics*, 110, 379-406.
- Stigler, George J. 1966. The Theory of Price. New York: Macmillan.
- Thaler, Richard H. 1985. "Mental Accounting and Consumer Choice," *Marketing Science*, 4, 199-214.
- Thaler, Richard H. and H.M. Shefrin. 1981. "An Economic Theory of Self-Control," *Journal* of *Political Economy*, 89(2), 392-406.
- Tversky, Amos and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice," *Science*, 211, 453-458.
- Tversky, Amos and Daniel Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference Dependent Model," *Quarterly Journal of Economics*, 106(4), 1039-1061.
- Tversky, Amos and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297-323.
- Wales, Terence J. 1973. "Estimation of a Labor Supply Curve for Self-Employed Business Proprietors," *International Economic Review*, 14, 69-80.

Weil, Philippe. 1989. "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401-421.

Chapter 31

# THE CONFLICT BETWEEN MONEY AND SELF-ESTEEM IN INVESTMENT<sup>\*</sup>

# Liqing Zhang<sup> $\dagger$ </sup>

Peking University, Beijing, China

# ABSTRACT

The present theory proposes that investors not only think of future monetary benefits, but also value the choices' implications regarding their self-esteem in decision making. Self-esteem is one's subjective evaluation of the self. Most people want to maintain a positive self-image. When they decide to invest in a project, people expect to receive financial rewards, and they also hope to enhance their self-esteem through the success of the project. Thus, when their initial investment produces negative economic return, they not only suffer financial loss, but also encounter challenges to their self-esteem. They can withdraw from the project to minimize their monetary loss, or they may keep throwing additional money into the project to demonstrate that their initial decision was correct. It is painful to admit a mistake because it poses negatively to the investors' self-concept. As a result, investors may be entrapped within a losing project and suffer accumulated financial loss. The present theory suggests that when investors encounter conflicts between money and self-esteem.

Keywords: Money, Self-esteem, Investment, Motivation, Decision Making.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Economics, Business and Finance*, edited by Klaus P. Hofmann, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence regarding this article should be addressed to Liqing Zhang, Department of Psychology, Peking University, 5 Yi-He-Yuan Avenue, Beijing, P. R. China, 100871. Email: liqingzhang@pku.edu.cn.

To some extent, we are all investors. We invest money and time in education. We devote much effort into specialized training before we start a career. The return that we expect to get is not merely to receive some regular paycheck and make a living, but also to validate our self-worth. Similarly, in financial investment, people not only expect to receive economic benefits, but also hope that their investments succeed and thus glorify their self-concepts.

# SELF-ESTEEM

Self-esteem refers to the subjective evaluation of the self. It is a fundamental motivation of human beings. Research in psychology has provided abundant empirical evidence that people want to defend, maintain, and enhance a positive self-image (Baumeister, 1998; Crocker and Park, 2004; Greenwald, 1980; James 1890; Steele, 1988; Taylor and Brown, 1988). People attribute success to their internal traits, but blame external factors for their failure (Miller and Ross, 1975). They derogate outgroup members in order to enhance their own group identity (Crocker and Luhtanen, 1990). They may compare themselves with others inferior to them so that they can enhance their self-esteem (Wills, 1981). In many situations, self-esteem plays an adaptive role in human functioning. Self-esteem fosters confidence, optimism and controllability, and thus it helps people to survive hardships (Taylor and Brown, 1988). For instance, cancer patients with high self-esteem tend to live longer after they are diagnosed with the disease than those with low self-esteem (Taylor and Brown, 1988). However, the pursuit of self-esteem may have costs. Attributing failure to external factors defends people's self-esteem, but it also precludes them from learning from their experiences. The belief that one's contribution is valuable may validate their self-worth, but it produces conflicts in resource allocation. For instance, in distributing the reward of a collaborative work, people tend to come up with different fair allocations, and the solution each side proposes usually favors their own position (Messick and Sentis, 1979; Van Avermaet, 1974). The present paper argues that economic loss may be a cost in pursuing selfesteem.

# **CONFLICT BETWEEN MONEY AND SELF-ESTEEM IN INVESTMENT**

The conflict between money and self-esteem becomes salient after people have chosen to invest in a project and the project produces negative economic return. When they decide to invest money into the project, they tie their pride to it, at least to a certain degree. When the project does not go on well, they can either withdraw or continue. Withdrawal from the project implies that the decision makers have made a bad decision. Therefore, instead of terminating a losing project, people may put additional money into the project to justify their original decision (Brockner and Rubin, 1985; Brockner, 1992; Staw, 1976, 1997). They may assume that the difficulty is temporary and there is still the prospect of making a profit. The motivation to defend their positive self-image thus produces costly over-investment in failing projects. In sum, people encounter conflict between money and self-esteem when they receive negative financial reports of their chosen project. They may defend their self-esteem at the expense of money.

1.

Research has shown that ego-threat may increase the tendency to incur monetary loss to defend people's self-esteem. Ego-threat refers to the situation in which people perceive threats to their self-esteem. Ego-threat is usually produced by negative information about the self, such as failure on an important task. People become more defensive after their self-esteem is challenged (Baumeister, 1998; Steele, 1988). They often make efforts to restore their positive self-image. If they encounter a decision dilemma of terminating or continuing a losing project, ego-threated people may be more likely to invest repeatedly so as to defend their self-esteem.

Recent empirical research supports the idea that people may sacrifice material payoffs in order to protect their self-esteem in entrapping situations. Ego-threat increases people's desire to restore self-esteem, and thus promotes their over-investment in a failing project. For instance, in Experiment 1 conducted by Zhang and Baumeister (2006), ego-threat was manipulated by suggesting that participants may have a negative personality-choking under pressure. In the ego-threat condition, participants were told, "If you're the kind of person who usually chokes under pressure or if you don't think that you have what it takes to win the money, then you might want to play it safe. But it's up to you." In the control condition, participants were not informed of any ego-threatening information. The entrapping situation was a gamble procedure that was adapted from Brockner and Rubin (1985). It represents chance-based situations, or other situations where people are not in control, such as the stock market, or waiting for a bus that does not come on time. In this game, participants were given \$5 dollars as deposit. They were told that they had the opportunity to win a jackpot of \$10. They were not informed of the exact probability of winning. This experiment was conducted using a computer program. A counter on the computer screen ran from 0 to 500. For every 25 numbers that the counter ran, participants invested in a quarter. When the counter reached a multiple of 25, such as 25, 50,---, the counter stopped and participants had to click on one of the buttons: continue or exit. Thus, each time participants decided to continue, they invested an additional quarter. They were told that they would win \$10 if a bell rang when the counter stopped. This game finished when participants decided to withdraw or when they used up all of their deposit (the counter reached 500). In fact, the program was set up so that no beep would be sound. Thus, each time they invested in some money, participants received negative economic return. The results demonstrated that participants in the ego-threat condition (M =3.67, SD = 1.59 invested a larger amount of money than those in the control condition (M =2.43, SD = 1.94). Thus, people who were eagerly motivated to restore their self-esteem invested and lost more money in the failing project than other people.

The motivation to maintain their self-esteem also impacts people's investment when their success depends on their ability and effort. In Experiment 2 conducted by Zhang and Baumeister (2006), ego-threat was manipulated by informing participants that they may have the negative personality of choking under pressure. The entrapping situation was a skill-related task, specifically solving a jigsaw-puzzle (Brockner and Rubin, 1985). Thus, participants might have thought that their skills and efforts could improve their probability of success. The puzzle was solvable, though the time limit--15 minutes made the task difficult. In fact, no participants solved it. Participants were given \$4 as deposit. The prize for solving the puzzle was \$10. Participants were told that solving a complicated puzzle under time pressure reflected one's capacity to perform well in challenging situations. They were first given 19 free puzzle pieces. They then had to purchase additional puzzle pieces required to

solve the puzzle. If they solve the puzzle within 15 minutes, they would gain a \$10 prize, and the money they used to purchase the puzzle pieces would be given back. There were no differences in the time that participants spent in solving the puzzle. Participants in the ego-threat condition (M = 2.12, SD = .91) invested more money in buying the puzzle pieces than those in the control condition (M = 1.47, SD = .77).

In the jigsaw-puzzle situation, people face the conflict between money and self-esteem after they have purchased some puzzle pieces and the puzzle is still not solved. The money that they have spent to buy the puzzle pieces is forgone. They can stop their investment in this task so that they minimize their financial loss. The findings suggest that they may rather continue investing in order to demonstrate that they are able to deal with the challenging situation. Ego-threat spurs the desire for self-esteem and increases the entrapment in losing endeavors.

It is conceivable that ego-threat influences choices in the potential investing situations because withdrawal poses negatively to one's self-concept (Staw and Ross, 1987). In Experiments 1 and 2 in Zhang and Baumeister (2006), the manipulation of ego-threat was related to the subsequent investment. It linked the decision of playing it safe to a personality weakness—choking under pressure. Therefore, it raised the question of whether ego-threat still increased the over-commitment to a losing project when the ego-threat was produced by a source irrelevant to the potential entrapping situation.

Empirical research shows that self-esteem influences people's investment choices, even when the ego-threat manipulation does not have any relevant connection with the subsequent decision making. Experiment 4 in Zhang and Baumeister (2006) was conducted to address this question. In it, ego-threat was manipulated by failure performance feedback on a creativity test (Baumeister, Tice and Heatherton, 1993). All participants completed a creativity test. In the ego-threat condition, participants received failure performance feedback. In the control condition, participants were given success performance feedback. The entrapping situation was the counter game as discussed earlier in this paper. The results showed that participants in the ego-threat condition (M = 4.04, SD = 1.34) invested more money than those in the control condition (M = 2.40, SD = 1.94). Therefore, ego-threat increased entrapment in losing projects even when the source of the ego-threat was irrelevant to the entrapping situations.

The motivation to defend self-esteem also contributes to the costly entrapment in interpersonal competition. For instance, in Experiment 3 conducted by Zhang and Baumeister (2006), ego-threat was produced by failure performance on a creativity test. The procedure that represented interpersonal competition was called dollar auction (Brockner and Rubin, 1985; Shubik, 1971; Teger, 1980). In this game, participants competed with another person in buying a dollar. A special feature of the game was that the person who paid the second highest price also had to pay their bid, although they did not get the dollar. In this experiment, participants were run in groups of four people. They were told that they were randomly paired with another person to bid for one dollar. They were given \$5 as deposit. They were told that they could decide how much to bid. The rule of the game was that the person who bided the highest paid their bid, and attained the dollar. The person who had the second highest bid was also charged for their bid, even though this person did not gain anything in return. In the opening round of the auction, each person wrote down their own bid. Starting from the second round, participants were told how much their opponent had bided and they decided whether to bid more or not. The game

ended when one person decided to withdraw from the auction. The highest bid that one could offer was their total deposit of five dollars. In fact, each time the participant offered a bid, the experimenter told them that their opponent had bided more. The results showed that participants in the ego-threat condition (M = 3.71, SD = 1.58) bided more than those in the control condition (M = 2.46, SD = 1.89).

In the dollar auction, participants may first want to put in a little bit of money to gain one dollar. As the auction went on, their motivation changed from gaining a small amount of money to winning the game (Teger, 1980). When they bided more than one dollar, it was obvious that the motivation to win, instead of earning money, was the driving force of their decisions. By additionally analyzing the data of Experiment 3 in Zhang and Baumeister (2006), participants' highest bids were divided into two categories: below (or equal to) one dollar, and over one dollar. The results of the Chi Square analysis showed that participants in the ego-threat condition (30 out of 35) were more likely to bid more than one dollar than those in the control condition (21 out of 34),  $Chi^2$  (1, N = 69) = 5.13, p < .02.

Thus, when people encounter conflicts between money and self-esteem in interpersonal competition, their reluctance to withdraw may produce financial loss for both parties. This theory provides a perspective in understanding a variety of destructive behaviors in interpersonal and intergroup conflicts, such as price wars, strikes, and costly law suits in divorce (Teger, 1980).

In the counter game and the jigsaw-puzzle game, it is conceivable that people want to put additional money into the losing project so that they can change the losing situation into a profitable one by winning the jackpot or solving the puzzle. However, in the dollar auction, it is apparent that no one is able to turn the losing situation into a winning one if people bid more than a dollar in exchange for one dollar. It is obvious that people increase their bid in the dollar auction because they do not want to be defeated by others.

Maximizing monetary payoffs is usually considered the bench mark for rational choices. The present theory indicates that people may consider factors other than money during decision making, and specifically, people may strive for self-esteem at the expense of economic payoffs. It should be noted that the pursuit of self-esteem at the expense of money may help people defend their pride in the short-run, but it may eventually produce much larger costs of both money and self-esteem (Baumeister, Heatherton, and Tice, 1993; Fox and Staw, 1979; Ross and Staw, 1986, 1993). For example, British Columbia's decision to host a world fair (Expo 86) in Vancouver ended up in a costly entrapping situation (Ross and Staw, 1986). The initial budget in 1978 was \$78 million, but the cost eventually accumulated to \$1.5 billion with a deficit of \$311 million. The Provincial Premier William Bennett originally decided to hold the fair and tied his pride to it. He declined the fair director's suggestion to cancel it despite the negative financial reports about the project. Therefore, giving up monetary payoffs may not be a best strategy to defend one's self-esteem.

# **GENERAL DISCUSSION**

Research has demonstrated that people throw good money away into losing projects (Bazerman, Giuliano, and Appelman, 1984; Brockner and Rubin, 1985; Brockner, 1992; Garland, 1990; Staw, 1976, 1997; Teger, 1980). Entrepreneurs over-invest in projects that

have produced negative financial return (Antonides,1995; McCarthy, Schoorman and Cooper, 1993). People fail to change their unpromising career choices (Drummond and Chell, 2001). The present paper argues that people may forgo financial benefits in chasing their positive self-image.

### Implications to Other Theories on Entrapment

The present idea is consistent with the self-justification theory on entrapment (Brockner, 1992; Brockner, Houser, Birnbaum, etc, 1986; Brockner and Rubin, 1985; Staw, 1976, 1979, 1981, 1997). Self-justification theory deems that people put more money into a previously chosen course of action because they want to justify that their initial decision of investment is wise. The notion that self-esteem is the motivation underlying self-justification was once suggested by Staw (1981). Empirical research also suggests that people are more likely to be entrapped into a losing project when the project is diagnostic of an important ability than when the task is irrelevant to any central attribute (Brockner and Rubin, 1985). Decision makers may perceive that their self-esteem is more at stake when the success of the project signifies their ability than when the failure of the task does not pose any threat to their competence. Research also demonstrates that people put in a larger second investment when they themselves decided the initial investment than when someone else chose the project (Schoorman and Holahan, 1996; Staw, 1976). Apparently, people's self-esteem was more at stake when they themselves made the original choice. To some extent, the present theory advances the self-justification theory by demonstrating self-esteem as the underlying motivation of self-justification.

Furthermore, the present theory suggests that motivational theories are able to provide new predictions that are not easily explained by competing cognitive theories. Even though self-justification theory is an influential theory on entrapment, alternative cognitive theories have been proposed (for review, see Brockner, 1992; Wilson and Zhang, 1997). Specifically, prospect theory (Kahneman and Tversky, 1979; Whyte, 1986) and decision dilemma (Bowen, 1987) theory provide two competing alternatives. According to prospect theory, decision makers prefer to take risks when they encounter a choice between accepting a sure loss and gambling to turn the situation around. After they have invested a small amount of money and receive negative financial return, decision makers have to accept the sure loss of the previous investment if they withdraw. Due to loss aversion, they would rather put additional money into the project so that they can have the opportunity (though small) to turn the situation around. It has been difficult to differentiate the driving forces between prospect theory and self-justification theory because the predictions of both theories are usually similar. Some researchers even proposes that prospect theory is able to account for all the findings that support self-justification theory and can also explain a broader range of phenomena (e.g., Whyte, 1986). Even though prospect theory offers an important alternative explanation, in particular for situations that do not involve the motivation to maintain an integrative selfconcept (e.g., Arkes, and Blumer, 1985), self-esteem contributes to costly entrapment. Prospect theory cannot explain the findings reviewed in this paper (Zhang and Baumeister, 2006). Risk seeking in loss domains cannot interpret why people become more locked into losing endeavors after their self-esteem is threatened.

Decision dilemma theory (Bowen, 1987) suggests that ambiguity may account for the over-investment in losing situations. Since the possibility of success is uncertain or unknown, people may think that they may eventually succeed by investing repeatedly.

Even though ambiguity may influence entrapment in some situations, the decision dilemma theory can not explain why challenges to self-esteem increase investors' risk taking. Additionally, in the dollar auction, there is clearly no possibility to turn bad into good after people pay more than one dollar for a dollar, but people, especially investors who have received ego-threat and thus have greater desire for self-esteem, bid much more than one dollar. Therefore, the present theory advances the motivational viewpoint of entrapment.

#### Indications for De-Escalation

The present theory suggests that the interpretation of the situation may influence investors' choices when their previously chosen project has produced economic loss. If they view withdrawal positively, investors may more likely stop investing in a failing project than if they regard withdrawal as a personal weakness (Brockner and Zubin, 1985; Simonson and Staw, 1992; Staw and Ross, 1987). If people do not view withdrawal as an indication of failure, they may be more likely to stop an unprofitable project. Re-interpretation of the situation reduces withdrawal's negative implication to the investors' self-concept and thus contriutes to de-escalation.

The present theory suggests that self-affirmation may de-escalate commitment to a losing project. Self-affirmation refers to the verification of one's positive self-image. Research has documented that self-affirmation reduces defensiveness (Steele, 1988). Future research should test empirically whether people become less entrapped in a failing project after their self-esteem has been assured.

# CONCLUSION

The present theory proposes that people may encounter conflicts between money and self-esteem in investment. The conflict between money and self-esteem becomes salient when people have invested in a project and the project produces monetary loss. People thus face the tough choice: withdrawal or continuing. Withdrawal may minimize financial loss, but it may also pose negatively to investors' self-esteem. Continual gives investors a small possibility to turn the losing situation into a winning one, but it may also more likely produce a large amount of accumulated loss. Furthermore, continuing may indicate that investors' initial decisions were wise and thus protect their self-esteem. The present paper reviews evidence showing that people are entrapped more deeply in losing endeavors when their self-esteem is threatened and thus they are more motivated to restore their positive self-image. Therefore, people may sacrifice monetary payoffs in pursuing self-esteem.

# REFERENCES

- Antonides, G. (1995). Entrapment in risky investments. *The Journal of Socio-Economics*, 24, 447-461.
- Arkes, H. R., and Blumer, C. (1985). The psychology of sunk cost. Organizational Behavior and Human Decision Processes, 35, 124-140.
- Baumeister, R. F., Dale, K., and Sommer, K. L. (1998). Freudian defense mechanisms and empirical findings in modern social psychology: Reaction formation, projection, displacement, undoing, isolation, sublimation, and denial. *Journal of Personality*, 66(6), 1081-1124.
- Baumeister, R. F., Heatherton, T., and Tice, D. M. (1993). When ego threats lead to selfregulation failure: Negative consequences of high self-esteem. *Journal of Personality and Social Psychology*, 64, 141-156.
- Bazerman, M. H., Beekun, R. I., and Schoorman, F. D. (1982). Performance evaluation in a dynamic context: A laboratory study of the impact of a prior commitment to the ratee. *Journal of Applied Psychology*, 67, 873-876.
- Bowen, M. G. (1987). The escalation phenomenon reconsidered: Decision dilemmas or decision errors? Academy of Management Review, 12, 52-66.
- Brockner, J. (1992). The escalation of commitment to a failing course of action: Toward theoretical progress. *Academy of Management Review*, 17, 39-61.
- Brockner, J., Houser, R., Birnbaum, G., Lloyd, K., Deitcher, J., Nathanson, S., and Zubin, J. Z. (1986). Escalation of commitment to an ineffective course of action: The effect of feedback having negative implications for self-identity. *Administrative Science Quarterly*, 31, 109-126.
- Brockner, J., and Rubin, J. Z. (1985). Entrapment in Escalating Conflicts: A Social Psychological Analysis. New York: Springer-Verlag.
- Crocker, J., and Park, L.E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin*, 130(3), 392-414.
- Crocker, J., and Luhtanen, R. (1990). Collective self-esteem and ingroup bias. *Journal of Personality and Social Psychology*, 58, 60-67.
- Drummond, H., and Chell, E. (2001). Life's chances and choices: A study of entrapment in career decisions with reference to Becker's side bets theory. *Personnel Review*, 30, 186-202.
- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, *35*, 603-618.
- Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, 47, 263-291.
- James, W. (1890). *The principle of psychology (Vol.1)*. Cambridge, MA: Harvard University Press.
- McCarthy, A. M., Schoorman, F. D., and Cooper, A. C. (1993). Reinvestment decisions by entrepreneurs: Rational decision-making or escalation of commitment? *Journal of Business Venturing*, 8, 9-24.
- Messick, D. M., Sentis, K.P. (1979). Fairness and preference. *Journal of Experimental Social Psychology*, 15, 418-434.

- Miller, D. T., and Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, 82, 213-225.
- Ross, J., and Staw, B. M. (1986). Expo 86: An escalation prototype. *Administrative Science Quarterly*, *31*, 274-297.
- Schoorman, F. D., and Holahan, P. J. (1996). Psychological antecedents of escalation behavior: Effects of choice, responsibility, and decision consequences. *Journal of Applied Psychology*, 81, 786-794.
- Shubik, M. (1971). The dollar auction game: A paradox in noncooperative behavior and escalation. *Journal of Conflict Resolution*, 15, 109-111.
- Simonson, I., and Staw, B. M. (1992). Deescalation strategies: A comparison of techniques for reducing commitment to losing course of action. *Journal of Applied Psychology*, 77, 419-426.
- Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to chosen course of action. *Organizational Behavior and Human Performance*, 16, 27-44.
- Staw, B. M. (1979). Rationality and justification in organizational life. In B. M. Staw and L. L. Cummings (Eds.) *Research in organizational behavior. (Vol. 2)*. Greenwald Connecticut: JAI Press.
- Staw, B. M. (1981). The escalation of commitment to a course of action. Academy of Management Review, 6, 577-587.
- Staw, B. M. (1997). The escalation of commitment: An update and appraisal. In Z. Shapira (Ed.) Organizational decision making (pp191-215). New York, NY: Cambridge University Press.
- Staw, B. M., and Ross, J. (1987). Behavior in escalation situations: antecedents, prototypes, and solutions. *Research in Organizational Behavior*, *9*, 39-78.
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261– 302). New York: Academic Press.
- Taylor, S. E., and Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193-210.
- Teger, A. I. (1980). Too Much Invested to Quit. Elmsford, New York: Pergamon Press.
- Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1, 39-60.
- Van Avermaet, E. (1974). Equity: A theoretical and experimental analysis. *Doctoral Dissertation*, University of California, Santa Barbara.
- Whyte, G. (1986). Escalating commitment to a course of action: A reinterpretation. *Academy* of Management Review, 11, 311-321.
- Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90, 245-271.
- Wilson, R. M. S., and Zhang, Q. (1997). Entrapment and escalating commitment in investment decision making: A review. *British Accounting Review*, 29, 277-305.
- Zhang, L., and Baumeister, R. F. (2006). Your money or your self-esteem: Threatened egotism promotes costly entrapment in losing endeavors. *Personality and Social Psychology Bulletin*, 32, 881-893.

Chapter 32

# **RISK MITIGATION: INDIVIDUAL AND MARKET BEHAVIOR**\*

Jamie Brown Kruse<sup>1</sup> and Kevin M. Simmons<sup>2</sup>

<sup>1</sup> Wind Science and Engineering Research Center, Texas Tech University, Lubbock, TX, USA and Center for Natural Hazards Mitigation, East Carolina University, Greenville, NC, USA
<sup>2</sup> Department of Business and Economics, Austin College, Sherman, TX, USA

# **INTRODUCTION**

Wind research has focused mainly on the meterological aspects of storms and the engineering challenges to construct resistant structures. Over the last several years, however, an effort has been launched to understand the behavior of people living in areas threatened by violent windstorms. The aftermath of a significant windstorm event brings instant media attention to the affected area and government aid is usually forthcoming to lessen the financial blow to residents and businesses. It is rare, but some storms produce damage so severe that it alters the economic conditions, not only of those living and working in the path of the storm but third party insurers as well. Hurricane Andrew is a relatively recent example. As losses from these events mount, several questions continue to arise. What can be done to lessen the effect of a tornado or hurricane? And second, will people adopt mitigation measures that may be available to them?

The second question has been of interest to economists and there is some debate on the potential response of people living in vulnerable areas. These events are rare and it is possible that people simply ignore the risk. If this tendency prevails, then storm resistant construction will not be adopted due to the increased cost. Essentially, the safety offered by the storm resistant engineered structures is a form of insurance and the value placed on that insurance

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

may or may not correspond to the perceived risk of an event. This paper uses experimental economics techniques to examine the human response to changes in risk and how this response aggregates to become market demand for risk mitigation measures.

We will examine the relationship between individual bidding behavior and the aggregative properties of a market for full insurance. We utilize a Vickrey auction to sell from one to six policies. A Vickrey auction is a sealed-bid auction in which participants each simultaneously submit bids. The auctioneer discloses the identity of the highest bidder(s) who is declared the winner. The price paid by the winning bidders is equal to the highest bid that does not win the object(s). This format is named after William Vickrey who first described it and pointed out that bidders have a dominant strategy to bid their true values. Participants in the laboratory experiment reported here either acquire one of the available insurance policies in the auction or face exposure to the risk of a two-dollar loss. There is a rich literature in Psychology and Economics that has contributed to our current understanding about decision under risk. We describe a few of the important papers in the next section. This is followed by our experimental design, results and conclusions.

# **REVIEW OF RELATED LITERATURE**

An early lab experiment performed by Slovic, Fischhoff, Lichtenstein, and Corrigan (1977) tested the attractiveness of insurance to consumers at various probability levels. The subjects were given an amount of money, which was put at risk through the drawing of a ball from an urn filled with red and blue balls. If a blue ball was drawn the subject lost his/her money. A red ball would signal that there would be no penalty. Full insurance was available for purchase at a fixed premium equal to the expected value of the loss. The results of this experiment indicated that as the probability of loss increased, the percentage of subjects purchasing the insurance increased. At low probabilities, only about 20% of the subjects purchased the insurance. When the probability of loss increased to over 0.2, the percentage of subjects purchasing insurance increased to over 80%. A second treatment studied was the level of the premium. It was found that as the premium was lowered, through a subsidy, the percentages purchasing insurance at all probability levels increased. Finally, the authors tested the effect of compound lotteries. A compound lottery has nested probabilities. The subjects were told that the probability of loss on each draw was the same, for instance 5%, but the number of draws increased thus increasing their overall probability of loss. An interesting finding resulted from this manipulation of the experiment. At a probability level of 0.01, little change was observed in the percentage purchasing insurance. However, beginning at the .05 level, almost all subjects purchased the insurance. The general conclusion of the study was that as probabilities increased either directly or through a compounding effect, consumer's willingness to purchase insurance also increased.

Shoemaker (1979) reports the results of an experiment that explored several issues regarding insurance. First, the experiment was designed to measure the level of aversion to a high-probability-low-loss event versus a low-probability-high-loss event. Second, the willingness to pay for varying premium/deductible combinations was elicited and finally, the willingness to pay for comprehensive insurance was measured. Subjects for the experiment came from two sources. One group was undergraduate students at the University of

Pennsylvania. The other group was solicited by mail from the client list of a Philadelphia insurance agency.

To test differences concerning risk aversion, the subjects were asked to choose between a 60% chance of losing \$100 or a 1% chance of losing \$6,000. Losses to the subjects were hypothetical in contrast to the first experiment, which used real money losses. The expected value of both choices is the same. The majority of students chose the low probability, high loss option. The majority of insurance clients, mostly adults, chose the high probability, low loss option. A possible explanation for this difference, offered by the authors, was that the adults had a larger accumulated wealth. This increased wealth influenced their decision in that they had more to lose than young college students.

The second group of questions tested various premium/cost combinations of insurance policies. It is interesting to note that for both groups, there was a clear preference for very low deductible and very high deductible policies. The least preferred policies were those in between the two extremes. One possible explanation for this result is the inability to process choices rationally with two variables, premiums and deductibles.

Regarding the willingness to pay for comprehensive insurance, the subjects were asked the amount they were willing to pay for insurance to cover each of three losses. They were then asked how much they would be willing to pay for a policy to cover all losses. Of the student group, 49% indicated that they would pay less for the comprehensive policy. Similarly, 53% of the client group also said they would pay less.

A final set of questions tests preferences for policies where the only difference is in scale. Two choices are offered to the subjects: a premium of \$100 to mitigate a 25% chance of a \$500 loss or a premium of \$1 to mitigate a 25% chance of losing \$5. The expected value of the two choices is the same but 49% of the clients and 47% of the student subjects preferred the \$1 premium. Only 13% of the students and 7% of the clients indicated consistent choices.

More recently, an experiment performed by McClelland, Schulze, and Coursey (1993) used a Vickrey auction to allocate four insurance policies that protect the subject from a negative result in a lottery. Similar to the urn experiments of Slovic et al., the subjects could earn monetary awards based on the results of the lottery. This differed from the Slovic et al. (1977) study in that the price of the insurance policy could vary whereas in the earlier study the price was fixed. Once the four policies were allocated to the highest bidders, the lottery was held. If a red chip was drawn, all subjects without a policy lost \$4. The draw of a white chip resulted in a gain of \$1 for all subjects. Six levels of probabilities were used ranging from .01 to .9. Ten rounds for each probability level were conducted.

A bid to EV ratio was calculated for each individual in each round. These ratios were then compared at the different levels of loss probability. One result of the study was that mean Bid/EV declined as the probability of loss increased. Further, in low probabilities (less than 0.10) a bi-modal distribution was observed with a zero bid forming one node and Bid to EV ratios greater than 1 forming the other extreme. At low probabilities, subjects either ignored the risk and bid zero, or focused predominately on the loss and submitted a bid in excess of the expected value of the loss. As probabilities increased the distribution of Bid/EV began to take the shape of a normal distribution.

A second experiment was used to test differences in bids as the subjects gained experience with the lottery. Fifty rounds, all at a probability of 0.01, were conducted. On round 33, a red chip was deliberately drawn from the jar. The authors observed that as the subjects gained experience with the process, the bids began to drop. Subsequent to the

drawing of a red chip in round 33, bids began to increase. Decreased bids prior to round 33 were attributed to the subjects lowering their perception of the risk until a loss occurred. The increase in the bids after the loss was attributed to a fear that once a loss has occurred, it seems more likely to happen again.

Ganderton, Thurston, Brookshire, Stewart, and McKee (1997) used lab experiments to test consumers' willingness to purchase insurance protection from a disaster such as a flood. The authors created three possibilities in the lottery. One possibility was no event, the second was a periodic flood and the third was an episodic flood that results in severe damage. Further, in many disasters, some homes are destroyed while neighboring structures suffered little damage. To model this possibility, once the nature of the event was determined by the lottery, a second drawing was held to determine if the damage was slight or severe.

Prior to each treatment, the subjects were told the possibilities for each event, potential losses and the cost of the insurance to protect them from the losses. A total of eighteen treatments were tested with different combinations of probabilities, losses, and cost of insurance. The authors found that as the level of expected loss increased, subjects purchased more insurance. Also, at a given level of loss, insurance was more likely to be purchased as the cost of the policy decreased.

One of the benefits of experimental research is the ability to identify individual behavior that determines the characteristics of a market. Market data aggregates the many decisions made by market participants. It is quite possible, that individual decisions may violate rational decision rules while the market does not. A study by Evans (1997) constructs a set of experiments where participants bid on a set of choices with the same expected value. If their choices exhibit inconsistency the author calculates the deviation of the individual bid from the expected value. These deviations are violations of a form of rationality known as the "betweenness property" of the *independence axiom* of Expected Utility Theory. The study concluded that betweenness violations were high for individuals.

Evans (1997) examined market behavior using a fifth price Vickrey auction. A fifth price auction is a device, which assigns the fifth highest bid as the market price in an auction. In the market setting, while betweenness violations continued to exist, the magnitude of the violations dropped dramatically. One of the reasons, suggested by the author, for the improved market behavior was the mechanism to set the market price. A fifth price auction selects a price from the middle of the bid distribution. This ignores the tails of the distribution that yield most violations. Evans states that it may be that research on individual bidding behavior is less relevant to market behavior than previously thought. Market mechanisms, which target the middle of the bid distribution, perform well. By this logic, other mechanisms which choose extreme bids, for instance a first price auction, may not perform as well.

Camerer (1987) examines the role of probability bias in setting market prices. In this experiment, participants were allowed to bid on lottery tickets that paid a state dependent dividend. The expected value of the dividend could be determined by multiplying the payoff by the state probability. Bids for the tickets were compared to the expected values to determine individual bias. Several competing theories were tested to determine which theory did the best job of predicting market price. The first theory uses Bayes Rule to determine the probability. The probability meeded to correctly assess the state probability in the initial drawing. Competing theories assume participants either overestimate or underestimate the probability in the initial drawing.

Results indicated that while individual bias was prevalent, generally market prices converged to Bayesian expected values. This convergence occurred more quickly with experienced subjects than with inexperienced subjects. As each type of subject participated in additional rounds, bids more closely tracked expected values. This result suggests that participants learn from the market.

## METHODOLOGY

The experiment we report was divided into five sessions with 8 to 12 participants in each session. Subjects were recruited from undergraduate and graduate courses at Texas Tech University as well as the local community. A total of 50 volunteers participated in the three sessions. There were 17 women and 34 men who participated in the experiments. All subjects were offered the same incentives. The length of each session ranged from  $1\frac{1}{2}$  to 2 hours. Payment per subject ranged from \$18 to \$25.

Each session began with volunteers signing an alphabetical participant list for the session. Volunteers were then given a copy of the Consent Form to read and sign. When all subjects had arrived and signed the consent form they were assigned a seat and given an envelope containing the experiment instructions, bid forms and their individual record sheet. The instructions were explained and a time for questions regarding those instructions was allowed prior to the session.

Each session contained 11 rounds, 1 practice and 10 actual rounds. After the practice round, each participant received \$2.00 at the beginning of each round. Before they could claim the money, a lottery was held in which 1 ball was drawn from a bingo cage containing 100 balls numbered sequentially 1-100. Every participant was assigned a range of numbers on his/her record. If the ball drawn was within the assigned range, the subject must forfeit the \$2.00. The probability for loss was the same for all participants in each round, however, the range of numbers was different for each individual. Five different loss probabilities were used in the experiment, 1%, 5%, 10%, 15%, and 20%. These probabilities were not explicitly given on the record sheets, but could be calculated by examining the range of numbers assigned. Each probability level was used in one of the first five rounds and the last five rounds. The sequence of probabilities in each half of the experiment was randomly determined prior to the experiment.

Prior to the lottery, all participants were given an opportunity to purchase a license (an insurance policy), which would protect their \$2.00, regardless of the outcome of the lottery. A Vickrey auction mechanism was used to allocate 1, 2, 3, 4, 5, or 6 licenses. Bids were posted from highest to lowest. All bids were public knowledge but the identity of the bidder was not. After the bids were posted, a six-sided die was rolled to determine the number of licenses available in that round. With the protocol we used, subjects did not know the number of available licenses at the time they submitted their bids. The license price was the bid whose rank was one more than the number of available licenses in that round. For example, if the roll of the die was 3, three licenses were sold to the highest bidders. The price of the license was equal to the fourth bid. In the event of a tie, tied bids were randomly queued for license allocation purposes.

Once the number, price and allocation of licenses were determined, the bid forms were returned with an indication of whether the participant had purchased a license. A ball was then drawn from the bingo cage, its number announced, displayed, and then returned to the bingo cage. For those with a license, the earnings for that round were \$2.00 minus the price of the license. All other participants were exposed to the results of the lottery and thus had earnings of \$2.00 or zero depending on the draw from the bingo cage. At the conclusion of the round, the participants were asked to total their earnings, update their accumulated balance and prepare their bid form for the next round. After all ten rounds were complete, the participant earnings were verified and paid in cash.

# RESULTS

In a previous study, McClelland et al. (1996) demonstrated that Bid/EV at low probabilities exhibited a bi-modal distribution. One node was at zero Bid/EV and the second node was a Bid/EV ratio well in excess of 1. Our study corroborates this finding. As is shown in Figure 1, there is a large cluster of bids at zero, few bids at middle ranges and a second cluster at Bid/EV in excess of 5. In fact at a probability of 1%, three quarters of the bids were zero. This confirms the observation made by McClelland et al. (1996) that at very low levels of probability, subjects either ignore or exaggerate the risk. The effect of subjects who exaggerate the risk is particularly noteworthy at these very low levels of probability. Since the expected value of the loss is minimal, it takes a small increase in the bid to increase dramatically the Bid/EV ratio at low probabilities. This fact is illustrated in Figure 2. The average Bid/EV at a probability of 1% is almost 3 while the percentage of zero bids was 75%. Some of the participants submitted bids in which the Bid/EV ratio was over 40. There are actually two forces at work here. First, as noted in McClelland et al. (1996), subjects who value mitigation at low probabilities exaggerate the risk. A second reason for the high Bid/EV ratio is the fact that the Bid/EV measure at low probability levels is extremely sensitive to very small changes in the bids. As the probability increases, the bid distribution begins to take a more normal shape. Again, two reasons account for this. First, the number of subjects bidding zero declines and second, the sensitivity of the Bid/EV ratio to small changes in the bids is lower. As Figure 2 indicates, beginning at a 5% probability, average Bid/EV tracks closer to one. In fact, at 20% probability, average Bid/EV is not significantly different from 1 [t-statistic = 1.35589 assuming unequal variance. P(T<=t) = .08911 (one tail test)].

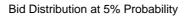
To address the question of whether the subject bidding behavior responds to increases in probability we performed a regression with bid as the dependent variable and probability as the independent variable.

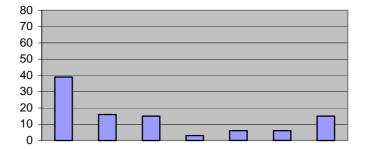
Bid =  $\beta_0 + \beta_1 Prob + \beta_{2-51} Subject + Error$ 

Even though our database contains 500 bid decisions by the experiment participants, the sample contains only 50 individuals. To control for the small number of participants we used a fixed effects regression model. This model controls for possible error correlation within the observations for a given individual by adding a dummy variable for each individual in the experiment. As can be seen in Table 1 the regression is significant. The intercept is very close

to zero, although it is not significant. Probability is highly significant. The 2.196 coefficient of Probability indicates a Bid/EV ratio of 1.098 over the entire distribution, since EV = \$2.00 x Probability. This corresponds to the observation seen in the graph in Figure 2. Clearly, an increase in probability motivates subjects to increase their bid for the license.

Bid Distribution at 1% Probability





Bid Distribution at 10% Probability

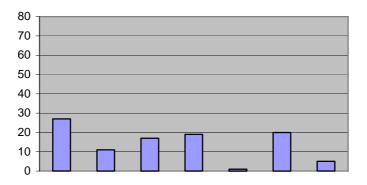
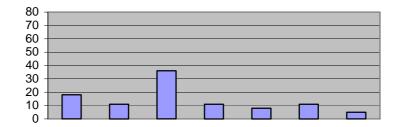


Figure 1. Continued on next page.

Bid Distribution at 15% Probability



Bid Distribution at 20% Probability

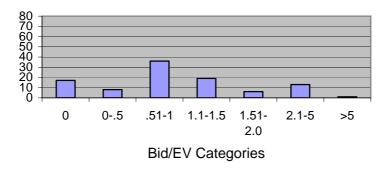


Figure 1.



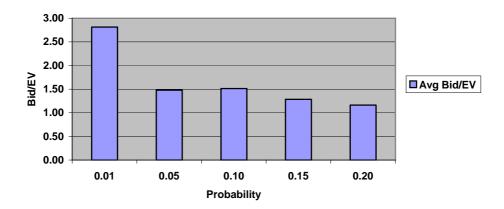


Figure 2.

Since the zero bids are prominent in this analysis, we plotted the changes in the percentage of subjects bidding zero at increasing probabilities. McClelland et al. (1996) noted that the number of zero bids declined as probability increased but never disappeared entirely.

This observation is consistent with the results of our study. As shown in Figure 3, zero bids at 1% probability are 78% of the total bids. The percentage of zero bids decreases as probability increases but never completely disappears. Small increases in subjects bids translate to large increases in Bid/EV. A similar effect is created by the zero bids. Since negative bids were not an option the bid distribution has a floor, which at low probabilities creates a distribution that is asymmetrical. As probability increases, the effect of the floor diminishes.

## **Table 1. Regression Results**

	<b>Parameter Estimate</b>	Std Error	T Value	Prob>T	R-Sq	Adj.R-Sq	
Bid/Probability Model (All Participants)							
Intercept	0.0470	.0999	0.475	.6353	.5003	.4446	
Probability	2.1961	.2016	10.892	.0001			
Bid/Hit Model (All Participants)							
Intercept	0.0752	.1018	0.739	.4606	.4818	.4220	
Probability	2.3550	.2438	9.661	.0001			
Bid/Hit Model (Participants which Experienced a Hit)							
Intercept	0.0860	.1056	0.815	.4157	.5016	.4448	
Probability	2.2206	.2029	10.944	.0001			
EV = 2.00 X							

.01 = .02

### Proportion of Zero Bids to Total Participants

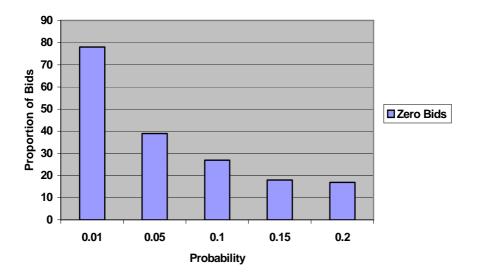
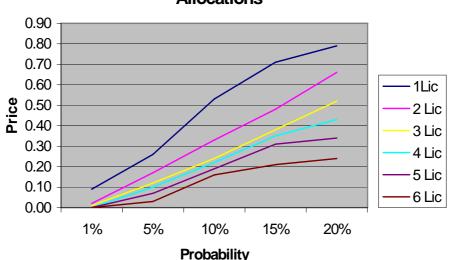


Figure 3.

This experiment used a Vickrey auction to allocate licenses. A unique feature of the study was to withhold information about the allocation of the licenses. Subjects did not know the number of available licenses until their bids were submitted. As a result, we can examine the behavior of prices at each probability for six possible allocations. Figure 4 shows the behavior of the prevailing price, averaged across sessions, for each of six allocation possibilities. As

expected, prevailing price declines as allocation increases. This models the effect of the scarcity of the license. When only 1 license is allocated, price begins and remains high. As the probability increases, price rises, although for 1 license the relationship is more erratic than when 2, 3, 4, or 5 licenses are available. When six licenses are available, the price increase across probabilities begins to lose some stability. This is an interesting result and it reflects the effect of bids on price at the upper and lower end of the bid distribution. As noted by Evans (1997), individual bids do not always track expected value. However, the market price is remarkably consistent. Her study used a fifth price auction to determine market price. The design of our study allows for analysis of six different prices and confirms her result at the middle of the bid distribution. As allocation of licenses increases, the price function approaches a linear function. But when allocation exceeds 4 licenses, the instability seen in the 1 license allocation reappears. At these high levels of available licenses, price is largely determined by the low and sometimes zero bids. To examine this effect further, we performed regressions on each allocation using price as the dependent variable and probability as the independent variable. (Table 2) These regressions contain few observations, so only the general behavior of the relationship between price and probability can be examined. Each of the six regressions indicated that probability was highly significant and the value of the intercept was very close to zero. The difference seen in the graph, however, is reaffirmed by differences in the value of the adjusted R-Square, for each regression. R-Square for the 1 license and 6 license allocation was .31 and .32. The middle allocations, 2 through 5 licenses, had adjusted R-Square values ranging from .41 to .48. These middle allocations minimize the effect of the high and low bidders on price and indicate that market price responds well to increases in the probability of a loss.



# Average Market Price for different License Allocations

Figure 4.

	Parameter	Std. T		Prob>T	R-Sq	Adj.
	Estimate	Error	Value	1100 1	R Sq	R-Sq
One License	Allocation					
Intercept	0.0866	.0974	0.889	.3786	.3244	.3103
Probability	3.8176	.7951	4.801	.0001		
Two License	Allocation					
Intercept	-0.0079	.0682	-0.117	.9075	.4261	.4140
Probability	3.3232	.5567	5.96	.0001		
Three Licen	se Allocation					
Intercept	-0.0186	.0541	-0.310	.7577	.4313	.4194
Probability	2.6626	.4413	6.033	.0001		
Four License	e Allocation					
Intercept	-0.0186	.0446	-0.416	.6790	.4596	.4483
Probability	2.3231	.3636	6.389	.0001		
<b>Five License</b>	Allocation					
Intercept	-0.0167	.0346	-0.482	.6321	.4933	.4827
Probability	1.9302	.2824	6.836	.0001		
Six License A	Allocation					
Intercept	-0.0109	.0342	-0.320	.7502	.3370	.3232
Probability	1.3801	.2794	4.939	.0001		
<b>Total Regres</b>	ssion					
Intercept	0.0026	.0274	0.095	.9241	.3083	.3060
Probability	2.5728	.2232	11.525	.0001		

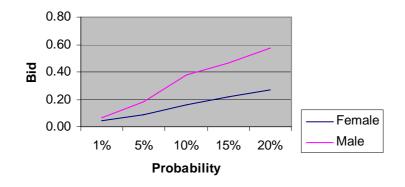
 Table 2. Regression Results: Market Price/Probability Model

 for Different License Allocations

While the main focus of this study was the response of subjects to increases in the probability of loss, two side issues were examined. First, does experiencing a "hit" from the lottery have an effect on subsequent bidding behavior? McClelland et al. (1996) devised a scheme to test this effect by contriving a hit to all subjects on the 33<sup>rd</sup> round of a 50-round experiment. Their observation was that bids increased after the loss. Our study tests the effect of a loss, but the hit is allowed to occur naturally from the lottery itself. For all subjects, we create a variable whose value is zero if no loss occurs in the previous round and has a value of 1 if a loss does occur in the previous round. We also created a variable that had a value of 1 in every round subsequent to a loss. This variable is analogous to the method used by McClelland et al. (1996). A graph comparing the Bid/EV ratio for subjects experiencing a hit and those who did not is shown in Figure 5. We used a fixed effects regression model similar to the one described earlier to test the effect of a hit on bidding behavior.

Bid =  $\beta_0 + \beta_1 Prob + \beta_2 Hit + \beta_{3-53} Subject + Error$ 

The first model regressed bid against hit for all participants in the study, regardless of whether they experienced a hit. Secondly, we regressed bid against hit for only the participants who experienced a hit. As can be seen in Table 1, HIT is not a significant explanatory variable for bidding behavior in either model. We therefore conclude, that experience with a hit in the lottery has no effect on subsequent bidding behavior.



## **Gender Comparison - Average Bid**

Figure 5.

A second issue was the effect of gender on bidding behavior. We assigned a value to each subject indicating whether the subject was male or female. It was then possible to compare the bids by gender. Figure 6 shows average bids by women and by men at increasing probability levels. At every probability, bids by women were lower than the bids submitted by men. To test the strength of this observation, we ran a regression where bid was the dependent variable and the independent variables were probability and whether or not the subject was female.

 $Bid = \beta_0 + \beta_1 Prob + \beta_2 Gender + Error$ 

The variable for female was highly significant and the sign of the coefficient was negative. We also conducted a t-test on the bids of men and women. As Figure 6 shows, the bids submitted by men generally increased as probability increased. However, the increase was large between 5% and 10%, declined at 15% and then increased again at a probability of 20%. Compared to the pattern for women, these increases were more erratic. Some of this difference can be seen in the variances for bids from women and bids for men. The variance for women's bids was under \$0.03 whereas the variance for men's bids exceeded \$0.23. Further, men submitted bids that were consistent with higher risk aversion (Figure 7). This result is consistent with other experimental results and gives an instance where men appear to be more conservative toward risk than women. Our conclusion is that there is a significant difference in the bidding behavior between men and women.

# Comparison of Bid/EV by Hit and No Hit

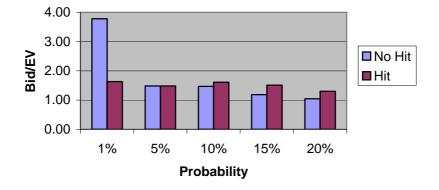


Figure 6.



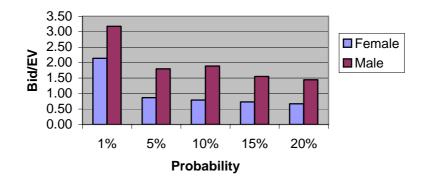


Figure 7.

# CONCLUSION

This study generated three important results. First, individual bids and consequently market price are affected by probability of loss. As shown in the McClelland et al., subjects do respond to increasing loss probabilities in their willingness to buy protection. This behavior, at low probabilities, indicates a bimodal response to risk with most subjects ignoring the risk and a minority over-valuing the risk. At higher probabilities, the distribution appears to take a shape similar to a normal distribution. Further, bid/EV ratios correspond closely to the expected value of the loss.

Second, in contrast to the McClelland et al. study, we find that experience with a hit in the lottery has no discernible effect on subsequent bidding behavior. This may be due to the contrast in the way that a loss was implemented in the two studies. Although the probabilities for an individual were the same for the two studies, in McClelland et al., when a loss occurred, it hit everyone whereas a hit was isolated to only one or two individuals in our study. This could be likened to when a single property is damaged by a tornado versus an entire community that could be damaged by a hurricane. Finally, we find a difference in the bidding behavior of men and women. Men tend to bid higher than women indicating that men value the protection of the license more highly than women do.

Extreme wind events such as hurricanes and tornadoes are low-probability-highconsequence events. The annual probability that a tornado will touch down at a given location is less than 1% even in the states in "tornado alley." The annual probability that a hurricane force storm will strike a location is 10% or less even in hurricane-prone regions. The results of this study give us insight on mitigation investment behavior when the probability of loss is low. Admittedly, a two dollar loss in the laboratory is not a high consequence event. However, the preponderance of zero bids at low probabilities match up with the syndrome the Kunreuther calls "it can't happen to me." Likewise, those individuals that overweight the probability are likely to be the first to adopt mitigation measures that otherwise fail cost effectiveness tests. Individuals that overweight the probability are likely to unnecessarily evacuate when a hurricane threatens and contribute to congestion on evacuation routes. Public policy that takes into account individual biases will be more effective in managing mitigation efforts and protecting the population when a disaster occurs.

### ACKNOWLEDGMENTS

This work was performed under the Department of Commerce NIST/TTU Cooperative Agreement Award 70NANB8H0059.

## REFERENCES

- Camerer, Colin F. (1987). "Do Bias in Probability Judgment Matter in Markets? Experimental Evidence," *American Economic Review*, Vol. 77, No. 5, pp. 981-977.
- Evans, Darla A.(1997). "The Role of Markets in Reducing Expected utility Violations," *Journal of Political Economy*, Vol. 105, No. 3, pp. 622-636.
- Ganderton, P., Thurston, H., Brookshire, D., Stewart, S., and McKee, M. (1997). "Buying Insurance for Disaster-Type Risks: Experimental Evidence," Manuscript. Albuquerque: University of New Mexico.
- McClelland, Gary H., Schulze, William D., and Coursey, Don L. (1993). "InsuranceFor Low-Probability Hazards: A Bi-Modal Response to Unlikely Events," *Journal of Risk and Uncertainty*, Vol. 7, pp. 95-116.
- Schoemaker, Paul J. H., and Kunreuther, Howard (1979). "An Experimental Study of Insurance Decisions," *The Journal of Risk and Insurance*, Vol. 46, No. 4, pp. 603-618.

Slovic, P., Fischhoff, B., and Lichtenstein, S. (1977). "Behavioral Decision Theory," *Annual Review of Psychology*, Vol. 28, pp. 1-39.

Chapter 33

# FREQUENCY-DOMAIN VERSUS TIME-DOMAIN ESTIMATES OF RISK AVERSION FROM THE C-CAPM: THE CASE OF LATIN AMERICAN EMERGING MARKETS<sup>\*</sup>

# Ekaterini Panopoulou<sup>†</sup>

University of Piraeus, Piraeus, Greece and Trinity College Dublin, College Green, Dublin

# ABSTRACT

Campbell (2003) confirms the equity premium puzzle in an international context based on the Consumption-CAPM and cross-country evidence on implausibly large coefficients of relative risk aversion. In this paper we adopt a spectral approach to reestimate the values of risk aversion over the frequency domain for six Latin American emerging markets. We complement our analysis with the traditional time series approach and confirm the results of existing literature of large coefficients of relative risk aversion. Our frequency domain findings, however, indicate that at lower frequencies risk aversion falls substantially across countries, thus yielding in many cases reasonable values of the implied coefficient of risk aversion.

**Keywords:** Equity premium puzzle; Consumption-CAPM; risk aversion; frequency domain; emerging markets.

**JEL Classification:** G10; C13

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: Ekaterini Panopoulou, Department of Statistics and Insurance Science, University of Piraeus, 80 Karaoli & Dimitriou str., 18534, Piraeus, Greece. Email: apano@unipi.gr. Tel: 0030 210 4142728. Fax: 0030 210 4142340.

# **1. INTRODUCTION**

Economists accept as a stylized fact that average stock returns are substantially higher than average returns on short-term debt instruments. The failure of financial theory to explain the magnitude of these excess returns has led to this phenomenon being labeled as the "equity premium puzzle" by Mehra and Prescott (1985). Standard asset pricing models can only match the data if investors are extremely risk averse. In particular, the coefficient of relative risk aversion must be implausibly large for traditional models to reconcile the large differential between real equity returns and real returns available on short-term debt instruments.<sup>1</sup> Of course, we expect different financial assets to deliver large variations in returns, but typically financial economists have explained such differentials by attributing them to differences among the covariances of asset returns and investors consumption, e.g. the Consumption Capital Asset Pricing Model (C-CAPM) of Lucas (1978) and Breeden (1979). The more traditional version of CAPM assumes a perfect correlation between the stock market return and the consumption path of the typical investor. This allows us to measure asset risk as its covariance with the market return. However, in their path breaking work, Mehra and Prescott (1985), using annual US data from 1889 to 1978, showed that the covariance of equity returns with consumption growth was insufficient to explain the observed equity premium of over 6%. In fact, they could only account for a premium of approximately 0.35%.

Much of the resulting empirical literature has focused on the US markets where longer data series exist, but Campbell (1996, 2003) focuses on some smaller stock markets and finds evidence that an equity premium is also a feature of these markets. Specifically, Cambpell (2003) reports evidence from 11 countries that imply extremely high values of risk aversion, which usually exceed many times the value of 10 considered plausible by Mehra and Prescott (1985), and claims "...that the equity premium puzzle is a robust phenomenon in international data". Using the same theoretical setup as in Campbell (2003) and a dataset from six Latin America emerging markets, we adopt both a time series and a spectral approach to re-estimate the values of risk aversion over the frequency domain.<sup>2</sup>

Our approach allows for long-term consumption dynamics by performing a dynamic analysis of the link between consumption and returns at several frequencies rather than over the time domain. As pointed out by Granger and Hatanaka back in 1964, according to the spectral representation theorem a time series can be seen as the sum of waves of different periodicity and, hence, there is no reason to believe that economic variables should present the same lead/lag cross-correlation at all frequencies. We incorporate this rationale into the context of the single-factor C-CAPM by using well-developed techniques to estimate the coherency (the analog of the correlation coefficient in the time domain) between returns and consumption over the frequency domain along with the related spectra of the series (the analog of variance).<sup>3</sup> In this way, we can separate different layers of dynamic behavior of consumption and returns by distinguishing between the short run (fluctuations of 2 to 6 quarters), the medium run or business cycle (lasting from 8 to 32 quarters), and the long run

<sup>&</sup>lt;sup>1</sup> Kocherlakota (1996) provides an excellent survey of the topic.

<sup>&</sup>lt;sup>2</sup> Some studies have investigated the implications of spectral analysis within economic applications, mostly by interpreting (high) low-frequency estimates as the (short) long-run component of the relationship under scrutiny; see, for instance, Engle (1974, 1978).

<sup>&</sup>lt;sup>3</sup> See Hamilton (1994) for a general overview of spectral analysis.

(oscillations of duration above 32 quarters). Our findings indicate that at lower frequencies risk aversion falls substantially across countries, thus yielding in many cases reasonable values of the implied coefficient of risk aversion.

The paper is structured as follows. Section 2 presents the theoretical model, while our frequency domain technique is developed in Section 3. Section 4 discusses our data and its properties, while Section 5 presents the empirical results. Concluding remarks are contained in Section 6.

# 2. MEASURING RISK AVERSION AND THE EQUITY PREMIUM

The equity premium puzzle can be presented in different ways. We adopt the approach of Campbell (1996, 2003) which follows the seminal papers of Lucas (1978), Mehra and Prescott (1985) on the equity premium puzzle. Specifically, we assume that there is a representative agent who maximizes a time-separable utility function:

$$MaxE_t \sum_{j=0}^{\infty} \delta^j U(C_{t+j})$$
<sup>(1)</sup>

where  $\delta$  is the discount factor,  $C_{t+j}$  is the investor's future consumption stream and  $u(C_{t+j})$  is the period utility derived from such consumption. This problem yields the following Euler equation to describe the optimal consumption and investment path of the investor;

$$U'(C_t) = \delta E_t[(1 + R_{i,t+1})U'(C_{t+1})]$$
(2)

with  $1 + R_{i,t+1}$  representing the gross rate of return available on asset *i*. The investor equates the loss in current consumption with the expected gain in discounted consumption next period.

Consistent with other studies, we employ a time-separable power utility function:

$$Max \sum_{j=0}^{\infty} \delta^{j} \frac{C_{l+j}^{1-\gamma}}{1-\gamma}$$
(3)

where  $\gamma$  is the coefficient of relative risk aversion. The features of this utility function are well known and have had their validity questioned in the literature.

Specifically, this utility allows risk premia to remain stable over time even if both aggregate wealth and the economy increase. In this respect, aggregation of different investors with different wealth levels but the same power utility function into a single representative investor can be performed. On the downside, however, the tight link between the elasticity of intertemporal substitution which is equal to the reciprocal of the coefficient of relative risk

aversion is preserved. Epstein and Zin (1991) and Weil (1989) introduced a more general utility function that breaks the aforementioned link.

Despite some shortcomings of the power utility specification, we retain this specification in this study for two reasons. Firstly, it facilitates comparison with other studies. Secondly, Kocherlakota (1996) reports that modifications to preferences such as those proposed by Epstein and Zin (1991), habit formation due to Constantinides (1990) or "keeping up with the Joneses" as proposed by Abel (1990) fail to resolve the puzzle.

Combining equations (2) and (3), we get the familiar expression,

$$1 = E_t \left[ (1 + R_{i,t+1}) \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \right].$$
(4)

Following Hansen and Singleton (1983) and Campbell (1996), we assume that the joint conditional distribution of asset returns and consumption is lognormal, with time-varying volatility. Taking logs of equation (4), we get

$$0 = E_t r_{i,t+1} + \log \delta - \gamma E_t [\Delta c_{t+1}] + 0.5 (\sigma_{it}^2 + \gamma^2 \sigma_{ct}^2 - 2\gamma \sigma_{i,ct})$$
(5)

where  $r_{i,t} = \log(1 + R_{i,t})$ ,  $c_t = \log(C_t)$ ,  $\sigma_{it}^2$  and  $\sigma_{ct}^2$  denote the conditional variance of log returns and log consumption growth respectively, and  $\sigma_{i,ct}$  represents their conditional covariance. The log risk premium is

$$E_t \left[ r_{i,t+1} - r_{f,t+1} \right] + \frac{\sigma_{it}^2}{2} = \gamma \sigma_{ict} \tag{6}$$

where the variance term on the left-hand side of equation (6) is Jensen's Inequality adjustment term due to using expectations of log returns. Therefore the log risk premium is a function of the coefficient of relative risk aversion multiplied by the covariance of stock returns with consumption growth. In this respect, riskier assets get a larger risk premium as these assets are the ones that have a high covariance with consumption and produce low returns when consumption is low.

Letting then  $e_{i,t+1} \equiv E_t[r_{i,t+1} - r_{f,t+1}]$  denote the excess return over the risk free rate, we get that the excess return on any asset over the riskless rate is constant and therefore the risk premium on all assets is linear in expected consumption growth with the slope coefficient,  $\gamma$ , given by:

$$\gamma = \frac{e_{i,t+1} + 0.5\sigma_i^2}{\sigma_{i,c}} \tag{7}$$

Equation (7) provides a working formula for the estimation of the coefficient of relative risk aversion by simply calculating the sample moments of excess returns and consumption growth. This type of calculation is refereed to as the 'time-domain' estimate of relative risk aversion.

# **3. ECONOMETRIC METHODOLOGY**

In this section, we present an alternative econometric methodology employed to study the behavior of the equity premium and the coefficient of relative risk aversion over the frequency domain.

Departing from the time domain to the frequency domain we can rewrite (7) for each frequency.<sup>4</sup> After dropping the time subscript for simplicity, we get that the coefficient of risk aversion over the whole band of frequencies  $\omega$ , where  $\omega$  is a real variable in the range  $0 \le \omega \le \pi$ , is given by:

$$\gamma_{\omega} = \frac{e + 0.5 f_{ee}(\omega)}{f_{ec}(\omega)} \tag{8}$$

where e denotes the excess log return of the stock market over the risk-free rate. As is well known, the cross-spectrum,  $f_{ec}(\omega)$ , between e and c is complex-valued and can be decomposed into its real and imaginary components, given here by:

$$f_{ec}(\omega) = C_{ec}(\omega) - iQ_{ec}(\omega),$$

where  $C_{ec}(\omega)$  denotes the *co-spectrum* and  $Q_{ec}(\omega)$  the *quadrature spectrum*. The measure of comovement between returns and consumption over the frequency domain is then given by:

negative definite matrix, i.e.  $F_{yx}(\omega) = F_{yx}^*(\omega)$ , with \* denoting the complex conjugate transpose since  $f_{yx}(\omega) = \overline{f_{xy}(\omega)}$ . See Hamilton (1994) for a more detailed presentation of spectral analysis.

<sup>&</sup>lt;sup>4</sup> In general, the spectrum of a process, say  $x_t$ , can be written as  $f_{xx}(\omega) = \rho_0 + 2\sum_{k=1}^{\infty} \rho_k \cos(k\omega)$ , where  $\rho_k$  is the k-order autocovariance function of the series. In turn, we can consider the multivariate spectrum,  $F_{yx}(\omega)$ , for a bivariate zero mean covariance stationary process  $Z_t = [y_t, x_t]^T$  with covariance matrix  $\Gamma(\cdot)$ , which is the frequency domain analog of the autocovariance matrix. The diagonal elements of

 $F_{yx}(\omega)$  are the spectra of the individual processes,  $f_{yy}(\omega)$  and  $f_{xx}(\omega)$ , while the off-diagonal ones refer to the cross-spectrum or cross spectral density matrix of  $y_t$  and  $x_t$ . In detail,  $F_{yx}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \Gamma(k) e^{-ik\omega} = \begin{bmatrix} f_{xx}(\omega) & f_{yx}(\omega) \\ f_{xy}(\omega) & f_{yy}(\omega) \end{bmatrix}$ , where  $F_{yx}(\omega)$  is an Hermitian, non-

$$c_{ec}^{2}(\omega) \equiv \frac{\left|f_{ec}(\omega)\right|^{2}}{f_{ex}(\omega)f_{cc}(\omega)} = \frac{C_{ee}^{2} + Q_{cc}^{2}}{f_{ee}(\omega)f_{cc}(\omega)}$$
(9)

where  $0 \le c_{ec}^2(\omega) \le 1$  is the squared *coherency*, which provides a measure of the correlation between the two series at each frequency and can be interpreted intuitively as the frequency-domain analog of the correlation coefficient.<sup>5</sup>

The spectra and co-spectra of a vector of time-series for a sample of T observations can be estimated for a set of frequencies  $\omega_n = 2\pi n/T$ , n = 1, 2, ..., T/2. The relevant quantities are estimated through the periodogram, which is based on a representation of the observed time-series as a superposition of sinusoidal waves of various frequencies; a frequency of  $\pi$  corresponds to a time period of two quarters, while a zero frequency corresponds to infinity.<sup>6</sup>

# 4. TIME-DOMAIN AND FREQUENCY DOMAIN PROPERTIES OF THE DATA

#### 4.1. Data

Our dataset comprises quarterly equity and macroeconomic data for the Latin American emerging markets, namely Argentina, Brasil, Chile, Colombia, Mexico and Peru. The source for equity data is the Morgan Stanley Capital International (MSCI) indices adjusted for dividends for emerging markets. The indices are denominated in local currency and the domestic money market rate is employed. In this respect, we calculate the equity premium that a domestic investor faces in contrast to calculating returns in a foreign currency, for example in US dollars and adjusting them for the US risk free rate.<sup>7</sup> Data on population, household consumption, GDP deflator, and a short-term interest rate, mainly a 3-month T-bill rate are available at the IFS Statistics Database (International Monetary Fund). Real personal consumption per capita was calculated by dividing personal consumption with the population and adjusting it for price changes utilizing the GDP deflator.<sup>8</sup> Additionally, seasonal adjustment of the consumption series was also undertaken for the cases that consumption data were unadjusted. Our dataset is the longest available for each country at hand. Specifically, the samples employed are as follows: Argentina (1988:Q1- 2005:Q4), Brasil (1988:Q1-

<sup>&</sup>lt;sup>5</sup> Engle (1976) gives an early treatment on the frequency-domain analysis and its time-domain counterpart.

<sup>&</sup>lt;sup>6</sup> Consistent estimates of the spectral matrix can be obtained by either smoothing the periodogram, or by employing a lag window approach that both weighs and limits the autocovariances and cross-covariances used. We use here the Bartlett window, which assigns linearly decreasing weights to the autocovariances and crosscovariances in the neighborhood of the frequencies considered and zero weight thereafter, with the lag, k, set using the rule  $k = 2\sqrt{T}$ , as suggested by Chatfield (1989).

<sup>&</sup>lt;sup>7</sup> The approach of calculating US denominated excess returns would signal the equity premium faced by an international investor investing in the emerging markets. Such an approach is taken by Salomons and Grootveld (2003).

<sup>&</sup>lt;sup>8</sup> Population data are reported annually. The series were converted to quarterly by means of linear interpolation.

2005:Q4), Chile (1988:Q1- 2005:Q4), Colombia (1993:Q1-2005:Q4), Mexico (1988:Q1-2005:Q4) and Peru (1993:Q1-2005:Q4).

### 4.2. Descriptive Statistics

Table 1 reports summary statistics for stock market excess returns and consumption growth. Excess returns are calculated by subtracting the return on the relevant short-term risk-free asset from the return in the stock market index and consumption growth is the difference in the logarithm of seasonally adjusted personal consumption per capita GDP deflated. For each country we report the mean and standard deviation in annualized percentage points of the series and countries under scrutiny.<sup>9</sup> The table shows that Latin American countries share the same stylized facts with developed countries (Campbell, 2003).

First, Latin American stock markets have delivered average returns of 7% or more in excess of the risk free rate. Specifically, Chile, Colombia and Peru just exceed 7%, whereas the excess return in Mexico is 17%. These figures are roughly equivalent to the figures for the developed markets examined by Campbell (2003). In his dataset, excess returns range between 3% and 14%. Quite interestingly, the respective figures for Argentina and Brasil reach the impressive 170% and 146% for the period under examination.

Second, the annualized volatility of stock returns ranges from 25% to 35% for Chile, Colombia, Mexico and Peru. As expected, the higher Argentinean and Brasilian excess returns are followed by higher risk as depicted in their volatility that reaches 461%. Comparing the volatility of these emerging markets with the developed ones, we have to note that the volatility of the former is on the upper part of the latter which experience a volatility of 15% to 27%.

Country	Excess return (mean)	Excess return (standard deviatio	1	Consumption Growth (standard deviation)
Argentina	170.44	461.01	7.50	13.31
Brasil	146.45	131.78	18.04	21.40
Chile	7.75	24.99	14.46	4.48
Colombia	7.22	35.25	14.85	3.51
Mexico	16.60	29.58	22.09	9.42
Peru	7.06	29.30	12.63	5.97

#### **Table 1. Descriptive Statistics**

Notes: Means and standard deviations are given in annualized percentage points. Means are multiplied by 400 and standard deviations by 200.

Third, consumption growth is rather elevated and depicts higher volatility compared to developed countries. This is expected since developing countries normally experience faster

<sup>&</sup>lt;sup>9</sup> To annualize the raw quarterly data, means are multiplied by 400 and standard deviations by 200 given that in serially uncorrelated data standard deviations increase with the square root of the interval.

growth that undergoes larger swings. The fastest growth in consumption is for Mexico (22%), followed by Brasil (18%), while the lowest one is the Argentinean one with annualized growth of almost 8%. In any case these figures are 4 to ten times greater than the growth in developed countries. The least volatile economy in our dataset is Colombia, while the ones with the greater volatility are Argentina and Brasil.

# **5. EMPIRICAL RESULTS**

#### 5.1. Time-domain Techniques

We follow Campbell (2003) and present two measures of risk aversion. The first one termed RRA1 is calculated directly from (7), while the second one, denoted by RRA2 assumes a unitary correlation of excess returns and consumption growth. Although this is a counterfactual exercise, we follow closely Campbell (2003) and postulate a unitary elasticity between returns and consumption growth to account for the sensitivity of the implied risk aversion on the smoothness of consumption rather than its low correlation with excess returns. Table 2 reports the average equity premium, the unconditional correlation between excess returns and consumption along with the two estimates of relative risk aversion for the countries at hand.

In detail, the second column reports the mean equity premium given by the left part of equation  $(6)^{10}$  (i.e. adjusted with the respective variance term).<sup>11</sup> Quite interestingly, there is a great dispersion among the values of the equity premia for the Latin American countries. The higher one is the one of the Argentinean market that exceeds 1200% per, followed by Brasil which is over 200% with the lowest one being the one that corresponds to Peru of just over 10%. The huge numbers for Argentina and Brasil combine two effects, both the increased mean of average excess returns and the respective increased volatility. The correlation between the returns and consumption also exhibits great variation among countries. The lowest is reported in Argentina and Mexico (4%) and the highest one in Brasil (51%). The lower correlations should normally correspond to higher estimates of risk aversion while the higher ones to the lower estimates. This is true as column 3 (Table 2) suggests that reports these estimates.

Specifically, this traditional approach yields the usual finding (see also Campbell 1996, 2003). The estimated coefficients of relative risk aversion of 502 for Argentina, 175 for Mexico and 81 for Peru are much larger than suggested by economic theory.<sup>12</sup> This is often interpreted as an equity premium puzzle. While still outside the economically accepted range, the estimates of 36 and 16 for Chile and Brasil, respectively certainly move the relative risk parameter in the right direction.<sup>13</sup>

<sup>&</sup>lt;sup>10</sup> This approach has been advocated by Jobson and Korkie (1981), amongst others, in relation to the application of Modern Portfolio Theory.

<sup>&</sup>lt;sup>11</sup> The equity premium is given in annualized percentage points. The calculations, however, are done in natural units.

<sup>&</sup>lt;sup>12</sup> Mehra and Prescott (1985) impose an upper limit of 10 on 'reasonable' values of the coefficient of RRA. However, others such as Kandel and Stambaugh (1991) argue that this parameter could be much higher (up to 30) in financial markets and still imply reasonable economic behaviour.

<sup>&</sup>lt;sup>13</sup> The negative correlation between excess returns and consumption for Colombia does not allow us to calculate the RRA coefficient.

We repeat the analysis for the hypothetical situation where the correlation between stock returns and consumption growth is unity. This case gives the model the best chance of explaining the premium. Indeed, the estimated coefficients are reduced. With the exception of Argentina (RRA= 20) and Colombia (RRA=11), the remaining estimates of risk aversion are all within the economically acceptable values. What it shows is that the extreme risk aversion displayed in the actual data is driven mainly by an almost complete dis-connect between stock returns and consumption.

Country	Equity premium (me	an) Correlation	RRA1	RRA2
Argentina	1233.11	0.04	502.01	20.10
Brasil	233.29	0.51	16.12	8.27
Chile	10.87	0.27	35.51	9.71
Colombia	13.43	<0		10.87
Mexico	20.97	0.04	175.48	7.53
Peru	11.36	0.08	81.43	6.49

**Table 2. Time Domain Estimates of Relative Risk Aversion** 

Notes: Equity premium is given in annualized percentage points. RRA1 is calculated from (7) and RRA2 by imposing a unitary correlation between excess returns and consumption growth.

#### 5.2. Frequency-domain Techniques

Before moving on with the estimates of relative risk aversion, we report some evidence on the comovement between returns and consumption growth in the frequency domain along with the estimated spectra of the series at hand. Tables 3A-3F report the coherency and the variance decomposition of excess returns and consumption growth over the frequency domain (columns 3-5). Column 1 states the respective frequency as a fraction of  $\pi$ , while column 2 reports the time-domain analogue of the frequency in quarters. Zero frequency corresponds to an infinite horizon, while a frequency of  $\pi$  corresponds to a 2-quarter horizon.

Frequency	(Quarters)	Coherency $f_{ec}(\omega)$	Returns Variance f <sub>ee</sub> (w)	Consumption Variance $f_{cc}(\omega)$	RRA1	RRA2
0	(inf)	0.995	1.663	0.610	1.25	1.25
1/16	(32.000)	0.987	1.437	0.476	1.39	1.38
1/8	(16.000)	0.940	1.067	0.233	1.99	1.92
3/16	(10.667)	0.776	0.895	0.070	3.95	3.48
1/4	(8.000)	0.448	0.834	0.025	8.74	5.85
3/8	(5.333)	0.792	0.733	0.056	4.39	3.90
1/2	(4.000)	0.936	0.706	0.144	2.52	2.44
5/8	(3.200)	0.981	0.794	0.245	1.88	1.87
3/4	(2.667)	0.986	0.925	0.290	1.73	1.71

 Table 3A. Frequency Domain Estimates of Relative Risk Aversion (Argentina)

Frequency	(Quarters)	Coherency $f_{ec}(\omega)$	Returns Variance $f_{ee}(\omega)$	Consumption Variance $f_{cc}(\omega)$	RRA1	RRA2
13/16	(2.462)	0.976	1.007	0.243	1.90	1.88
7/8	(2.286)	0.953	0.938	0.139	2.53	2.47
15/16	(2.133)	0.876	0.755	0.053	4.28	4.01
1	(2.000)	0.651	0.640	0.023	7.67	6.19

#### Table 3A. Continued

Notes: Frequency is expressed as a fraction of  $\pi$ . RRA1 is calculated from (8) and RRA2 by imposing a unitary correlation between excess returns and consumption growth.

Frequency	(Quarters)	Coherency	Returns Variance	Consumption Variance	RRA1	RRA2
0	(inf)	0.989	0.634	1.294	0.76	0.75
1/16	(32.000)	0.969	0.413	0.852	0.98	0.97
1/8	(16.000)	0.826	0.150	0.318	2.22	2.02
3/16	(10.667)	0.606	0.085	0.170	4.37	3.40
1/4	(8.000)	0.740	0.067	0.149	4.65	4.00
3/8	(5.333)	0.799	0.043	0.164	5.14	4.59
1/2	(4.000)	0.646	0.030	0.200	6.14	4.93
5/8	(3.200)	0.637	0.035	0.404	4.05	3.23
3/4	(2.667)	0.602	0.032	0.591	3.60	2.79
13/16	(2.462)	0.209	0.027	0.434	7.72	3.53
7/8	(2.286)	0.587	0.044	0.220	5.17	3.96
15/16	(2.133)	0.548	0.051	0.146	6.16	4.56
1	(2.000)	0.270	0.025	0.121	13.34	6.93

# Table 3B. Frequency Domain Estimates of Relative Risk Aversion (Brasil)

Notes: See Table 3A.

Frequency	(Quarters)	Coherency	Returns Variance	Consumption Variance	RRA1	RRA2
0	(inf)	0.976	0.003	0.012	3.51	3.47
1/16	(32.000)	0.943	0.001	0.000	45.37	44.07
1/8	(16.000)	0.993	0.003	0.000	48.73	48.54
3/16	(10.667)	0.593	0.006	0.000	85.63	65.96
1/4	(8.000)	0.354	0.008	0.000	185.63	110.38
3/8	(5.333)	0.864	0.001	0.000	218.04	202.72
1/2	(4.000)	0.964	0.003	0.000	119.81	117.66
5/8	(3.200)	0.846	0.002	0.000	215.19	197.92
3/4	(2.667)	0.943	0.004	0.000	77.07	74.86

Frequency	(Quarters)	Coherency	Returns	Consumption	RRA1	RRA2
			Variance	Variance		
13/16	(2.462)	0.668	0.006	0.000	179.88	146.98
7/8	(2.286)	0.987	0.003	0.000	177.33	176.20
15/16	(2.133)	0.868	0.001	0.000	411.44	383.23
1	(2.000)	0.331	0.000	0.000	1696.00	975.16

### Table 3C. Continued

Notes: See Table 3A.

# Table 3D. Frequency Domain Estimates of Relative Risk Aversion (Colombia)

Frequency	(Quarters)	Coherency	Returns	Consumption	RRA1	RRA2
			Variance	Variance		
0	(inf)	0.093	0.010	0.001	22.34	6.81
1/16	(32.000)	0.061	0.009	0.001	30.71	7.61
1/8	(16.000)	0.035	0.008	0.001	51.13	9.53
3/16	(10.667)	0.059	0.007	0.000	55.82	13.60
1/4	(8.000)	0.170	0.006	0.000	62.59	25.79
3/8	(5.333)	0.304	0.005	0.000	83.65	46.10
1/2	(4.000)	0.371	0.004	0.000	103.01	62.70
5/8	(3.200)	0.268	0.003	0.000	155.69	80.63
3/4	(2.667)	0.178	0.003	0.000	244.71	103.19
13/16	(2.462)	0.119	0.002	0.000	356.00	123.03
7/8	(2.286)	0.089	0.002	0.000	441.08	131.25
15/16	(2.133)	0.039	0.003	0.000	643.67	127.02
1	(2.000)	0.053	0.004	0.000	531.75	122.52

Notes: See Table 3A.

# Table 3E. Frequency Domain Estimates of Relative Risk Aversion (Mexico)

Frequency	(Quarters)	Coherency	Returns	Consumption	RRA1	RRA2
			Variance	Variance		
0	(inf)	0.674	0.006	0.006	8.60	7.06
1/16	(32.000)	0.618	0.005	0.004	12.90	10.15
1/8	(16.000)	0.389	0.003	0.001	36.69	22.88
3/16	(10.667)	0.188	0.004	0.001	60.99	26.43
1/4	(8.000)	0.345	0.005	0.001	45.56	26.76
3/8	(5.333)	0.504	0.003	0.000	56.87	40.36
1/2	(4.000)	0.588	0.002	0.000	93.57	71.75
5/8	(3.200)	0.477	0.002	0.000	107.61	74.33
3/4	(2.667)	0.242	0.006	0.000	114.33	56.23
13/16	(2.462)	0.233	0.008	0.000	108.09	52.20
7/8	(2.286)	0.030	0.006	0.000	334.33	58.33
15/16	(2.133)	0.310	0.004	0.000	116.19	64.68
1	(2.000)	0.285	0.003	0.000	136.24	72.68

Notes: See Table 3A.

Frequency	(Quarters)	Coherency	Returns Variance	Consumption Variance	RRA1	RRA2
0	(inf)	0.324	0.003	0.001	15.51	8.83
•	(inf)					
1/16	(32.000)	0.320	0.004	0.001	16.51	9.34
1/8	(16.000)	0.258	0.003	0.001	22.68	11.53
3/16	(10.667)	0.174	0.003	0.000	38.08	15.89
1/4	(8.000)	0.116	0.003	0.000	72.84	24.80
3/8	(5.333)	0.150	0.002	0.000	102.77	39.79
1/2	(4.000)	0.156	0.003	0.000	120.48	47.57
5/8	(3.200)	0.141	0.003	0.000	137.56	51.66
3/4	(2.667)	0.126	0.004	0.000	152.96	54.22
13/16	(2.462)	0.126	0.004	0.000	159.15	56.48
7/8	(2.286)	0.090	0.004	0.000	207.61	62.26
15/16	(2.133)	0.081	0.004	0.000	237.79	67.71
1	(2.000)	0.112	0.004	0.000	216.75	72.47

Table 3F. Frequency Domain Estimates of Relative Risk Aversion (Peru)

Notes: See Table 3A.

Overall our estimates suggest that the correlation (measured by coherency) between returns and consumption growth exhibits an upward trend as we move from high to low frequencies with the exception of Colombia. More in detail, coherency in the long-run is around 99% for Argentina, Brasil and Chile, 67% for Mexico, 32% for Peru and just 9% for Colombia. The Colombian co-spectrum is rather inverse exhibiting higher correlation in the short-run rather than the long-run. On the whole, the short-run correlation (2 to 6 quarters) between returns and consumption growth ranges from 13% in Peru to 90% in Argentina, while the business-cycle correlation (8 to 32 quarters) fluctuates between 8% in Colombia to 78% in Argentina and Brasil.<sup>14</sup>

As aforementioned the spectra of the series under scrutiny (reported in columns 4-5 of Tables 3A-3F) can be interpreted as the variance decompositions over various frequency bands (stated as a fraction of  $\pi$ ). As can be readily observed, the variability of returns and consumption growth exhibit substantial changes over the frequency domain. Specifically, the variability of consumption is generally muted for 2 to 16 quarters; however, for horizons exceeding 16 quarters a steep increase is prevalent. As the time horizon approaches infinity, the variance of consumption is thirty times greater than its short-run value in Brasil for example.<sup>15</sup> Similar patterns are observed in all the countries under scrutiny. Turning to the spectra of returns, we have to note that in general the variance of returns increase as we move to the end of spectrum (long-run) but these increases are milder compared with the ones prevailing the consumption growth.

These findings have direct implications for the subsequent analysis, i.e. the estimation of the coefficient of risk aversion, since the variance of consumption growth and returns are inversely related to the coefficient of risk aversion and coherence is positively related to

<sup>&</sup>lt;sup>14</sup> The respective figures are calculated by averaging their frequency domain counterparts over the relevant frequencies.

<sup>&</sup>lt;sup>15</sup> The concentration of variance in low frequencies is an indication of short-term correlation in consumption growth, such as an AR(1) with a positive coefficient, rather than an indication of non-stationarity of the process, which can be ruled out for the series at hand.

relative risk aversion. Consequently, we expect that as the lower frequencies are taken into account, risk aversion will decrease.

Column 6 (Tables 3A-3F) reports the coefficient of relative risk aversion calculated from (8) (RRA1) while column 7 reports the respective coefficient after imposing a unitary correlation coefficient between returns and consumption (RRA2). Starting with Argentina, we have to note that both RRA1 and RRA2 are sufficiently reduced even at high frequencies. Specifically RRA1 reduces to 1.25 in the long run from 7.67 in the short run. Marginal reductions prevail when RRA2 is considered. This stems from the high coherency between returns and consumption growth over roughly the entire spectrum.

Turning to Brasil, a similar pattern is observed with RRA1 decreasing from 13.34 to 0.76 over the spectrum, while the respective figures are 6.93 and 0.75 for RRA2. Our results for Chile are more impressive as the short-term RRA1 is estimated at the enormous value of 1696 in the short-run and falls to 3.51 in the infinite horizon with similar reductions prevailing RRA2 as well. With respect to Colombia, the low correlation between consumption and returns in the short-run leads to high relative risk aversion of 531.75, whereas in the low frequency a somewhat increased coherency and increased variability in returns and consumption leads to a reduced estimate of 22.34. These gains are further amplified when the unitary coefficient is imposed and as such RRA2 reduces to 6.81. Our results for Mexico and Peru paint a similar picture suggesting that risk aversion at high frequencies is found to be extremely large while it considerably reduces at the low frequencies with estimates at 8.60 and 15.51 respectively. This picture continues to hold under the assumption of a unitary elasticity between excess returns and consumption growth that leads to further reductions to 7.06 and 8.83, respectively.

On the whole, the estimates of relative risk aversion improve substantially at low frequencies and range from 0.76 (Argentina) to 22.34 (Colombia). When a unitary correlation coefficient is imposed, these estimates are slightly reduced for all the countries at hand and range from 0.75 to 8.83. This improvement in the estimates of relative risk aversion for the lowest frequency is driven by the spectral properties of the data at hand. As we move to lower frequencies, the variability of consumption growth increases significantly matching the variability of excess returns and as such the covariance of returns and consumption increases. This property is coupled for most of the countries at hand with a rise in the estimated coherency (i.e. the correlation in the frequency domain) between consumption and returns.

# **6.** CONCLUSION

We re-investigate the presence of an equity premium puzzle in six Latin America emerging stock markets. We attempt to re-address the empirical issue of implausibly high risk aversion within the context of the C-CAPM by looking at the pattern of risk aversion over the frequency domain. Our results show that as lower frequencies are taken into account, risk aversion falls substantially across countries and, in many cases, is consistent with more reasonable values of the coefficient of risk aversion.

Specifically, the traditional time series approach yields the usual finding of high risk aversion with estimates of relative risk aversion of 502 for Argentina, 175 for Mexico and 81 for Peru. However, the estimates of 36 and 16 for Chile and Brasil, respectively certainly

move the relative risk parameter in the right direction. When repeating the analysis by imposing a unit coefficient between stock returns and consumption growth, the estimated coefficients are reduced. With the exception of Argentina (RRA= 20) and Colombia (RRA=11), the remaining estimates of risk aversion are all within the economically acceptable values.

Turning to our frequency domain estimates, we get substantially improved estimates of relative risk aversion at low frequencies that range from 0.76 (Argentina) to 22.34 (Colombia). Our findings mainly stem from the fact that the variability of consumption growth increases significantly matching the variability of excess returns and as such the covariance of returns and consumption increases, as we move to lower frequencies. This property is coupled for most of the countries at hand with a rise in the estimated coherency (i.e. the correlation in the frequency domain) between consumption and returns. Furthermore, when a unitary correlation coefficient is imposed, these estimates are slightly reduced for all the countries at hand and range from 0.75 to 8.83.

This evidence shows some improvement towards understanding the dynamics of the C-CAPM in the Latin America countries by reconciling its standard single-factor version with lower values of risk aversion and thus the equity premium over the frequency domain appears to be less of a puzzle.

# REFERENCES

- Abel, A.B. (1990). Asset Prices under Habit Formation and Catching up with the Joneses. *American Economic Review*, 80, 38-42.
- Breeden, D.T. (1979). An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities. *Journal of Financial Economics*, 7, 265-296.
- Campbell, J.Y. (1996). Consumption and the Stock Market: Interpreting International Experience. *Swedish Economic Policy Review*, 3, 251-299.
- Campbell, J.Y. (2003). Consumption-Based Asset Pricing, Handbook of the Economics of Finance, George Constantinides, Milton Harris, and Rene Stulz eds., North-Holland, Amsterdam, Vol 1B, 803-887.
- Cochrane J. (2005) 'Financial markets and the real economy', *Foundations and Trends in Finance*, 1, 1, 1-101.
- Constantinides, G. (1990). Habit Formation: A Resolution to the Equity Premium Puzzle. *Journal of Political Economy*, 98, 519-543.
- Chatfield C. (1989). The Analysis of Time Series, London: Chapman and Hall.
- Engle R.F. (1974). Band spectrum regression, International Economic Review, 15, 1, 1-11.
- Engle R.F. (1976). Interpreting spectral analysis in terms of time-domain models, *Annals of Economic and Social Measurement*, 5, 89-109.
- Engle R.F. (1978). Testing price equations for stability across spectral frequency bands, *Econometrica*, 46, 4, 869-881.
- Epstein, L. and S.E. Zin (1991). Substitution, Risk Aversion and the Temporal Behaviour of Consumption Growth and Asset Returns: An Empirical Investigation. *Journal of Political Economy*, 99, 263-286.

- Granger C.W.J. and M. Hatanaka (1964). *Spectral Analysis of Economic Time Series*, Princeton: Princeton University Press.
- Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 357-384.
- Hamilton J.D. (1994) Time Series Analysis, Princeton: Princeton University Press.
- Hansen, L.P. and K.J. Singleton (1983). Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns. *Journal of Political Economy*, 91, 249-268.
- Jobson, J.D. and B. Korkie (1981). Putting Markowitz Theory to Work. Journal of Portfolio Management, 70-74.
- Kandel, S., and R.F.Stambaugh (1991). Asset returns and intertemporal preferences. *Journal of Monetary Economics*, 27, 39-71.
- Kocherlakota, N.R. (1996). The Equity Premium: It's still a Puzzle. *Journal of Economic Literature*, 34, 42-71.
- Lucas, R. (1978). Asset Prices in an Exchange Economy. Econometrica, 46, 1429-1445.
- Mehra, R. and E. Prescott (1985). The Equity Premium: A Puzzle. Journal of Monetary Economics, 15, 145-161.
- Salomons, R. and H. Grootveld (2003). The equity risk premium: emerging vs developed markets. *Emerging Markets Review*, 4, 121-144.
- Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics*, 24, 401-421.

Chapter 34

# PREPARING CONSUMERS FOR SHARED DECISIONS: ANALYSING THE EFFECTIVENESS OF A DECISION-AID FOR WOMEN MAKING CHOICES ABOUT BIRTH AFTER CAESAREAN<sup>\*</sup>

Allison Shorten<sup>1,†</sup>, Brett Shorten<sup>2</sup> and Sandra West<sup>3</sup>

<sup>1</sup>School of Nursing, Midwifery and Indigenous Health, University of Wollongong, Wollongong, NSW, Australia

<sup>2</sup>School of Economics, University of Wollongong, Wollongong, NSW, Australia <sup>3</sup>Faculty of Nursing, University of Sydney, Australia

# ABSTRACT

*Background*: The *Making Choices for Childbirth* study was undertaken to examine the effectiveness of a decision-aid to support women making choices for birth after caesarean section.

*Methods*: A prospective randomized controlled trial (RCT), involving 227 women recruited at 18 weeks gestation, measured levels of decisional conflict, knowledge and preferences for mode of birth at 28 weeks gestation and 36 weeks gestation. The decision-aid was administered to women in the intervention group after completion of the 28 week survey. A follow-up survey at 6-8 weeks postpartum focused on birth outcomes, health state and factors relating to preparation for decision-making, satisfaction with choice, decision support and overall birth experience.

*Results*: Of the 227 women, 193 (85%) completed surveys at both 28 and 36 weeks, with 169 (74%) returned at 6-8 weeks postpartum. Results for the prenatal components of the RCT are reported in Shorten et al (2005). The chapter will explore the relationship between various decision factors such as knowledge, decisional conflict and perceived preparation for decision-making. Analysis of postpartum data will include elements

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Health Care*, edited by Elizabeth P. Blakely, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> This project was supported by a NSW Midwives Association Research Scholarship, an MBF Health Fund Research Grant and University of Wollongong New Researcher Grant.

related to women's views on the decision support they received during pregnancy. The extent to which the decision-aid improved the likelihood of women feeling informed and being prepared to participate in decision-making will be discussed.

*Conclusions*: Evidence from the RCT suggests that the decision-aid is potentially valuable in assisting women who have experienced previous caesarean section to consider the risks and benefits of their options for birth after caesarean. However, this level of preparation for 'informed' decision making did not appear to translate into helping women discuss their options with their healthcare provider and others. In order to facilitate a process of shared decision making, stand alone products may need to be enhanced by strategies to better prepare practitioners to share decisions with consumers.

# **INTRODUCTION**

This chapter extends upon work undertaken to evaluate the effectiveness of a decision-aid booklet as a strategy for supporting women in making informed choices about their birth after previous caesarean section. The extent to which informed choice and shared decision making can be a reality in healthcare is addressed in the context of a randomised controlled trial. Evidence about the extent to which mode of birth was driven by women's choice, and the degree to which this was informed, was an important outcome of earlier work. The questions addressed within this chapter relate to the role that decision-aids may play in preparing consumers for decision making about their healthcare. In particular, whether the decision-aid through increasing knowledge, translates to consumers feeling more informed about their options and better prepared for making decisions in reality. The complex nature of the healthcare organizations and the impact this has in decision making should be taken into account when developing strategies to support shared decision making in healthcare contexts.

#### BACKGROUND

#### Support for Shared Decision-Making in Healthcare

From the beginning of the 1990's, alongside the development of the 'evidence-based' health care movement in Australia, paternalistic approaches to medicine have been challenged by initiatives focused on 'shared decision-making'. Concomitantly, organisations such as the National Health and Medical Research Council (NHMRC) have included the consumer in the process of information exchange, as part of the development of policies and guidelines relating to the formulation of clinical practice guidelines (National Health and Medical Research Council, 2000). Information that facilitates the process of consumer participation is now seen as crucial to a 'partnership' approach to medical decision making, but such initiatives are new in an Australian context and as such require careful evaluation.

Historically, imbalances in information between doctors and their patients (asymmetric information) have been accepted as inevitable, given the complexities of medical science. A paternalistic approach to medical care was thus justified in terms of protecting patients from their own ignorance. Acceptance of the professions control over knowledge and power in decision-making was therefore a prerequisite to the establishment of such medical

relationships (Charles, Whelan, & Gafni, 1999). In addressing the information imbalance, the medical profession has more recently moved towards a more inclusive decision relationship. Such relationships of necessity include much greater patient autonomy, which Nessa and Malterud (1998) consider to be a medico-ethical right. Thus, practicing in the context of informed consent when this is based on the ethic of patient autonomy, challenges practitioners to provide information about the risks and benefits of healthcare so that consumers are able to understand their options and make choices about them (Coulter, 1999).

The assumption that 'patients' cannot understand the complexities of their healthcare is slowly being eroded. Consumer organisations such as the Consumer Focus Collaboration (CFC) provide evidence that consumer participation, both as part of individual care and in healthcare services, leads to improvements in individual outcomes and to more accessible and effective healthcare services (Consumer Focus Collaboration, 2001). The suggestion that potential benefits to healthcare outcomes come as a result of increasing consumer control over their own healthcare is thus raised for discussion (Breemhar & van den Borne, 1991; Consumer Focus Collaboration, 2001), with further evidence needed regarding the impact of initiatives involving partnership approaches to shared clinical decision-making.

The formulation of key consumer roles in the development and evaluation of health care has been acknowledged by national bodies such as the NHMRC (2000) and international bodies such as the Cochrane Collaboration. The Australasian Cochrane Centre, established in 1994, now successfully melds the philosophy of evidence-based healthcare and the importance of consumer involvement in healthcare. Its aims reflect activities that;

"...promote the equitable provision of effective health care in Australasia by facilitating the preparation and maintenance of systematic reviews and their dissemination and application to influence service provision and clinical practice...relating to government, health professionals and consumer groups in Australasia ...facilitating the dissemination and application of information about the effects of health care to communities, health professionals and policy makers" (http://www.cochrane.org.au/<24Aug1998>; http://www.cochrane.org.au/aboutus/ <19April 2007>).

An important part of this is the Cochrane Consumer Network which serves to promote research that meets community priorities and to inform the community about results of healthcare research in a relevant way for consumer decision-making (http://www.cochrane.org.au/). Professional and government bodies, in embracing evidence-based approaches to care, have created a need and a niche for information development for both professionals in the form of 'clinical practice guidelines' and for consumers in the form of educational decision support strategies.

Strategies that are effective in supporting patient decisions extend beyond information and address the 'non-clinical' factors that impact upon an individual when faced with competing choices and uncertain outcomes. In addition to descriptive material, benefits and risks of options and likelihood of outcomes must be blended with personal values and individual clinical profiles. Decision support strategies developed and evaluated by researchers of the Ottawa Health Decision Centre Canada, and further endorsed by the NHMRC (2000), help to address these aims.

#### Decision Theory Underpinning a Framework for Support

Decision theory is based on an amalgamation of concepts and empirical information that helps to predict or describe the behaviours or actions of individuals when faced with choices. The notion of rational decision-making comes from a prediction about what human beings will do in circumstances of decision-making (Lee, 1971). 'Rational' decisions cannot be made without an understanding of the choices in terms of probable objective and subjective outcomes and information relevant to the decision (Lee, 1971). It is thought that individuals will value consequences of their decisions in different ways. Theorists use the term utility to note the expected value of an option, that is the weighted sum of payoffs or benefits for the individual, given a balancing of subjective probabilities of risks and benefits (Beach & Beach, 1982; Lee, 1971). The development of the science of decision-making behaviour, under conditions of uncertainty, although applied in the context of contemporary healthcare, has its foundations in disciplines such as psychology (Tversky & Kahneman, 1981) and economics (Arrow, 1965; Arrow, 1983).

Social psychologists in particular have studied the interaction between values, attitudes and behaviour in the context of decision-making. Fishbein and Ajzen (1975) in their 'expectancy-value' model acknowledge the relationship between individual attitudes and behaviours and the weighting of subjective probabilities, importance of perceived consequences of behaviours and expectations as well as subjective norms (influence of social environment) (Feather, 1982). The important issue to note here, in the context of healthcare, is that individuals will value outcomes of their healthcare options differently and will use information in different ways to guide their decisions. In acknowledging this complex process it is therefore important that values are in some way incorporated into the process of decisionmaking, which strengthens the argument that 'patient information' alone is likely to be inadequate for facilitating patient decision-making under conditions of uncertainty.

Conflict theory is also relevant in the context of healthcare decisions, because it is concerned with how individuals behave when faced with decision-making and how they can be actively involved in decision processes. Compatible with expectancy value theory, it acknowledges that individuals will avoid states of conflict or stress and move towards more pleasurable or desirable psychological outcomes. The manner in which this occurs is again an individual one and is influenced by individual coping styles or 'coping patterns'. This is believed to be especially important in conflict resolution and the possible impact of strategies to facilitate this process. It is thought that high quality decisions come from a strategy of 'vigilance', where the decision-maker searches for information that is relevant to the decision, assimilates information in an unbiased manner and then appraises the alternatives before making a choice (Feather, 1982). This is different to adaptive and defensive patterns which are thought to be less desirable as they involve a range of behaviours such as unquestioned acceptance, responsibility shifting, rationalisation, bolstering of the least stressful alternative and inattention to additional information that would involve change (Feather, 1982).

The aim of strategies to modify the way individuals cope with decisions is therefore to generate a vigilant approach and to counteract less effective patterns of coping (Feather, 1982). It follows then that interventions that involve information as well as procedures to modify decision-making behaviours may reduce decisional conflict (Feather, 1982). In aiming for an ideal such as 'vigilance' in decision-making, individual characteristics and past behaviours may still be difficult to counteract. In providing strategies to prevent the tendency to make a choice and 'bolster' that choice using defensive tactics such as minimising risks to promote the benefits of the choice (Janis & Mann, 1977), it should be acknowledged that individuals may still use these strategies to cope with decision stress. Therefore any evaluation of such strategies for improving the process of decision-making should

acknowledge such behaviours in their evaluation. In fact 'real-life' decisions may evoke different behaviours to hypothetical decisions which are often made during research. Therefore analysis of 'real-life' decision-making is important in assessing the effectiveness of strategies designed to assist individuals in this process. The issue of adequate time as well as social or organisational environment on decision behaviours should also be acknowledged as having a potential effect on individual decision-making, and should always be taken into account.

#### Caesarean Section: A Growing Dilemma

Events that occur during pregnancy and childbirth can have significant and prolonged effects on women and their families. Surgical procedures associated with childbirth, such as caesarean section (CS), impact upon all future pregnancies and subsequent childbirth decisions. Consequently decision-making about mode of birth for women who have already experienced CS challenges women, their families and their healthcare providers. It is in acknowledging that both CS and vaginal birth are associated with risks and benefits for women, that supporting a process of informed choice for childbirth after previous CS becomes important.

Caesarean section is growing in use and popularity despite recommendations to curtail its use. The 1985 WHO consensus statement about CS recommended a rate of between 10-15 percent, based on an estimation that higher rates would not provide benefits in maternal and perinatal morbidity and mortality (Senate Community Affairs References Committee, 1999). Despite the 1985 WHO recommendation for a rate no higher than 15 percent, in 1999 the Australian CS rate had reached 21.9 percent. According to historical figures from the Australian Institute for Health and Welfare (AIHW), the CS rate initially increased rapidly from 5 percent in the 1960's to 15 percent in the 1980's, with a steady increase in the 1990's of a further 5 percentage points to a rate of over 20 percent (Senate Community Affairs References Committee, 1999), continuing upward to a rate of 27 percent in 2002 (Laws & Sullivan, 2004), rising again to 29.4 percent of all births in 2004 (Laws, Grayson, & Sullivan, 2006). This phenomenon is not exclusive to Australia. The Centre for Disease Control (CDC) in the US reported increasing rates of caesarean section reaching 27.5 percent in 2003 (Martin et al., 2005), 29.1 percent in 2004 (Martin et al., 2006) up to 30.2 percent in 2005 (Hamilton, Martin, & Ventura, 2006).

International and Australian literature is consistent in presenting the major clinical determinants of CS as dystocia (failure to progress), foetal distress, breech position and repeat CS. Other factors include organisational characteristics (including hospital type and size); obstetrician characteristics (including attitudes, experience and litigation concerns) and characteristics of women themselves (including insurance status, education level, parity, age, personal preferences) (Health Department of Victoria, 1990; Senate Community Affairs References Committee, 1999; Thomas, Paranjothy, & Royal College of Obstetricians and Gynaecologists Clinical Effectiveness Support Unit, 2001). Rising CS rates do not appear to be due to increased clinical need and it is clear from the literature that the elective CS rate in particular is not determined by clinical/medical factors alone. Individual consumers and their practitioners appear to be considering a range of non-clinical factors in the process of making decisions about modes of birth.

#### Birth after Caesarean Section

The clinical issue of birth after CS is a significant one as 55,550 Australian women were reported to have experienced CS in 1999 (Nassar & Sullivan, 2001) growing to 74,300 in 2004 (Laws et al., 2006). More than 24,640 women who gave birth in 2004 were known to have experienced at least one previous caesarean section, and thus faced a decision about whether to attempt vaginal birth after caesarean (VBAC) or undergo repeat CS (Laws et al., 2006). During 2004 repeat CSs were performed for 81.6 percent of mothers who had experienced previous CS.

Even though both options are considered to be 'safe' for most women, the choice between attempted VBAC and elective repeat CS is not straightforward. Both options entail some degree of risk for mother and baby, although the weight of evidence favours trial of vaginal birth as an appropriate option that should be considered (American College of Obstetricians and Gynecologists, 1999 (ACOG); Flamm, Goings, Liu, & Wolde-Tsadik, 1994; Martel, MacKinnon, & SOGC, 2005; National Institute of Clinical Excellence (NICE), 2004; Perveen & Shah, 1997). Both the ACOG and The Society of Obstetricians and Gyneacologists of Canada (SOGC) reviewed the issue of birth after CS and concluded that most women with one previous low-transverse CS are potentially eligible for VBAC in an institution equipped to respond to emergencies such as scar rupture. The review recommended that the decision should be made after consideration of individual risks and benefits, with the ultimate decision to be made jointly by the woman and the obstetrician (American College of Obstetricians and Gynecologists, 1999). Studies have demonstrated that between 60 to 80 percent of women who attempt VBAC will achieve vaginal birth (Appleton et al., 2000), with the SOGC indicating that the rate of successful vaginal delivery ranges from 50 percent to 85 percent of the attempted trials (Martel et al., 2005 p 165). However, success is influenced by the indication for primary CS and attitudinal factors related to both women and practitioners (American College of Obstetricians and Gynecologists, 1999; Martel et al., 2005).

Given the availability of clinical recommendations for birth after CS, one may ask why this decision scenario presents a clinical dilemma for women and their practitioners. Part of the clinical problem comes from the presence of a surgical scar on the uterus, which is associated with a risk of rupture. Although the rate of rupture of scar is small (less than 0.5% (Appleton et al., 2000)), the consequences are serious with possible complications including hysterectomy and neonatal death (Appleton et al., 2000). The statement "Once a caesarean always a caesarean", however, is no longer clinically justified due to growing evidence about benefits of VBAC (Flamm et al., 1994; Norman, Kostovcik, & Lanning, 1993; Perveen & Shah, 1997). Trial of VBAC, also referred to as trial of labour (TOL) is therefore considered to be a viable option for most women. However, TOL, if unsuccessful, leads to complications such as emergency CS, which produce inferior outcomes (in terms of associated morbidity) to an elective caesarean procedure (McMahon, Luther, Bowes, & Olshan, 1996; Shorten, Lewis, & Shorten, 1998).

For many women and their practitioners, the decision is not purely clinical. Women's choices are characterised by a difficult balance between family or social commitments and relationships. Obstetricians must consider women's individual risk profiles and assess the likelihood of complications requiring emergency caesarean occurring, within a context of potential litigation (McClain, 1985, 1987, 1990). Even when the decision to undertake TOL

results in a vaginal birth, women will not necessarily feel satisfied with the decision or the outcome (Abitbol et al., 1993; Joseph, Stedman, & Robichaux, 1991). Women's views are important in determining method of birth after caesarean and attitudinal factors play a significant role in the choice and outcome (Fraser, Maunsell, Hodnett, Mountquin, & Childbirth Alternatives Post Caesarean Study Group, 1997).

Educational programmes both within Australia and overseas have been designed and implemented in an effort to encourage women to attempt VBAC and reduce overall CS rates, with limited success. The effect of facilitating informed choice rather than promoting one mode of birth over another, on women's pattern of preference/choice for birth, remains unclear (Eden, Hashima, Osterweil, Nygren, & Guise, 2004; Horey, Weaver, & Russell, 2004). Reports such as the "Changing Childbirth Report" (Department of Health, 1993) and the "Rocking the Cradle Report" (Senate Community Affairs References Committee, 1999), communicate the importance of consumer information in pregnancy and childbirth with the aim of facilitating 'informed decision-making'. Informed decision-making requires that women be presented with reliable information about the possible consequences of their choices, particularly when faced with a number of options (National Health and Medical Research Council, 1996).

#### The Ottawa Decision Support Framework

Healthcare decisions are often made in a context of uncertainty about the outcome. The Ottawa Decision Support Framework (DSF) acknowledges that many health care decisions are not straightforward and that non-medical factors influence consumer preferences and choices. There is an imbalance in information between the 'expert' provider of care and the consumer. Reliance upon the healthcare provider to make all decisions in cases where comparable options exist carries the risk that the decision will not be appropriate to the individual or that it is influenced by the bias of the practitioner and their interpretation of research, without consideration of the values and beliefs of the individual consumer. The DSF facilitates a more evidence-based approach to information sharing in a consumer-centred format.

O'Connor et al (1998) base the DSF on expectancy value, decisional conflict and theories of social support with the purpose of addressing health decisions that are;

- (1) stimulated by a new circumstance, diagnosis, or developmental transition;
- (2) require careful deliberation because of uncertain and/or value-sensitive nature of the benefits and risks; and
- (3) need relatively more effort during the deliberation phase than the implementation phase. "(O'Connor, Tugwell, Wells, Elmslie, Jolly, Hollingworth, McPherson, Drake et al., 1998:268).

The framework itself addresses three key areas including assessment of the determinants of decisions, a decision support strategy and evaluation of the decision support strategy. According to O'Connor's 1996 DSF (O'Connor, Tugwell, Wells, Elmslie, Jolly, Hollingworth, McPherson, Bunn et al., 1998) determinants of decisions include sociodemographic and clinical characteristics, perceptions of the decision (knowledge, expectations, values, decisional conflict), perceptions of significant others and resources for decision-making (including personal skills and characteristics). Decision support varies in

format and structure but includes research-based information about risks and benefits of options, values clarification exercises, opportunities for modifying expectations and enhancement of personal resources to cope with pressure from others in implementing decisions.

Evaluation focuses on the quality of decisions and the process of decision-making. Therefore indicators of good decisions include "knowledge, realistic expectations, clear values, congruence between values and choice, low decisional conflict, decision implementation, satisfaction with the decision and decision making process" (O'Connor, Tugwell, Wells, Elmslie, Jolly, Hollingworth, McPherson, Bunn et al., 1998:271). With the uncertainty in outcome associated with many health decisions however, it is important to acknowledge that although a decision may contain the components of what may be considered to be a "good quality decision", the actual outcomes may be adverse or negative in nature. Health may have deteriorated or the prognosis may have worsened. There are no guarantees that clinical outcomes will be better as a result of using a decision-aid. Given the element of chance associated with uncertain outcome it has been suggested that it is unfair to judge the quality of decisions by the clinical outcome (Ratliff et al., 1999). Rather, the process used to make the decision is crucial and a 'good decision' reflects principles of shared decision making including the acceptance of the patient's decision by the practitioner (Ratliff et al., 1999).

#### Decision-Aids as Shared Decision Strategies for Health

Decision-aids are more than providers of patient information, they are multi-dimensional tools that focus on decisions and the decision process. Decision-support strategies, although evolving from the field of 'patient education', can be distinguished from the discrete action of 'informing' because of the combination of detailed descriptions of evidence-based benefits and risks of options, use of explicit probabilities of risks and benefits (such as illustrated or numerical), inclusion of overt values identification and clarification exercises, personal identification of the importance of individual values in the decision, emphasis on the notion of 'choice' and the underpinning principle of shared decision-making (O'Connor, 1997). Items which are excluded from the decision-aid criteria include passive material (such as for informed consent), interventions designed to promote compliance to a recommended option or material not focused on making a decision (O'Connor, Drake et al., 1999; Ottawa Health Decision Centre, 2001).

Better quality decisions are thought to be made when patients have knowledge about options, have realistic expectations of outcomes, and are clear about their own personal values interacting with the decision (O'Connor, Drake et al., 1999). This includes the additional aspects of environment, emotion and culture. Decision-aids are therefore appropriate when there is a need for careful deliberation of alternatives for care, possibly because the decision involves making value judgments about the benefits relative to the risks and there is uncertainty in the outcome of options presented (O'Connor, Tugwell, Wells, Elmslie, Jolly, Hollingworth, McPherson, Bunn et al., 1998).

This is true for the choice of birth after CS, given the need for careful consideration of associated risks and benefits within the context of each woman's own clinical scenario. Women's preferences for birth relate closely to their individual needs, values, expectations and experiences, therefore information alone would be inadequate support for the decision-making process.

In essence, it is thought that the potential benefits of decision-aids are that they are capable of assisting individuals who are undecided about their choices to be objectively more informed and reach a decision. For those who already have strong preferences, decisions are more likely to be based on improved knowledge, realistic expectations and be consistent with personal values (O'Connor, Drake et al., 1999). Decision aids appear to have little effect on levels of satisfaction with the process and do not appear to increase anxiety, however their impact on actual choices vary depending on the decision context (O'Connor et al., 2002). It is clear from the review that decision aids vary in format and that it is still difficult to compare many decision aids due to the variations in methods used to study their effectiveness. Future research should examine measures that enable comparison with studies so far, such as knowledge, decisional conflict and perhaps satisfaction with the decision making experience. Adequately powered randomised controlled trials involving participants actually facing treatment decisions, are likely to reveal better information about the effect of decision-aids in facilitating active participation in informed choices for healthcare consumers.

#### **Research Questions**

In exploring the literature on determinants of method of birth the choice appears to be multi-dimensional, involving non-clinical as well as clinical factors and varied involvement of women and their practitioners in the decision-making role. The notion that practitioners are acceding to women's demand for CS over vaginal birth is not quantitatively justified. The degree to which women are informed of the risks and benefits of CS and vaginal birth is still unclear. The effect of increasing women's knowledge of their options for birth and their perception of feeling informed and prepared for decision making remains spectulative. The questions to be addressed in this chapter are;

- What is the relationship between increased knowledge and women feeling informed about their options for birth?
- To what extent does a decision-aid contribute to women's preparation for shared decision making for birth?

The literature supporting the use of decision support strategies, indicates that healthcare decision-aids have a potential benefit for women in pregnancy. Their strength lies in improving knowledge levels of users and reducing the decisional conflict experienced when faced with actual decisions (Shorten, Shorten, Keogh, West & Morris 2005). The extent to which they prepare women for shared decision making is less clear.

# METHOD

This use of a decision-aid was evaluated through a randomised controlled trial (RCT) using multiple study sites. The use of an RCT was considered important in terms of the level of evidence that can be provided by a well designed RCT when compared with that achieved through use of a 'before and after cohort' design (National Health and Medical Research

Council, 1995; Wallace, Shorten, & Russell, 1997). This RCT thus comprised a control group receiving strictly 'routine care' and an intervention group, whose only departure from 'routine care' was provision of a decision-aid. Development and piloting of the decision-aid has been reported elsewhere (Shorten, Chamberlain, Shorten, & Kariminia, 2004).

Multiple study sites were used because of the variations in rates of CS between hospitals and practitioners within Australia (Appleton et al., 2000). Therefore a multi-site study was necessary to provide a level of comparison between patterns of practice and also to determine the usefulness of the decision-aid under conditions of different demographic characteristics of patient populations. Sample size calculations were conducted to ensure that each study site achieved the sample size required to support individual analysis of data on key measures, thus ensuring that comparison between sites as well as randomised groups was possible.

All women participating in the study were surveyed three times during their pregnancy at 12-18 weeks (S1); 28 weeks (S2) and 36 weeks (S3), and then after the birth at 6-8 weeks (S4). The control group women received 'routine care' only, according to their specified model of pregnancy care. Women in the intervention group, in addition to their 'routine care', received the decision-aid booklet from the researcher at 28 weeks of pregnancy. A description of the study protocol, including eligibility criteria, recruitment process, randomization protocol, determination of sample size and survey administration protocol can be found in Shorten et al (2005).

#### Data Collection Tools

The following provides a description of the selected measures and the survey in which they appeared. The measures of interest were knowledge, decisional conflict and satisfaction with information.

#### Knowledge of Health EFfect Score (Survey 2 and 3)

A baseline test of knowledge of health outcomes for trial of vaginal birth and elective caesarean birth was constructed using the format utilised by the Ottawa Health Decision Centre. (A generic knowledge test for the issue of birth after caesarean section was not available). The test contained 15 brief statements about the major risks and benefits of trial of labour and CS. Women were asked to circle their response as True, False or Unsure. The questions were constructed to relate closely to the information contained in the decision-aid (major pros and cons of trial of vaginal birth versus elective caesarean).

#### Decisional Conflict Score (DCS) (Survey 2, 3 and 4)

The DCS is an 18-item scale developed by Annette O'Connor using 5-point Likert format, including subscales on *Certainty, Feeling Informed, Values Clarity, Decision Quality and Feeling Supported* (O'Connor, 1995, 1999). Its test-retest coefficients and alpha coefficients are >0.80. It measures the degree of uncertainty about a course of action with scores of 2 or less being associated with decision-making, therefore discriminating between decision delay and decision-making (O'Connor, 1995, 1999). Uncertainty may arise from inherent decision factors such as risks and benefits and modifiable factors such as inadequate knowledge, values and expectation (Ottawa Health Decision Centre) (O'Connor, 1995). The

DCS has been used extensively in decision-aid trials and has been documented to be sensitive to changes following decision-aid interventions (O'Connor, 1995; O'Connor et al., 2002) and was therefore thought to be the most appropriate available measure of possible impact of the decision-aid.

#### Postnatal satisfaction with information (Survey 4)

In order to assess level of satisfaction with information received during their pregnancy, a set of questions was adapted from a generic survey exploring satisfaction with preparation for decision-making (Graham & O'Connor, 1999). The generic survey alpha coefficient was stated to be >0.90 and it was documented as being able to discriminate between interventions such as a pamphlet and a decision-aid with an effect size of 1.7 (p=0.001) (Graham & O'Connor, 1999). This part of the survey consisted of 9 statements about how information received during pregnancy may have assisted with factors associated with preparation for decision making. The responses were then analysed according to three categories 1 = 'Not at all/Very little'; 2 = 'Somewhat'; 3 = 'A lot'/ 'A great deal'.

#### **Ethics Approval**

Ethics approval was given by Human Research and Ethics Committees of The University of Wollongong, University of Sydney and participating hospitals. Obstetric and Gynaecological Committees of each hospital were provided with the decision-aid prior to full approval to ensure appropriate clinical content and consistency with individual hospital policy and practices.

# RESULTS

#### Sample

Women were recruited over a period of two years, between May 2001 and June 2003. Of the 252 women approached to participate, only 25 declined, with 227 women recruited in total. Participants were randomly allocated to either control or intervention (decision-aid) groups, resulting in 115 women being allocated to the intervention group and 112 allocated to the control group. A summary of the recruitment and flow of participants through the study protocol is presented in Figure 1.

The demographic characteristics of participants were similar for both intervention and control groups, with no statistically significant difference reported for a range of socioeconomic factors. The clinical characteristics of women were also similar except for the responses to a question regarding perceived problems experienced after previous CS, for which 46.1% of the intervention group women answered 'yes' whereas only 33% of the control group women answered 'yes'. A full description of participant characteristics can be found in Shorten et al (2005).

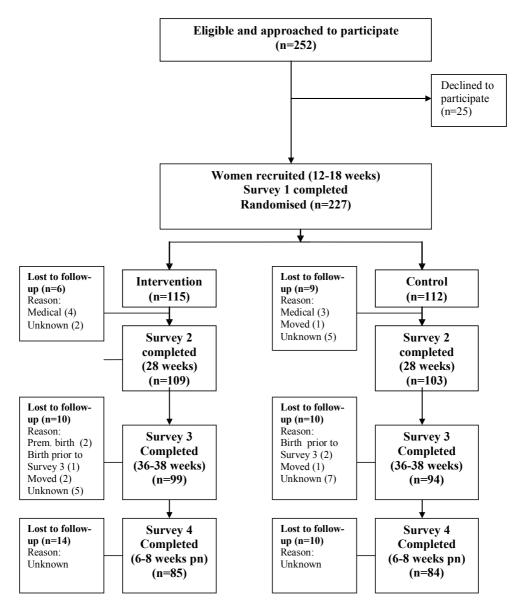


Figure 1. Flow of participants through trial (summary). Note: Only 93 control group women completed both Survey 2 and Survey 3; pn =postnatal.

#### Relationship between Knowledge and Feeling Informed

The decision-aid was the most reliable predictor of positive change in knowledge score for women during pregnancy. As reported in Shorten et al (2005), the mean increase in women's knowledge scores across the pregnancy was 2.17 for the decision-aid (intervention) group versus 0.42 for the control group, this difference of 1.75 points being statistically significant (p<0.001, 95% CI 1.15-2.35). The majority of women (78.8%) in the decision-aid group increased their knowledge score, with most scores increasing by at least 3 questions

(58%), compared with less than half (46.8%) of the women in the control group, of whom only 25% increased their score by 3 or more. Over one third (34.8%) of the control group women recorded a decrease in knowledge score at Survey 3 (36 weeks). In addition, women in the decision-aid group were significantly less likely to circle 'unsure' responses on the knowledge test at Survey 3 when compared with women in the control group ( $\chi^2$ , 3*df* 32.52 (p<0.001).

As reported in Shorten et al (2005) women who received the decision-aid experienced greater reduction in decisional conflict scores (DCS) between Survey 2 and Survey 3. The mean DCS scores for the decision-aid group at Survey 3 were 1.94 compared to 2.17 for the control group (p<0.01). Women who received the decision-aid experienced a reduction in DCS of -0.40 compared to -0.08 for controls (p<0.01, 95% CI -0.51 to -0.29).

The following analysis explores the relationship between a measured increase in 'knowledge score' and the DCS subscale of feeling "informed". In Table 1, average "informed" scores within the DCS are tracked across time, comparing scores at Surveys 2, 3 and 4. For all respondents, women on average felt more informed at Survey 3 (mean score = 1.96) than at Survey 2 (mean score = 2.25), and more informed again at Survey 4 (mean score = 1.77). However, due to the attrition rate during the study average "informed" scores need to be recalculated for the 164 women completing all 3 surveys only. Table 1 indicates however, that the mean scores barely changed after recalculation, suggesting that women who dropped out of the study were a random group, at least with respect to their opinions about the level of information they possessed relevant to the decision faced.

Survey	ey All Respondents		Answer Surveys	ed All 3		ol Group Surveys)		n-Aid Group arveys)
	n	Mean	n	Mean	n	Mean	n	Mean
S2	211	2.25	164	2.28	80	2.33	84	2.23
<b>S3</b>	193	1.96	164	1.93	80	2.11	84	1.76
<b>S4</b>	168	1.77	164	1.77	80	1.88	84	1.66

Table 1. Average scores for "Informed" sub-scale of DCS for S2, S3 and S4

Analysing the information from women completing all surveys, via matched pairs comparisons of means, it was found that the reduction over time in "informed" scores was statistically significant (S2-S3 = 2.28-1.93, p<0.000 and S3-S4 = 1.93-1.77, p<0.000). Hence, the finding is that, on average, all respondent women felt better informed about the birthmode decision they faced making as their pregnancy progressed.

The next question was whether these feelings of being better informed related to increases in objective knowledge about the birthmode issues faced by women who had experienced previous CS. As a first step to answering this question, Table 1 takes the group of 164 women who answered Surveys 2, 3 and 4 and splits them into the decision-aid and control group, on the basis that the former were found to have significantly greater knowledge of the relevant decision-making issues subsequent to receiving the decision aid booklet (Shorten et al, 2005).

Factor	Survey 2	Survey 3	Survey 4
	(n=199)*	(n=184)*	(n=156)*
Mean Informed Score	2.24	1.97	1.77
Mean Knowledge Score (/15)	8.74	10.15	10.33 (S3)
Booklet (Control) Group (%)	51 (49)	51 (49)	51 (49)
Birth Preference (%)			
Trial Of Labor	49	48	N/A
Elective CS	25	35	N/A
Undecided	26	17	N/A
Mode Of Care (%)			
Midwives Clinic	30	31	28
Team Midwifery	14	14	14
GP Shared Care	32	32	34
Doctors/High Risk Clinic	5	4	3
Private Obstetrician	19	19	21
Study Site 1 (Site 2) (%)	67 (33)	66 (34)	68 (32)
Australian Born (Other) (%)	67 (33)	68 (32)	69 (31)
Level Of Education (%)			
Year 10 or less	20	19	15
Year 12	8	8	9
Certificate/Diploma	34	35	36
Degree or Higher	38	38	40
Employment At S1(%)			
Full Time	14	14	15
Part Time	36	35	39
Other	6	6	4
Home Duties	44	45	42
EPDS Category (%)			
Low	69	69	71
Medium	22	22	19
High	9	9	10

Table 2. Summary Statistics for Regression models of "Informed" Scores

\* Numbers of women at each Survey vary somewhat from those reported in Table 1 due to missing data on some variables.

There was no significant difference in feeling "informed" between the two groups at Survey 2, prior to the intervention. Average scores for the booklet (n=84) and control (n=80) groups were 2.23 and 2.33 respectively (p=0.40 according to a comparison of means). Overall, women in both groups felt better informed at Survey 3 than at Survey 2. For the control group, the mean "informed" score fell from 2.33 to 2.11 (p<0.01 according to matched pairs comparison of means) whilst for the intervention group, the improvement was from 2.23 to 1.76 (p<0.000). However, it can be seen that the improvement was more dramatic for the booklet group, so that the difference in mean scores at Survey 3 of 1.76 versus 2.11 for the control group was highly statistically significant (p<0.000). Given the

significantly higher knowledge scores of the booklet group at Survey 3, it thus appears that the positive feelings regarding information held (feeling informed) may well be related to measured levels of objective knowledge. It is also evident from Table 1 that this advantage for women who received the decision aid booklet persisted beyond the pre-natal period. At 6-8 weeks post-partum (survey 4), the difference in mean "informed" scores between intervention and control group women was still statistically significant (p=0.02). Interestingly, statistically significant improvements in feeling informed occurred in both groups between Surveys 3 and 4. For the intervention group, the mean score fell from 1.76 to 1.66 (p = 0.05), whilst analogous results for the control group were 2.11 to 1.88 (p<0.000).

To further explore the associations between knowledge, access to the decision aid booklet and "feeling informed", a series of regression models, using "informed" scores from Surveys 2, 3 and 4 as dependent variables was conducted. Table 2 provides summary statistics for all variables used in these regressions.

Table 3 summarizes regression results for reported "Informed" scores. Note that, in the model for Survey 4, Knowledge scores from Survey 3 were used, as this quiz was not administered as part of Survey 4. Similarly, Edinburgh Postnatal Depression Scale (EPDS) questions were not asked at Surveys 2 and 3, so scores from Survey 1 were used in these regressions. Reference categories for categorical variables were control group, preference for Trial of Labor, Midwives Clinic mode of care, Year 12 level of education, Home Duties employment, born outside Australia and low EPDS score.

Factor	Survey 2	Survey 2 Survey 3 (n=199)	
	(n=183)		(n=156)
Knowledge Score	-0.027 (-1.34)	0.001 (0.07)	-0.01 (-0.35)
Booklet Group	-0.087 (-0.90)	-0.334 (-3.37)***	-0.18 (-1.61)
Birth Preference			
Elective CS	-0.114 (-0.94)	-0.053 (-0.52)	-0.15 (-1.31)
Undecided	0.393 (3.35)***	0.307 (2.48)**	0.04 (0.79)
Mode Of Care			
Team Midwifery	-0.086 (-0.52)	-0.028 (-0.19)	0.04 (0.22)
GP Shared Care	0.003 (0.03)	0.182 (1.60)	0.12 (0.96)
Doctors/High Risk Clinic	0.237 (0.96)	-0.081 (-0.33)	-0.12 (-0.40)
Private Obstetrician	0.289 (1.73)*	-0.034 (-0.23)	-0.11 (-0.65)
Study Site 1	-0.086 (-0.69)	-0.078 (-0.67)	-0.01 (-0.10)
Australian Born	-0.022 (-0.18)	-0.041 (-0.38)	0.03 (0.26)
Level Of Education			
Year 10 or less	0.105 (0.51)	-0.063 (-0.33)	0.14 (0.64)
Certificate/Diploma	0.087 (0.46)	-0.010 (-0.06)	0.11 (0.56)
Degree or Higher	0.046 (0.23)	-0.056 (-0.30)	0.34 (1.74)*

 Table 3. Regression models for factors associated with DCS "Informed" Scores (t scores in parentheses)

Factor	Survey 2 (n=183)		
Employment (%)			
Full Time	0.294 (1.95)*	-0.022 (-0.16)	-0.06 (0.37)
Part Time	-0.019 (-0.16)	-0.083 (-0.78)	-0.21 (-1.75)*
Other	-0.203 (-0.93)	0.097 (0.50)	-0.41 (-1.64)
<b>EPDS Category</b>			
Medium	0.074 (0.62)	0.166 (1.50)	0.24 (1.83)*
High	0.018 (0.11)	0.295 (1.82)*	0.08 (0.64)
R <sup>2</sup> Adjusted	0.073	0.122	0.03
F Statistic	1.87**	2.41***	1.30

#### Table 3. Continued

\* P < 0.10 \*\* P < 0.05 \*\*\* P < 0.01

# Table 4. Post-Natal Responses to Questions Regarding Helpfulness of "Birth Choices" Booklet (n=84)

Question: How much did the booklet help	Not at all/	Somewhat	A lot/
you to	Very Little		A great deal
	n (%)	n (%)	n (%)
1 "organise your own thoughts"	5 (6.0)	43 (51.2)	36 (42.9)
2 "consider pros and cons of each option"	5 (6.0)	20 (23.8)	59 (70.2)
3 "Identify questions you needed to ask"	7 (8.3)	25 (29.8)	52 (61.9)
4 "consider how involved you wanted to be"	14 (16.7)	29 (34.5)	41 (48.8)
5 "discuss options with family"	19 (22.6)	25 (29.8)	40 (47.6)
6 "discuss options with doctor/midwife"	15 (17.9)	30 (35.7)	39 (46.4)
7 "prepare you to make a decision"*	9 (10.8)	29 (24.9)	45 (54.2)
8 "know what to expect from your choice"	11 (13.1)	23 (27.4)	50 (59.5)
9 "help you to feel satisfied "	15 (17.9)	25 (29.8)	44 (52.4)

\* n=83

Results suggest that it is not knowledge, as reflected in Knowledge quiz scores, that appears to affect the subscale of "feeling informed". In none of the three Surveys is this variable a significant explanator of variations in DCS "informed" subscale score. However, at Survey 3 especially, and to some extent at Survey 4 as well, women in the intervention group (decision-aid), reported significantly better levels (i.e. lower scores) of DCS "informed" scores. This suggests that it may be other aspects of the decision aid, rather than the knowledge it imparts per se, that contributes to women feeling better informed about the decision they faced. Although the study group is moderately highly correlated with knowledge especially at Survey 3, the results reported in Table 3 do not appear to reflect problems of multicollinearity. For the Survey 3 regression, for example, the correlation between knowledge score and study group is 0.425. However, when the former is omitted from the regression model, the coefficient on study group barely changes (from -0.334 to - 0.331), as does the explanatory power of the model ( $R^2$  rises from 0.122 to 0.127). When

study group is omitted, the co-efficient changes from 0.001 to -0.031, and the t statistic changes to -1.63 (p = 0.11), but the explanatory power of the model reduces dramatically, with R<sup>2</sup> falling from 0.122 to 0.067. This suggests that study group (intervention or control) is important in explaining variations in perceptions of feeling informed, rather than knowledge scores per se. The decision-aid therefore appears to have functioned effectively in terms of both improving knowledge and reducing decisional conflict.

Women who received the decision-aid reported that the booklet had helped a lot or a great deal in helping them to consider the pros and cons of their options as well as in helping them to identify questions to ask and in knowing what to expect (Q2,Q3,Q8). The booklet was reported as less helpful in terms of assisting women in discussions with their family or healthcare provider (Q5,Q6).

	Not at All/Very Little		Somewhat		A Lot/Great Deal	
Question	Booklet	Control	Booklet	Control	Booklet	Control
	(%)	(%)	(%)	(%)	(%)	(%)
1*	7.1	23.8	36.5	38.1	56.5	38.1
2*	5.9	22.6	28.2	36.9	65.9	40.5
3**	11.8	25.0	36.5	34.5	51.8	40.5
4	20.0	17.1	30.6	32.9	49.4	50.0
5	20.0	17.9	30.6	42.9	49.4	39.3
6	17.9	19.0	28.6	31.0	53.6	50.0
7	15.3	23.8	32.9	31.0	51.8	45.2
8	12.9	16.7	34.1	28.6	52.9	54.8
9	14.3	16.7	35.7	32.1	50.0	51.2

 Table 5. Post-Natal Responses to Perceived Information Received About Birth Choices

 During Pregnancy by Study Group (n=169)

\* p<0.01;

\*\* p<0.10.

When comparing responses about the information received during pregnancy for women who received the decision-aid and those in the control group (summarized in Table 5), the strength of the decision-aid again appears to be in the provision of information as a means of preparation for decision-making.

Statistically significant differences were found for the first three questions (according to a  $\chi^2$  test with 2df), with booklet group women more likely to feel that information received had been helpful in the dimensions of organizing their thoughts about their decision for birth, considering the pros and cons of each option and helping to identify questions they needed to ask. However, women who had received the booklet were not significantly more likely to feel more positive about information received on the other six question, which related to other important aspects of shared decision making such as usefulness in discussion with healthcare providers.

# DISCUSSION

One of the key objectives of this research was to determine the effectiveness of a decision-aid in preparing women for participation in informed decision making about their mode of birth after CS. In order for this to be determined, a strategy for improving knowledge and for facilitating a process of informed choice was developed. The decision support framework was adopted for this purpose because of its potential value in promoting shared healthcare decision-making for childbirth.

#### Effect of Knowledge on Feeling Informed

Decision-aids have been shown to be effective in improving knowledge of healthcare options. Previous systematic reviews of decision aids have emphasised that decision-aids are better than 'usual' care in improving knowledge about healthcare options (A. O'Connor, Rostom et al., 1999; A. M. O'Connor et al., 2002). The decision-aid developed and evaluated by this RCT was the most important determinant of improvement in knowledge about options for birth after CS (A. Shorten, Shorten, Keogh, West, & Morris, 2005). However, results suggest that it is not knowledge as measured by the knowledge test and as reflected in actual test scores, that appears to affect the extent to which women felt informed about their options for birth.

Relatively few of the variables in the regression models proved to be statistically significant explanators of variations in women's perceptions about feeling informed. However, the following variables were found to have explanatory power in at least one of the three regression models (Table 3).

- Women who were undecided regarding their preference for Trial of Labor versus Elective CS at Surveys 2 and 3, exhibited significantly worse (higher) scores in terms of feeling informed. However, this apparent confusion appears to have largely vanished in the post-partum period, as no significant differences were found at Survey 4.
- Mode of pre-natal care appears to have played little role in women's perceptions of feeling informed, except that at Survey 2 those attending private obstetricians reported lower levels of information. This effect seems to have been resolved by Survey 3.
- Level of education was also found to play little role in perceptions of information held, except that, at Survey 4, those with the highest levels of education reported feeling less well informed. Perhaps this indicates a need for greater depth in information provided or a greater ability of these women to have reflected on their decision, the level of information that they had in making the decision, and their birth experience in the post-partum period.

Qualitative information in the form of written comments collected within the surveys added some further insight into the empirical information presented about the issue of feeling informed. When asked to provide an explanation of their choice for mode of birth and details about their birth experience, some of the women participating in the control group for this study revealed concerns about not feeling well informed and wrote additional statements on their surveys to express their unease about not receiving adequate information about their options.

"I am finding it difficult to know what is best to do. I haven't really got an understanding of how safe a trial of vaginal birth is. Also I feel I don't really have enough information about either option." (#1037 Control)

This comment below raises a similar concern, also from a woman in the control group.

"I realise how uninformed I was about the 2 options. Additionally the doctors didn't really approach the subject until 36 weeks and at this stage it was much too late to start a rational discussion and research process for myself. Certainly it is an important decision and the whole issue appears quite political varying from doctor to doctor and even country to country." (#2015 Control)

A perception of 'bias' came from a range of healthcare practitioners within different models of care. This appeared to have had an impact on how women felt and may have hampered their decision making.

"...changed to a private doctor. Really want a caesarean section [I] had no choice last time and feel angry that midwives make me feel that I am a failure and shouldn't want a caesarean." (#2014 Decision-aid)

The role that knowledge plays in the decision process is therefore potentially affected or even hampered by some practitioners and their own personal values and biases about birth. This was an issue for both medical practitioners and midwives who clearly communicated personal feelings about women's choices if they did not align with their own beliefs about what women should choose.

In contrast, some of the comments shared by women who received the decision-aid booklet and who were supported by their practitioners in making a choice, were very positive. In particular, if women had come to a decision that they were confident about, they expressed the positive feeling that came with that confidence.

The word *peaceful* was expressed on a number of occasions when explaining what choice they had made...

"I really felt the book was helpful. I now understand the terms and feel more informed about my choice. I took the booklet with me to my midwife and we looked at it together...I feel peaceful about this decision as it cuts out the mystery and the chance of an emergency Caesar which I want to avoid at all costs. It meets the needs of everyone in my family, me, the baby and my children and husband." (#1010 Decision-aid)

#### Decision-Aids as Preparation for Shared Decision Making

The attributes of decision-aids that are helpful for consumers in the decision process appear to be associated with the degree to which the decision-aid helps to inform and identify issues of importance to the women when making this decision. Women's perceptions of the value of the decision-aid in helping *their* decision making during pregnancy was closely linked to being able to organize *their* own thoughts about *their* decision, consider the pro and cons of *their* options and identify questions that *they* needed to ask. The decision-aid performed less effectively in relation to issues of communication between the woman and either her family or healthcare providers in terms of its use in facilitating a discussion about options for birth.

Qualitative information confirmed that the decision-aid was helpful in assisting women to consider risks and benefits of options. In fact some women wrote their explanation about their choice using the terms risks and benefits as seen in the two examples below;

*"For me, the benefits of c-section definitely outweighed those of trial of vaginal birth" (#2059 Decision-aid)* 

and

"My decision to have an elective caesarean was taken on the basis of the baby's size and my medical condition at the time. These two factors meant that the risks of a trial of scar became greater than the benefit..." (#2005 Decision-aid)

This overt weighing up of the 'risks and benefits' and applying this in the context of their individual situation is important and a potentially valuable aspect of the decision-aid. Knowledge in itself may not be useful unless it can be applied to an individual's personal set of circumstances, needs and values. The decision-aid may have encouraged women to organize their own thoughts and concerns about values and needs and facilitated their relating of these to the pros and cons that were specific to their experience.

McClain's work in the 1980's raised the question about the extent to which risk-benefit assessment or weighing the pros and cons played a role in the context of making choices about childbirth services and providers of care (McClain CS, 1983). The concept of bolstering (Janis IL & Mann LI, 1977), applied to the context of childbirth, explains the tendency for women to play up the advantages of one course of action in order to reduce conflict and avoid post-decision regret (McClain CS, 1983). This bolstering also involves downplaying the benefits of the alternatives and emphasising their risks (McClain CS, 1983). The risk-benefit assessment of options using the bolstering hypothesis indicates that women could be expected to articulate choices in terms of the least 'risky' and most attractive.

The results of the RCT, although it was not attempting to investigate the phenomenon of bolstering, lend some support to these decision behaviours. In the example below, the risk of ruptured uterus, although numerically small is stated as a significant risk in the justification for preferring caesarean section. The reflection on past negative experience adds weight to the likely event of further negative experiences so the caesarean section is the 'least risky' alternative in this case.

"I feel that having a caesarean birth would avoid the risk of having a rupture uterus and the bad experience I had with my first one. I just don't want to take the risk." (#1028 Decision-aid)

Women may use this process as a coping strategy. In fact some women referred to their decision as a way of coping. In stating why she had chosen caesarean section this participant said;

"Convenience, save strength of when the baby comes – Better coping strategy" (#2087 Decision-aid)

and another who selected caesarean to reduce stress;

"Feel that is the best for me and my husband. Less stress involved" (#2045 Control)

The role of the decision-aid in reducing decisional conflict (one aspect of which is in feeling informed) may have assisted women in 'bolstering' as a strategy to reduce conflict. Women may have been able to use the information on risks and benefits to assist them to bolster their preferences, however the extent to which this phenomenon occurred in the study is speculative.

The Extent to which Decisions were Shared

When one analyses the extent to which decision-aids prepare consumers for shared decision making the assumption is that the decision is one that could or should be shared. In the context of birth after caesarean section, this was not necessarily the case and may have impacted upon the perceived benefit of the decision-aid in preparing women to participate in their birth decision. In addition to the evidence that already exists suggesting that women's decisions were significantly influenced by the hospital site in which they were birthing (Shorten et al , 2005), women shared accounts of their decision experience which range from perceptions of a lack of choice to no choice at all.

The statement that they did not feel they had a choice at all was made by a number of women in the study;

"I've been told the choice is not mine. I have been told the decision is not mine as I don't have Private Health Insurance. The decision will be made at 36 weeks by an obstetrician. I strongly would like another caesarean due to the stress my last baby went through. I wish the decision was mine." (#1038 Control)

On a subsequent survey (at 36 weeks) this same participant stated...

"I'm afraid of my scar from previous surgery ripping. I've been <u>told</u> to have a trial of labour-not asked. It hasn't been my decision. I would like a repeat caesarean but have been told by Dr's I must have a trial of labour first. I've been told there's a 50% chance (of successful vaginal birth) which I don't feel is enough. I feel a caesarean birth is best for me...I'm afraid I will go through a painful labour and then end up have a caesarean anyway" (#1038 Control)

The suggestion that medical practitioners in each of the study sites were inclined towards either Vaginal Birth After Caesarean (VBAC) or Elective Repeat Caesarean Section (ERCS), and thus influenced decisions according to their preferences, rather than the preferences of the women, has some support in the literature. Barriers to patient participation in healthcare decision-making have been suggested to include the complex language of doctors which is difficult for patients to understand or use for decision-making, a tendency for doctors to dominate discussions with a limited capacity for listening to patients and patients being discouraged from asking questions (Molenaar, Sprangers, Postma-Schuit, & Rutgers, 2000). Therefore if women were prepared to ask questions but were not given the opportunity, then the decision-aid could be expected to be less effective in helping them to discuss options with their doctor or midwife.

Informal discussions with participants lend support to these ideas, with some women clearly articulating that they felt that they had been discouraged from asking questions and that they felt that they did not have a choice at all. Barriers clearly exist and in order for effective strategies to be developed to assist in the process of patient participation, such barriers need to be more clearly identified.

The potential impact of organisational culture and clinical practice patterns on consumer decision-making and the inhibiting effect this could have on true consumer choice is important. The notion of the impact of societal culture on choice for mode of birth has been suggested in an Australian context (Walker, Turnbull, & Wilkinson, 2004). Community perceptions or cultural norms portray caesarean section as an accepted "easy and convenient" way of giving birth (Walker et al., 2004). However the result of this RCT illustrates the level of influence that not only individual doctors or midwives may have in the decision process, but the impact that actual hospital site, area health service or organisational culture has in this decision.

In acknowledging the relevance of women's decision-making about their healthcare for the nursing and midwifery professions, Ruth Witmann-Price (2004) suggests that decision theories such as those underpinning decision-aids do not account for the societal norms that can affect individual value systems in a very oppressive way. The notion of oppression and the struggles of emancipation are raised as barriers to shared decision-making in the context of women's healthcare. The idea that historically healthcare decision-making reflects social norms of medical paternalism is raised again here as an important factor in the movement to a paradigm of shared decision-making between 'patient' and provider of care (Wittmann-Price R, 2004). Midwifery 'paternalism' is as much an issue as medical paternalism in this context and the negative comments directed towards women who chose caesarean section rather than trial of labour have been identified by women in the study. Receiving such comments appeared to have a negative impact on women's perception of the pregnancy and birth experience, and in some cases led to feelings of 'failure', such as in the comment "...midwives make me feel that I am a failure and shouldn't want a caesarean". The already difficult decision about mode of birth was made even more challenging if women felt they were not supported by their midwife, doctor or even their family.

One of the key components to an 'emancipated' decision-making process is a '*Flexible Environment*' (Wittmann-Price, 2004). In the true sense of this idea it would involve 'free'

choice and a 'non-judgemental' support of choice for healthcare options (Wittmann-Price, 2004). Without this flexibility, the effects of the improvement in personal knowledge would be lost to other powerful factors such as societal norms and the women's perceived levels of power within health care environments. The subjugation of knowledge-based choices to the influence of hospital culture and practitioner attitudes is suggested by the findings of this RCT. This is consistent with the issues raised by Kirkham and Stapleton (2004) in their discussion of the culture of maternity services in Wales and England and its role as a barrier to informed choice (Kirkham & Stapleton, 2004). Although the language of informed choice was always synonymous with the 'right' choice according to local established practices (Kirkham & Stapleton, 2004 p. 131).

This concept of problematic difference in power between participants in healthcare decision-making is not new and was raised as an important factor in limiting the success of tailored 'informed choice leaflets' in a recent UK study (Stapleton, Kirkham, & Thomas, 2002). The availability of information was not sufficient in itself, particularly as the information was withheld in some cases where, for example, obstetricians determined (without consultation) that a choice should not be available. The conclusion that such power imbalances need to be addressed to further enhance strategies based on the provision of information (Stapleton et al., 2002) is consistent with the findings of this study. Increasing the understanding of all practitioners about how to facilitate a process of 'shared' decision-making, given time constraints and associated work pressures, is important if healthcare services wish to move from a state of 'informed compliance' to genuine informed choice (Stapleton et al., 2002).

This raises the question about criteria for judging the effectiveness of decision-aids in the context of childbirth. The decision-aid in this RCT was effective in improving knowledge and reducing decisional conflict, however it may be inappropriate to assume that the presence of a decision-aid will in itself facilitate shared decision making. Important and influential determinants of such decisions which are related to the practitioners as well as the organisations within which they practice, are possibly beyond the scope of a decision-aid.

An alternative view might be that shared decision making is not appropriate for the clinical context of birth after CS. There is a level of uncertainty in pregnancy, where circumstances can change or including in the time period immediately prior to the birth. In some instances this may suddenly restrict birth options for women, for example in the case of a medical contraindication to a Trial of Labour, clinical indications of placenta praevia. Although the issue of medical complication was not evident in the available medical record data for the participants in this study who experienced situations where they did not receive the birth choice they had wanted, it still raises the notion that perhaps an informed choice must also be a flexible one in the context of childbirth. The value in being well informed and having knowledge about the risks and benefits of the various birth options, and being in a position to adapt to changes in circumstance or the provision of new information, could enable women to move from one mode of birth to the other without affecting other decision factors such as increasing levels of decisional conflict or reduced satisfaction with the birth experience. This is an issue worthy of future research in terms of the effect of decision-aids on enabling consumers to adapt to changes in decision context and personally desired outcomes.

#### Consumer Participation in Healthcare Decision-Making

The current findings regarding the effectiveness of the decision-aid in the process of healthcare decision making are consistent with the decision-aid literature in terms of improvement in knowledge and reduction in decisional conflict (O'Connor AM et al., 2002). The notion that this is clinically significant cannot however be confirmed in the context of making choices about birth after caesarean. It is not enough to merely inform and increase comfort with choice without increasing the capacity to follow through with actual choices made. This research has not established that a decision-aid improves long-term persistence with choices because the decision-aid in itself does not tackle the complex decision environment of obstetric care. Women are themselves aware that their practitioners attempt to influence their choices and often allow their biases to show. This leads to frustration and uncertainty for those women who wish to challenge and negotiate with caregivers.

The following quotation from Survey 3 illustrates the impact that such frustration can have on individual women who do not feel support in their choice for birth...

"I was sure I was going to trial VBAC. During my last appointment the Dr stated that I would be having a planned caesarean...I cried for about 24hrs after this appointment and now can hardly talk about the uncertainty..." (#1047 Control)

This woman actually achieved a normal vaginal birth and at Survey 4, upon reflection...

"...I demanded he re-read his notes from my last labour... it helped work out what went wrong last time (my son was transverse)...This labour felt good. There was pain but it was purposeful...I can't believe what a difference a good start makes to recovery and coping with a new baby ...I was well informed and felt confident about me and my body's ability to do this...which was why I felt able to challenge the doctor..."(#1047 Control)

It must be recognised that 'pressure' is not only directed from caregivers but family and friends. The degree of support for decision-making from those who are 'significant others' clearly has an impact on decision-making as well...

"I feel extremely satisfied and proud of myself. I feel very glad and pleased that I stuck to my birth choice as I did get a lot of pressure and negativity from family and friends about my decision to have a vaginal birth" (#2030 Decision-aid)

The extent to which family and friends contribute to decisions made about birth is not quantified in the study, although it is an important issue to consider. The decision-aid in some cases assisted participants to deal with non-supportive friends and relatives. The concept of feeling supported in making a decision features in the DCS, however this was the dimension where least impact was made by the decision-aid compared with other dimensions such as 'informed'.

#### Limitations of the Research

Sample size for the study was calculated for the main outcome measures associated with knowledge and preference for mode of birth. In the process of analysis of a number of other issues identified as important in this discussion small numbers may have limited the strength of analysis.

It was valuable that all women in the intervention group received their decision-aid and were informed about how to use it, given previous difficulties experienced by other researchers where practitioners were inconsistent in supplying educational material to participants (O'Caithain, Walters, Nicholl, J., & Kirkham, 2002). It was not possible to detect the extent to which women used the decision-aid booklet or whether they shared this information with other women they knew in the clinical trial. Information from surveys collected during the study did not suggest that contamination had occurred, but this was still possible and a limitation of a design based on randomisation within research sites rather than of research sites.

There was no validated tool for assessment of knowledge for the specific decision-aid developed within this study, and the issue of birth after caesarean. For this reason, the knowledge survey was developed simultaneously with the decision-aid and based specifically on content included in the decision-aid. At the time the study was designed, this was the recommendation of the Ottawa Health Decision Centre developers in terms of facilitating content validity. The knowledge survey was based on the format of generic knowledge surveys utilised by Annette O'Connor and associates at the Ottawa Health Decision Centre, and those used in assessment of Hormone Replacement Therapy research (O'Connor AM et al., 1998).

The knowledge survey was piloted at the time the decision-aid was piloted and the survey refined to more closely represent the content of the booklet. Cronbach's alpha was calculated for both Survey 2 and Survey 3 results of the knowledge test as an estimation of internal consistency. Although this was thought to be satisfactory for the purposes of this research, it would not be immediately transferable to other research on decision-aids. Therefore the knowledge survey developed for this research would be of limited general value for assessment of other decision-aids in the future. There is still no generic knowledge test for decision-aids because each decision-aid will have its own content and it will still be a focus of information are important to check for understanding using such a test.

The measurement of knowledge was a crucial element of the research, in terms of decision-aid effectiveness, and therefore caution regarding the clinical significance must be raised. The clinical impact of creating a capacity to answer questions correctly may be insignificant when compared to the impact of other powerful factors in decision-making and will have little relevance in the future if environmental and cultural factors remain the same.

In answering research questions associated with the process of decision making for women about birth after CS, the first underlying assumption was that women were in fact making a choice. The notion of 'consumer demand' for CS for instance raises the picture of women empowered to state their choice and doctors or midwives ready to comply. Although the literature provides a mixed account of women's involvement in decision-making about mode of birth, recommendations from midwifery and obstetric professions, as well as various government organisations, to facilitate informed choice (American College of Obstetricians and Gynecologists, 1999; National Health and Medical Research Council, 1996; Senate Community Affairs References Committee, 1999), reinforce the notion that most women are actively involved in decision processes about childbirth. Recommendations for shared decision-making or partnership in decisions about pregnancy and childbirth are embedded within clinical competency documents for Australian midwives (Australian Nursing and Midwifery Council, 2006). The extent to which recommendations mirror clinical reality remains open to discussion.

# CONCLUSION

Decision-aids as 'stand-alone' aids can contribute to decision-making through increasing knowledge, reducing uncertainty and assisting women to put the pros and cons into perspective according to important personal factors such as past experience, values and needs. This however is of limited value if the practitioner is not prepared to practice on the basis of an inclusive consumer relationship. If the expectations of the consumer differ to those of the practitioner, then this imbalance creates a situation of challenge for both parties. Therefore strategies are required that can assist practitioners to develop skills in decision-making alongside their consumers. Clinical Practice Guidelines such as those developed on Caesarean Section in the UK will help inform practitioners as well as consumers regarding risks and benefits of birthing options. (National Institute of Clinical Excellence (NICE), 2004). However without an effective strategy in place, the informed part of decision-making is of little value. Interactive computer-based decision-aids that can involve both consumers and their practitioners at the same time, rather than separately, may deserve future research attention. A series of paper-based practitioner worksheets and associated resources are already being developed by the Ottawa Health Decision Centre (A. O'Connor & Jacobsen, 2004) in response to a need for practitioner support. The problem that remains relates to the desire of practitioners to utilise such resources and the approaches practitioners develop to deal with situations where their beliefs conflict with those of the consumer and where medico-legal conflicts of interest can be anticipated.

Partnership approaches to care, although articulated in policy documents and professional codes of practice, are difficult to achieve without effective education and training. Even a commitment to the principles of informed choice will not equate to providing informed choices. Obstetricians and midwives in this study verbalised support for the notion of informed choice for women, yet women themselves revealed that they frequently felt that they did not have a choice or that they were being influenced by the individual biases of their practitioners.

Patient participation in healthcare decision-making has become an issue for debate amongst a broad range of nursing and medical literature. It is interesting that the focus has been on the role of the 'patient', whether or not they wish to be engaged in the decisionmaking about their healthcare and if they are in fact equipped to do so in any meaningful way. There is an acknowledgement, in upholding the principles of "self-determination', that patient participation is justified in the context of healthcare, and in this, individual preferences about their expected involvement needs to be elicited by the practitioner (Guadagnoli & Ward, 1998). It is important to acknowledge that there are different types of patients who display differing levels of interest in the degree to which they want to be involved in making medical decisions. However we would extend upon this notion and argue that there are also different types of practitioners. The argument to devise strategies to determine the readiness of patients to participate in decisions should also include strategies to determine the readiness of practitioners to engage in this process. In assessing what constitutes 'patient participation' for an individual doctor-patient or nurse/midwife-patient dyad, the assessment and process to follow should consider the attributes and expectations of both parties in the decision 'partnership'. Incompatible expectations may even lead to a decision that a different practitioner should be sought by the 'patient'.

#### Is it Possible to Change the Decision Environment?

The consumer decision-aid concept, as a strategy to facilitate informed choice, is not enough on its own. Attitudes of practitioners toward women's ability to participate in birth choice appear to vary and this has an effect on the level of involvement women can expect during their birthing decisions. Attitudes however can change and perhaps practitioners, rather than being expected to comply with policy, need to take a more active part in the process. During the RCT, overt antagonism toward the study by obstetricians, who were sceptical about women making a choice or of VBAC itself, often changed as they became more familiar with the purpose of the decision-aid.

How can the benefits of a decision-aid such as this be maximised in the current healthcare environment? Given the results of the UK study using 'informed choice' leaflets (O'Caithain et al., 2002; Stapleton et al., 2002), where women who needed information resources were not uniformly offered them, it is important that we neither rely on the practitioner nor the consumer alone to ensure such strategies are implemented. The nature of busy prenatal clinics and private obstetric services means that resources such as decision-aids may be used intermittently and infrequently unless there is a process that integrates such tools into the consultation format or clinical practice guidelines. The best way to facilitate this is yet to be determined, but is the next step in researching the effective incorporation of decision-aids into clinical practice.

#### REFERENCES

- Abitbol, M. M., Castillo, I., Taylor, U. B., Rochelson, B. L., Shmoys, S., & Monheit, A. G. (1993). Vaginal birth after caesarean section: The patients point of view. *American Family Physician*, 47, 129-134.
- American College of Obstetricians and Gynecologists. (1999). ACOG practice bulletin Vaginal birth after previous cesarean delivery. *International Journal of Gynecology and Obstetrics*, 64(October), 201-208.
- Amu, O., Rajendran, S., & Bolaji, I. (1998). Maternal choice alone should not determine method of delivery. *BMJ*, 317(August 15), 463-465.

- Appleton, B., Targett, C., Rasmussen, M., Readman, E., Sale, F., Permezel, M., et al. (2000). Vaginal birth after caesarean section: An Australian multicentre study. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 40(1), 87-91.
- Arrow, K. (1965). Aspects of the theory of risk-bearing. Helsinki: Yrjo Jahnsson Saatio.
- Arrow, K. (1983). Behavior under uncertainty and its implications for policy. Stanford: Stanford University.
- Australian Nursing and Midwifery Council. (2006). National Competency Standards for the Midwife. Dickson: ANMC.
- Beach, B., & Beach, L. (Eds.). (1982). Expectancy-based decision schemes:sidesteps toward applications. New Jersey: Lawrence Erlbaum Associates.
- Breemhar, B., & van den Borne, H. (1991). Effects of education and support for surgical patients: the role of perceived control. *Patient Education and Counseling*, *18*, 199-210.
- Charles, C., Whelan, T., & Gafni, A. (1999). What do we mean by partnership in making decisions about treatment? *British Medical Journal*, *319*, 780-782.
- Consumer Focus Collaboration. (2001). *The Evidence Supporting Consumer Participation in Health*. Bundoora: LaTrobe University.
- Coulter, A. (1999). Paternalism or partnership? British Medical Journal, 319(7212), 719-720.
- Department of Health. (1993). Changing Childbirth Report. London: HMSO.
- Eden, K., Hashima, J., Osterweil, P., Nygren, P., & Guise, J. (2004). Childbirth preferences after cesarean birth: A review of the evidence. *Birth*, *31*(1), 49-60.
- Feather, N. (1982). Human values and the prediction of action: an expectancy-valence analysis. In F. N (Ed.), *Expectations and actions: expectancy-value models in psychology*. Hillsdale New Jersey: Lawrence Erlbaum Associates.
- Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behavior: an introduction of theory and research. Reading: Addison-Wesley.
- Flamm, B. L., Goings, J. R., Liu, Y., & Wolde-Tsadik, G. (1994). Elective repeat caesarean delivery versus trial of labour. *Obstetrics and Gynecology*, 83, 927-932.
- Fraser, W., Maunsell, E., Hodnett, E., Mountquin, J. M., & Childbirth Alternatives Post Caesarean Study Group. (1997). Randomised controlled trial of a prenatal vaginal birth after cesarean section education and support program. *American Journal of Obstetrics* and Gynecology, 176, 419-425.
- Graham, I., & O'Connor, A. (1999). Satisfaction with preparation for decision-making. In A. O'Connor (Ed.), *Evaluation Measures*. Ottawa: University of Ottawa.
- Guadagnoli, E., & Ward, P. (1998). Patient participation in decision-making. Social Science Medicine, 47(3), 329-339.
- Hamilton, B., Martin, J., & Ventura, S. (2006). Births: Preliminary data for 2005. Health E-Stats: National Centre for Health Statistics.
- Health Department of Victoria. (1990). *Having a Baby in Victoria: Ministerial Review into Birthing Services in Victoria.* Melbourne: Victorian Government.
- Horey, D., Weaver, J., & Russell, H. (2004). Information for pregnant women about caesarean birth. *The Cochrane Database of Systematic Reviews*(1).
- Janis IL, & Mann LI. (1977). Decision Making: A Psychological Analysis of Conflict, Choice and Committment. New York: Free Press.
- Janis, I. L., & Mann, L. I. (1977). Decision Making: A Psychological Analysis of Conflict, Choice and Committment. New York: Free Press.

- Joseph, G. F., Stedman, C. M., & Robichaux, A. G. (1991). Vaginal birth after caesarean section: The impact of patient resistance to a trial of labour. *American Journal of Obstetrics and Gynecology*, 164, 1441-1447.
- Kirkham, M., & Stapleton, H. (2004). The culture of the maternity services in Wales and England as a barrier to informed choice. In M. Kirkham (Ed.), *Informed choice in maternity care* (pp. 117-145). Houndmills: Palgrave.
- Laws, P., Grayson, N., & Sullivan, E. (2006). Australia's Mothers and Babies 2004 (No. 18). Sydney: AIHW National Perinatal Statistics Unit.
- Laws, P., & Sullivan, E. (2004). *Australia's Mothers and Babies 2002* (No. 15). Sydney: Australian Institute of Health and Welfare National Perinatal Statistics Unit.
- Lee, W. (1971). Decision Theory and Human Behavior. New York.
- Martel, M., MacKinnon, C., & SOGC. (2005). Guidelines for vaginal birth after previous caesarean birth: Society of Obstetricians and Gynaecologists of Canada. *Journal of Obstetrics and Gynaecology Canada*, 27(2), 164-174.
- Martin, J., Hamilton, B., Sutton, P., Ventura, S., Menacker, F., & Kirmeyer, S. (2006). Births Final Data for 2004, National Vital Statistics Reports. Hyattsville: National Centre for Health Statistics.
- Martin, J., Hamilton, B., Sutton, P., Ventura, S., Menacker, F., & Munson, M. (2005). Births Final Data for 2003, National Vital Statistics Report: . Hyattsville: National Centre for Health Statistics.
- McClain, C. S. (1985). Why women choose trial of labor or repeat cesarean section. *The Journal of Family Practice*, 21(3), 210-216.
- McClain, C. S. (1987). Patient decision making: The case of delivery method after a previous cesarean section. *Culture, Medicine and Psychiatry, 11*, 495-508.
- McClain, C. S. (1990). The making of a medical tradition: Vaginal birth after cesarean. *Sco Sci Med*, *31*(2), 203-210.
- McClain CS. (1983). Perceived risk and choice of childbirth service. *Social Science and Medicine*, 17(23), 1857-1865.
- McMahon, M. J., Luther, E. R., Bowes, W. A., & Olshan, A. F. (1996). Comparison of a trial of labour with an elective second caesarean section. *New England Journal of Medicine*, 335, 689-695.
- Molenaar, S., Sprangers, M., Postma-Schuit, F., & Rutgers, E. (2000). Feasibility and effects of decision aids. *Medical Decision Making*, 20(1), 112-127.
- Mould, T. J., Chong, S., Spencer, J. D., & Gallivan, S. (1996). Women's involvement with the decision preceding their caesarean section and their degree of satisfaction. *British Journal* of Obstetrics and Gynaecology, 103, 1074-1077.
- Nassar, N., & Sullivan, E. A. (2001). *Australia's Mothers and Babies 1999*. Sydney: Australian Institute of Health and Welfare National Perinatal Statistics Unit.
- National Health and Medical Research Council. (1995). *Guidelines for the development and implementation of clinical practice guidelines*. Canberra: AGPS.
- National Health and Medical Research Council. (1996). *Options for Effective Care in Childbirth*. Canberra: Commonwealth of Australia.
- National Health and Medical Research Council. (2000). *How to present the evidence for consumers: preparation of consumer publications*. Canberra: Commonwealth of Australia.

- National Institute of Clinical Excellence (NICE). (2004). *Clinical Guideline 13: Caesarean Section (No. 13)*. London: NICE.
- Nessa, J., & Malterud, K. (1998). Tell me what's wrong with me: a discourse analysis approach to the concept of patient autonomy. *Journal of Medical Ethics*, 24(6), 394-400.
- Norman, P., Kostovcik, S., & Lanning, A. (1993). Elective repeat cesarean section: How many could be vaginal births? *Canadian Medical Association Journal*, *149*, 431-434.
- O'Caithain, A., Walters, S. J., Nicholl, J. P., J., T. K., & Kirkham, M. (2002). Use of evidence based practice leaflets to promote informed choice in maternity care: randomised controlled trial in everyday practice. *British Medical Journal*, 324, 1-5.
- O'Connor, A., Drake, E., Fiset, V., Graham, I., Laupacis, A., & Tugwell, P. (1999). The Ottawa patient decision aids. *Effective Clinical Practice*, 2(4), 163-170.
- O'Connor, A., & Jacobsen, M. J. (2004). *Decisional conflict: Assessing and supporting patients experiencing uncertainty about choices affecting their health*. Retrieved 14-2-05, from http://decisionaid.ohri.ca/docs/Training/Decisional\_Conflict.pdf
- O'Connor, A., Rostom, A., Fiset, V., Tetroe, J., Entwistle, V., Llewellyn-Thomas, H., et al. (1999). Decision-aids for patients facing health treatment or screening decisions: systematic review. *British Medical Journal*, *319*(18 September), 731-734.
- O'Connor, A. M. (1995). Validation of a decisional conflict scale. *Medical Decision Making*, 15, 25-30.
- O'Connor, A. M. (1997). Consumer/patient decision support in the new Millenium: Where should our research take us? *Canadian Journal of Nursing Research*, 29(3), 7-12.
- O'Connor, A. M. (1999). Decision Conflict Scale. Ottawa: University of Ottawa.
- O'Connor A.M., Stacey D, Rovner D, Holmes-Rovner M, Tetroe J, Llewellyn-Thomas H, et al. (2002). Decision aids for people facing health treatment or screening decisions. [Systematic Review]. Cochrane Consumers & Communication Group Cochrane Database of Systematic Reviews.(3).
- O'Connor A.M., Tugwell P, Wells GA, Elmslie T, Jolly E, Hollingworth G, et al. (1998). A decision aid for women considering hormone therapy after menopause: decision support framework and evaluation. *Patient Education and Counseling*, *33*, 267-279.
- O'Connor, A. M., Stacey, D., Rovner, D., Holmes-Rovner, M., Tetroe, J., Llewellyn-Thomas, H., et al. (2002). Decision aids for people facing health treatment or screening decisions. [Systematic Review]. Cochrane Consumers & Communication Group Cochrane Database of Systematic Reviews.(3).
- O'Connor, A. M., Tugwell, P., Wells, G. A., Elmslie, T., Jolly, E., Hollingworth, G., et al. (1998). A decision aid for women considering hormone therapy after menopause: decision support framework and evaluation. *Patient Education and Counseling*, 33, 267-279.
- O'Connor, A. M., Tugwell, P., Wells, G. A., Elmslie, T., Jolly, E., Hollingworth, G., et al. (1998). Randomized trial of a potable self-administered decision aid for postmenopausal women considering long-term preventive hormone therapy. *Medical Decision Making*, 18(3), 295-303.
- Ottawa Health Decision Centre. (2001). Guidelines for developing quality patient decision aids according to the pre-set criteria in the Cochrane review. Ottawa: Ottawa Health Research Institute.
- Perveen, F., & Shah, Q. (1997). Obstetric outcomes after previous caesarean section. Journal of Obstetrics and Gynaecological Research, 4, 344-346.

- Ratliff, A., Angell, M., Dow, R. W., Kuppermann, M., Nease, R. F., Fisher, R., et al. (1999). What is a good decision? {comment}. *Effective Clinical Practice*, 2(4), 185-197.
- Senate Community Affairs Reference Committee. (1999). *Rocking the Cradle: A report into childbirth procedures*. Canberra: Parliament of Australia.
- Shorten, A., Chamberlain, M., Shorten, B., & Kariminia, A. (2004). Making choices for childbirth: development and testing of a decision-aid for women who have experienced previous caesarean. *Patient Education and Counseling*, 52, 307-313.
- Shorten, A., Lewis, D. E., & Shorten, B. (1998). Trial of Labour versus elective repeat caesarean section: A cost-effectiveness analysis. *Australian Health Review*, 21(1), 8-28.
- Shorten, A., Shorten, B., Keogh, J., West, S., & Morris, J. (2005). Making choices for childbirth: a randomized controlled trial of a decision-aid for informed birth after cesarean. *Birth*, 32(4), 253-262.
- Shorten, B., & Shorten, A. (2005). *The public to private hospital shift: what happens to birthing women?* Paper presented at the Health Outcomes Conference, Canberra.
- Stapleton, H., Kirkham, M., & Thomas, G. (2002). Qualitative study of evidence based leaflets in maternity care. *British Medical Journal*, 324, 639.
- Thomas, J., Paranjothy, M., & Royal College of Obstetricians and Gynaecologists Clinical Effectiveness Support Unit. (2001). National Sentinal Caesarean Section Audit Report. London: Royal College of Obstetricians and Gynaecologists.
- Turnbull, D., Wilkinson, C., Yaser, A., Carty, V., Svigos, J. M., & Robinson, J. S. (1999). Women's role and satisfaction in the decision to have a caesarean section. *Medical Journal of Australia*, 170, 580-583.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453-458.
- Walker, R., Turnbull, D., & Wilkinson, C. (2004). Increasing cesarean section rates: exploring the role of culture in an Australian community. *Birth*, *31*(2 June), 117-124.
- Wallace, M., Shorten, A., & Russell, K. (1997). Paving the way: Stepping stones to evidencebased nursing. *International Journal of Nursing Practice*, 3, 147-152.
- Wittmann-Price, R. (2004). Emancipation in decision-making in women's health care. *Journal of Advanced Nursing*, 47(4), 437-445.

Chapter 35

## LEARNING, ACCEPTANCE AND THE PERCEPTIONS OF RISKS AND BENEFITS OF NEW TECHNOLOGY DEVELOPMENTS IN SPAIN\*

## Montse Costa<sup> $\dagger 1$ </sup> and Joan Costa-Font<sup>1,2</sup>

<sup>1</sup>London School of Economics, UK <sup>2</sup>Departament de Teoria Economica, Universitat de Barcelona, Spain

## ABSTRACT

The acceptance of risks associated with new technologies is a key issue that is likely to limit the extent of innovation in a 'risk society'. However, given the limited comprehensible information available to the public of new technologies, it is likely that risk information provision will have a heterogeneous effect on public perceptions. In order to examine this issue, we empirically examine the determinants of risk perceptions, benefit perceptions and risks acceptance of new technology developments in Spain. Our findings indicate that risk and benefits perceptions are not independent but affected by common information sources. Furthermore, by taking into account this effect individual's knowledge of science heterogeneously increases both risks and benefits perceptions.

Keywords: risks perceptions, benefits perceptions, bivariate probit, Spain.

## **1. INTRODUCTION**

The diffusion of the potential risks and benefits of new technology developments products (e,g., products resulting from genetic manipulation of crops and animals), stands in

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Economics, Business and Finance*, edited by Klaus P. Hofmann published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Contact address: Dr Joan Costa Font, LSE Health and Social Care, Cowdray House, London School of Economics, Houghton Street, WC2A 2AE, London. E-mail: j.costa-font@lse.ac.uk.

forefront of the current policy disputes in Europe. Important skepticism in the acceptance of some new technology applications has centered the debate on science communication in the European Union that tackles an increasing media audience. However still small is known on *how individuals learn on new science developments* as well as on the understanding of how individuals perceive (or mentally weight) information signals as being beneficial and /or prejudicial for their health, the environment or society as a whole. The demand for new methods of risk communication challenges how innovation should be disseminated and applied with a 'knowledge economy'. Indeed, new products penetrate the supply chain by improving some qualities of existing ones however pose some concerns regarding potential side effects to human health and the environment although often unproven. In addition to those risks, new technologies can potentially changes people's lifestyles (e.g., mobile phones, etc). Whilst some people might welcome such changes and envisage them as benefits, others might perceive them as harmful and risky.

Individuals both as societal stakeholders as well as consumers often lack sufficient information to make decisions on new technologies and/or information available is often provided by untrustworthy channels e.g., though corporate marketing campaigns. Often not even publicly provided information is trusted as a result of recent food scandals in Europe. According to Sheehy et al (1998) even highly educated individuals exhibit small knowledge in some areas of science and technology. Partial (dis) information on the benefits and risks of new products often casts an emotional response that leads individuals to exercise their 'exit voice' by the (dis) approving the commercialisation of new technologies. However, an appropriate understanding of how attitudes are form in new technology developments is important both for private and public bodies to design communication strategies.

Some studies have shown that although consumer has small information they view genetic engineering as a risky process (Wohl, 1998). People perceive environmental as well as safety risks. Furthermore, often risks rather than being objectively measurable are qualitative categorization resulting from lack of sufficient information and knowledge, and thus stand as unknown risks. Therefore, the decision on whether to consume certain goods produced though new technology developments results from benefits out weighting the risks.

In a world of perfect information, real and perceived risks arguably would perfectly comparable. However, even when information is available, normally individuals information update comes at some cost and, still education places and important barrier to entry for some individuals to understand how certain technologies improves individuals well being. Therefore, human beings are not always open to update their knowledge and even when they do so they might be selective in which information they collect and thus information will rely on the importance they confer to certain information channels. Therefore, role of information channels and especially the media place important effects in explaining the paths that guide individual's attitudes and perceptions towards new science applications. Given individuals lack of information (which exhibit known risks) consumption decisions are likely to be influenced by external aspects of individual perception. Indeed, it has been widely demonstrated that individuals exhibit some aversion to "the unknown" (Ellsberg, 1961). If this stands in the biotechnology area, the we might expect that people would value higher prospects involving known risks even when potential benefits are larger in alternative protects involving unknown risks. Therefore, knowledge is expected to play a key role in determining the extent to which individuals perceive the risks and benefits of certain technologies, and consequently on how they react to their exposure.

Public resistance to new technologies might harm growth and trade and thus, society will be excluded of potential benefits. It has been suggested that in order to increase acceptance by improving the benefits and reducing potential risks (Franks, 1999). Industry and the government need to have a better understanding of consumer's acceptance of biotechnology as far as " the full benefits of biotechnology will only realize if consumers and the food industry accept the use of these new technologies as safe and beneficial" (Hoban 1988). However, risks and benefits, which are supposed to guide attitudes for new biotechnology, cannot often be regarded as independent. It might be the case that certain risks are perceived as more intense due to the lack of benefits, or the other way around, lack of benefits results from small benefits and high risks. Kanheman and Tversky (1986) associate perception of risks and benefits by suggesting that people are risks averse when the outcome of a decision is perceived as a benefit rather than reduction of a loss. Therefore, a second important research question refers to testing whether risks and benefits perceptions of new technologies are impendent, and whether the determinants of risks and benefits differ when risks and benefits are examined simultaneously. An interesting country, which has been relatively less examined, is that of Spain. Spain is a southern European country that has progressively adapted to European core values although still high religiosity is highly influential (Inglehardt and Baker, 2000). However, the extent of modernization and dissemination of new technologies has taken place must faster than in center European countries (Atienza and Lujan, 1997).

People perceptions related with technological developments are likely to depend on the type and level of information individuals handle, which determine the extent which individuals are susceptible to alarms and exert some outrage reaction (Standman, 1992)". Prior research indicates that people tend to overestimate unfamiliar, less known, unfair, involuntary and artificial risks, as well as the one they have no control, the ones extremely publicized and the ones where enhancing moral underlying concerns.

This paper aims at providing some understanding of the risks and benefits perception, and therefore the acceptance of new technologies. We undertake and empirical analysis of survey data on the Spanish perceptions to new technology developments. First, we argue that due to the significant lack of information associated with new technology developments lead to risks and benefits perceptions to be influenced by similar underlying effects. Second, knowledge exerts important role as increasing the benefits perceptions and reducing the risks of technology developments. Finally, we examine the determinants of risks and benefits perceptions, and most notably the role of information acquisition and socioeconomic determinants.

The structure of the paper is the following. Next we provide a theoretical discussion on the independence of risks and benefits perceptions and issues concerning the empirical specification of the model on the determinants of risk and benefits perceptions. Next section describes the data and the preliminary evidence. Section four deal with results and section five concludes.

## 2. TECHNOLOGY ACCEPTANCE AND THE INDEPENDENCE OF RISK AND BENEFITS PERCEPTIONS

#### 2.1. The Theoretical Background

Yet, under imperfect information, individuals are likely to gather messages from several sources  $(\eta_1, ..., \eta_n)$  that convey the following information of risks in terms of a probability of a hazard  $(\pi_1, ..., \pi_2)$ , being the potential loss  $(L_1, ..., L_n)$  and the potential utility improvement  $(B_1, ..., B_n)$ . Individuals in deciding on new technologies consumption, they are assume to take into account all this information which is summarized in the net perceived utility of each product  $(P_i)$  subject to the existing public and private information  $(I_i)^I$ . Yet, in the event of potential harmful effects in the future, consumer's current utility can be summarized in two main arguments: H (that refers to health), and X (consumption of other goods) such as  $U_i = U_i(X, H)$ , such that  $\frac{\partial U}{\partial X} \ge 0$  and  $\frac{\partial U}{\partial H} \ge 0$ . On the other hand, future (expected) future utility can be conceptualised as  $EU_i = \pi_i U(X, H) + (1 - \pi_i)U(X, H - L_i)$ , whereby consumption results if the net expected tulity is positive, or in other words if the potentials benefits - both for health and consumption - assoicated with the consumption of a certain product overome the potential risks.

However, as far as consumers handle a limited amount of information on the possible consequences of certain events, potential information provision might exert an impact on individual's evaluation of the net expected utility of consumption. This effect takes place through a process of weighting new information, either public of private by sources  $\gamma_j$  determining the perception of risks ( $\pi_i = \gamma_0 I_0 + ... + \gamma_n I_n$ ) and perceptions of benefits ( $B_i = \gamma_0 I_0 + ... + \gamma_n I_n$ ). This process of information updating resembles the one of the Bayesian learning models (Viscusi, 1992), because information sources might affect both decisions at a time, one might argue that perceptions of risks and benefits are liable to be associated, and thus are nor independent as one might theoretically conceive.

#### 2.2. The Emprirical Specification

The decision to consume a certain new product following section 2.1 results from individuals' risks acceptance, or in other words, from perception of benefits overcoming risks. Let us specific the perception of risks as follows:

<sup>&</sup>lt;sup>1</sup> Availability of information available can be conceptualized as "signals" that might lead to some individuals to subjectively become aware of some possible "individual or societal benefits/looses".

$RISK^*_i = X_i\beta_i + u_i$		
$RISK_i = 1 \ if \ RISK_i^* > 0$	(1)	
$RISK_i = 0 \ otherwise$		

where RISK\* is not observable but only where the individual responds to a survey question on whether he perceives risks on science innovation. Similarly, as benefit perceptions (BP\*) are not observable but we only observe whether a dichotomous variable based on individuals responses as follows:

$BP^*_i = Z_i \delta_i + \varepsilon_i$	
$BP_i = 1 if BP_i^* > 0$	(2)
$BP_i = 0$ otherwise	

As is common practice for a probit model we assume that the errors are distributed N(0,1) and the two models' errors are independent of one another, so that  $\text{Cov}(u_i, \varepsilon_i) = 0$ . However, it might well be that  $u_i = \eta_i + v_i$  and  $\varepsilon_i = \eta_i + \omega_i$ , so that the errors in each model consist of a part  $v_i, \omega_i$  that is unique to that model, and a second part  $\eta_i$  that is common to both. If this is the case, the error terms are likely to be dependent. Thus, we're interested in the joint probability we use a bivariate normal distribution  $(u_i, \varepsilon_i) \sim \text{BiN}(0,0,1,1, \cdot)$ ; whereby  $\rho$  is a correlation parameter denoting the extent to which the two error term covary. Finally, a decision that stands as key information for risk policymaking is that of risk acceptance. Again, because net utility is unobservable (NU\*), we only observe whether individuals perceive that the benefits of science innovation overcome the risks as follows:

$NU_{i}^{*} = H_{i}\omega_{i} + \mathcal{G}_{i}$	
$NU_i = 1 if NU_i^* > 0$	(3)
$NU_i = 0$ otherwise	

## **3.** THE DATA AND PRELIMINARY EVIDENCE

The data that we employ in this study is gathered from public sources, and in particular it refers to a survey commissioned by the Spanish Centre for Sociological Research in 1996 on 'Attitudes towards Scientific and Technology Innovation'. This is a representative survey of the Spanish population between 18 and 64 years of age. The sample initially was made of 2552 respondents; it was personal interviews to individuals from 91 municipalities and 43 provinces. The variable definition and descriptive statistics are explain are in Table 1. As expected the vast majority of the Spanish population beliefs on the 'expert opinions', only 12% is techno-sceptic and knowledge of science. About 45 trusts the state ad doing a good

job, half of the same is made up of male and women, mean age is 41 years and 57% are married. Only 23% regard themselves as right wing, 2% practices religion and 37% is a family head.

Perceptions of risks (benefits) were measured from the response to the following question: "Do you expect in the next 20 years the technology developments to bring many/some/few/ no risks (benefits) to society?". This question provides in formation on the two hypothetical dimensions guiding individual acceptance of new science technologies. On the other hand, the survey includes another question, which goes like this: "Do you think that benefits of new technologies overcome the risks" (Yes/No).

Variable	Definition		Mean	s.e
Experts	Do not Belief on experts	D	0.79	0.01
Tecno	Belief that technology will not improve way of life	0		
			0.12	0.01
Know	Knowledge level <sup>2</sup>	С	6.96	0.05
estate	Government Trust =1	D	0.45	0.01
gender	Female==1	D	0.50	0.01
Age	Age in Years	С	41.1	0.03
Married	Married ==1	D	0.57	0.01
Politic	Leftist=1	С	0.23	0.74
Practice	Follows a religion	D	0.02	0.01
headfam	Head of household	D	0.37	0.01

#### Table 1. Variables definition

# Table 2. Perception of risks and benefits in the next 20 years (%), as a result oftechnological and scientific advances

Perception	Risks	Benefits	
Many	16.69	21.69	
Some	40.24	52.24	
Few	28.21	13.06	
None	4.11	2.98	
n.s.	10.70	9.57	
n.c.	0.04	0.47	

Table 2 measures the perception of risk and benefit of the technological and scientific advances. Interestingly, the vast majority of the population perceived both risks and benefits associated with technology innovation, although overall the share of those that perceive benefits overcomes that of risks. However, a sizeable share of the population perceives some risks which might arguably result from some resistance to new technologies, or an alternatively from the ignorance on the effects of new technologies, which are noted might lead to the so called 'risk ignorance or ambiguity aversion'. Furthermore, on the question of

<sup>&</sup>lt;sup>2</sup> Derived from 12 answers to questions on science and technology (information provided upon request).

nologies overcome the risks, 57% agreed this the assertion,

whether benefits of new technologies overcome the risks, 57% agreed this the assertion, indicating that although some individuals perceive technology related risks, the potential benefits seem to be larger.

In order to find some explanation to this evidence we examine further data. Interestingly, only the 57% of the population trusts on new technologies and 13% argue that new technologies cannot solve the problems of the oldest ones, 36% respond that life without technologies would be better, which could be regarded as 'techno skeptics'. 65% believe that the decisions related to technology cannot be exclusively based on consumer's knowledge but that need the intervention of experts. However, 92% agree that science and technology will on the whole improve the population quality of life. This, there is certainly some ambivalence in individuals responses to 'positive' and 'negative' dimensions of science developments. The clearest explanation relies in that science and technology might produce heterogeneous effects, both positive and negative, and that question on risks and benefits might stress specific effects. Accordingly, on the basis of our results, individuals in thinking on new technologies might suffer from a 'perceived aggregation effect'. Another explanation might lie in that individuals might be sensitive to the provision of information and might exhibit specific framing effects. However, given the general nature of the question and the fact that other survey questions referred to a vast array of technologies such as computers, cloning, space exploration, solar energy etc there is no reason to suggest that responses would be biased in some specific way.

### **4. RESULTS**

This section reports the results of the different specification of the determinants of risks and benefits perceptions of new scientific innovation in Spain.

#### 4.1. Risk and Benefit Perception

Let us begin with the determinants of risks perceptions by examining the significance of the estimated coefficients. As expected, individuals that are less likely to trust expects would perceive larger risks and the same applied for those individuals that exhibit a pessimistic belief on science and technology – those that perceive that technology will not improve way of life-. Thus, consistently with some previous literature risks perceptions convey relevant information on individuals trust although trust in the government did not place an effect on risks perceptions (Table3)<sup>3</sup>. Knowledge of science was significantly associated with lower risks perceptions, thus providing some support of 'ignorance aversion', whereby the lower the individuals knowledge the more likely she will be to perceive high science-related risks.

Interestingly, age does not exert a significant effect that indicates individuals of different ages, which are potentially exposed to different intensity of information acquisition, do not perceive risks systematically differently. On the other hand, Table 3 indicates that women are more likely to perceive risks than men, which is consistent with previous literature. Finally,

<sup>&</sup>lt;sup>3</sup> We have tested the existence of multi-colinearity amongst regressions by performing cross-correlation, but all of them were lower than 0.30.

whilst political orientation and religious are not significant predictors of technology risks perceptions, being married does exert a negative effect whilst family heads ado not lead to higher risks perceptions. The finding that married people would perceive fewer risks on technology developments might have to do with the fact that they are likely to have children that might benefit from future technology developments.

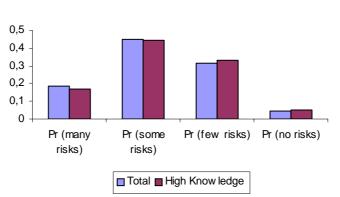
	Risk P	Risk Perception		Benefit Perception		Risk Acceptance	
	Coeff.	s.e	Coeff.	s.e	Coeff.	s.e	
NoExperts	0.262**	0.057	-0.189**	0.059	-0.284**	0.067	
Tecno	-0.702**	0.072	0.547**	0.082	0.547**	0.083	
Know	-0.019*	0.01	0.076*	0.012	0.077**	0.013	
estate	-0.031	0.046	0.022	0.047	0.012*	0.055	
gender	0.123**	0.052	-0.112	0.053	-0.132*	0.063	
Age	0.003	0.022	0.069*	0.023	0.054*	0.026	
Married	-0.117*	0.054	0.028	0.056	-0.105	0.066	
Politic	0.001	0.001	0.051	0.001	-0.001	0.001	
Practice	0.122	0.191	-0.267	0.224	-0.118*	0.060	
headfam	-0.071	0.057	0.106**	0.057	0.026	0.069	
Intercept					0.393	0.136	
RV Chi (2,11)	140.59		164.62		137.41		
Likelihood Ratio Test	-2603.45		-2313		-1442		
Pseudo R <sup>2</sup>	0.08		0.07		0.1		
% Corr	67%		71%		89%		

Table 3. Risk perceptions (ordered probit), Benenfit perceptions (ordered probit) and
Risk Acceptance (probit model)

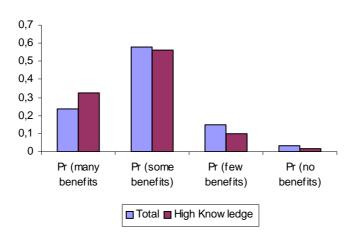
Overall, Table 3 reveals that benefits perceptions are influenced by similar variables as that of risks perceptions but they display opposite coefficients. This is the case of lack of trust in experts that is found to reduce the probability of perceiving benefits from science developments. As expected an optimistic belief on the effects of science in improving the quality of life s well as scientific knowledge increase the benefits perceptions. However, now although age does exert some positive effects, which indicate that possibly, older cohorts that have envisaged significant innovation changes tend to perceive higher benefits than younger cohorts. Gender exert just the opposite effect as in the risks perceptions case, female are less likely to perceive positive effects of new science developments. Interestingly, some political affiliation – those that regard themselves as left wing- is positively associated with larger benefits perceptions whilst those practicing a religious would perceive lesser benefits.

Figure 1 reports the predicted probabilities of each response conditioned on the knowledge effect. Interestingly, knowledge does exert some effect on reducing the probability of risk perception response and increasing the probability of a benefit perception response.

Therefore, if knowledge captures the capacity of agents of updating information on new technologies, our results indicate that pro-active information policies would have a strong impact on risks acceptance of new technologies.



**Risk perceptions** 



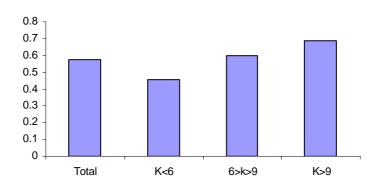
Benefit perceptions

Figure 1. Risks and benefit perception of new technologies.

#### 4.2. Risk Acceptance

According to theoretical model of previous sections, individuals in making decision on issues that convey some risks; they have to balance potential risks with accruing benefits. In Table 3, we examine using a probit model the determinants of individuals perceiving larger benefits that risks of science developments. Interestingly, trust in experts, beliefs on the potential improvements of quality of life, gender and age stand as the key risk acceptance drivers. Older individuals and especially men are more likely to accept technology related risks. Other relevant variables are that of religion, individuals practicing religion tend to be

less likely to accept technology related risks. An interesting issue, in risks acceptance is the significance of the intercept term, which according to the prospect reference theory conveys information on prior assessment or risks and benefits, which displays a positive effect. Figure 2, indicates that predicted risks acceptance increases with individuals knowledge, indicating the sensitivity of individuals attitudinal reactions to knowledge of science.



**Risk acceptance** 

Figure 2. Risk Acceptance and knowledge.

	Risk perceptions			Benenfit perceptions			
	Coef.	Std. Err.	Z	Coef.	Std. Err.	Z	
NoExperts	0.181	0.063	-2.878	-0.287	0.067	4.282	
Tecno	-0.733	0.087	-8.416	0.539	0.080	6.717	
Know	0.032	0.011	2.886	0.126	0.012	10.652	
estate	0.091	0.051	1.786	0.030	0.056	0.535	
gender	0.131	0.058	2.242	-0.071	0.065	1.087	
Age	-0.025	0.024	-1.051	0.079	0.026	3.031	
Married	0.133	0.060	2.204	-0.087	0.066	-1.317	
Politic	0.000	0.001	-0.410	0.001	0.001	1.958	
Practice	-0.238	0.205	-1.161	0.046	0.219	0.212	
headfam	0.082	0.064	1.290	-0.018	0.071	-0.256	
Intercept	0.066	0.122	0.540	-0.094	0.130	-0.722	
$\rho$	-0.20	0.01					
FV	-2.999.77						
Wald (2,22)	345.31						
RV ( $\rho$ =0)	32.45						

#### 4.3. The Independence of Risk and Benefits Perceptions

The fist results indicates that as expected both decision are jointly formed, as far as the correlation coefficient of the erros terns is significantly different from zero. Optimistic beliefs on technology would produce higher benenfit perceptions and lower isks perception, howewer the effect of knowledge of science is shenages. Interestingly, a higher kwnoweldge leads to higher risks and benenfits perceptions. Furthemore, the gender effect is only prevalent on risks perceptions, again increasing risks perception but not on benenfit perceptions. The religion specific effect disper whilst the political affiliation effect remain for benenfit perceptions. Overall, the results suggest that the process of risk and benenfit formatio is not independent and that certain information channels increasing risks perception might display an effect on benenfit perceptions.

## **5.** CONCLUSION

The acceptance of new technologies conveying benefits (and potential risks) to the population is determined by information sources, and primarily the individuals trust in experts as well as their knowledge to update prior information. However, in the light of our results, risks and benefits perceptions are not independent and that taking into account the potential dependence of similar information channels might affect the risk learning determinants. However, there were some information channels that that affect only benefits and risks separately. This was the case of age, religion practice and political affiliation which affected only benefits perceptions whilst being married that affected only risks perceptions while was not significantly associated with risks perceptions.

This study provides some issues for discussion in the light of risks communication of new technology developments. On the one hand, we have raised the point that benefits of new technologies although perceived are largely dependent on individuals knowledge. Second, we have shown that new information signals conveying risks information are likely to enhance lower benefits and the other way around and that when taking into account this feature, the effect of knowledge of science does determine the increase of both risk and benefits perceptions. However, if individuals are ambiguity or ignorance averse, they prefer known risks and thus they might be skeptical about the acceptance of new technologies until sufficient information is disseminated and their knowledge achieves the desired levels, which arguably are culturally determined.

#### REFERENCES

- Atienza, J and Lujan, JL (1997). Opiniones y Actitudes, N 14. La imagen social de las nuevas tecnologías en España, CIS, Madrid.
- Ellsberg, D. (1961), Risk, Ambiguity and the Savage Axioms, *Quarterly Journal of Economics* 75, 643-669.
- Franks, J (1999). The status and prospects for genetically modified crops in Europe. *Food Policy*, 24: 565-584.

- Inglehardt, R and Baker, W.E (2000). Modernization, Cultural Change and the Persistence of Traditional Values. *American Sociological Review*, February, 2000.
- Kanheman, D and Tversky, A (1986). Prospect Theory and Evaluation. *Journal of Cognitive Psychology*, 73: 34-48.
- Sheery, H Legault, Ireland, D (1998). Consumer biotechnology: a synopsis of survey and focus group research. *Journal of Consumer Policy*, 21: 359-386.
- Stadman, PM (1992). Hazard vs Outrage. IOMA Broadcaster, 6.
- Wohl, J.B (1998). Consumers' decision-making and risk perceptions regarding foods produced with biotechnology. *Journal of Consumer Policy*, 21: 387-404.

Chapter 36

## HURRICANE RISK PERCEPTIONS AND PREPAREDNESS<sup>\*</sup>

## John C. Whitehead

Department of Economics, Appalachian State University, Boone, North Carolina, USA

## INTRODUCTION

Preparedness is one component of effective hurricane risk mitigation. Emergency managers attempt to persuade residents of coastal zones to watch weather reports closely, develop and rehearse evacuation plans, store necessary supplies, and prepare property when storms threaten. Many coastal zone residents do not follow these recommended guidelines, even when recent experience suggests that the benefits of preparation exceed the costs. Understanding the factors that encourage these behaviors is an important step toward motivating coastal zone residents to become better prepared and minimize the losses from hurricanes. Toward this end we explore the relationship between wind and flood risk perceptions and hurricane preparedness.

Measures of objective and subjective risk factors are among the most important predictors of evacuation and other mitigating behavior during hurricanes (Baker, 1991). Those who do not feel safe staying home tend to evacuate (Dow and Cutter, 1997). Mobile home residents are more likely to evacuate than other housing residents (Baker, 1991; Whitehead et al., 2000). Storm intensity (Whitehead et al., 2000) and actual damages (Riad, Norris, and Ruback, 1999) are also factors in evacuation behavior. The purchase of mitigating goods is another type of mitigating behavior. Simmons and Kruse (2000) find that coastal residents value houses with hurricane shutters greater than inland residents. Ozdemir and Kruse (2000) find that the purchase of tornado shelters is based on risk perceptions.

In this paper we measure wind and flood risk perceptions and five determinants of hurricane preparedness in surveys of North Carolina coastal residents conducted after

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

hurricane Bonnie (1998) and hurricanes Dennis and Floyd (1999). These data allow an analysis of changes in risk perceptions and preparedness and their interrelationship. We explore a two-stage model. In the first stage, objective risk factors, hurricane experience, and demographics are determinants of risk perceptions. In the second stage, risk perceptions, hurricane experience, and demographics are the determinants of hurricane preparedness.

In the rest of the paper we describe the surveys and data. Then we use multivariate regression models to analyze the determinants of risk perception and preparedness. Finally we discuss our results and offer some conclusions.

## **THE SURVEYS**

During August of 1998 hurricane Bonnie approached the North Carolina coast prompting evacuations of all eight coastal counties. Hurricane Bonnie made landfall on the coast of North Carolina, near Cape Fear, 20 miles south of Wilmington (Avila, 1998). At landfall hurricane Bonnie was a low category 3 storm and quickly diminished to a category 1 storm on the Saffir-Simpson Hurricane Scale (Tropical Prediction Center, 1999). During the next two days a weakening hurricane Bonnie made its way up the entire North Carolina coast. Much of coastal North Carolina experienced hurricane Bonnie as a tropical storm.

During January 1999 we conducted a telephone survey of North Carolina residents who were affected by hurricane Bonnie in the summer of 1998. The survey used a random digit dial sample of households in the eight North Carolina ocean counties: Brunswick, Carteret, Currituck, Dare, Hyde, New Hanover, Onslow, and Pender. Of the households contacted, 76% completed the interview. The data has 1063 complete and partially complete cases.

During August of 1999, hurricane Dennis approached the North Carolina coast as a category 1 storm, missed the coast, stalled off the Outer Banks for two days, and then landed on the northeast North Carolina coast and quickly lost hurricane intensity (Beven, 2000). Again, portions of the entire North Carolina coast received evacuation orders. Less than one month later, hurricane Floyd traveled north along the Atlantic coast as a strong category 4 storm, weakened as it reached North Carolina, landed near Wilmington as a category 3 storm, and left the state the next day (Pasch, Kimberlain, and Stewart, 2000).

During January 2000 we attempted to contact the same member of the household that we contacted after hurricane Bonnie. Of the 1063 survey respondents, 66 had moved out of the sample counties, 20 had a non-working number, 99 had disconnected numbers, 16 had changed their number, and 29 numbers had become either a business, government, organization, or institute number. Subtracting these unavailable respondents from the total leaves 826 potential contacts. Of the potential contacts, 542 interviews were completed with another 23 interviews partially completed. These 565 contacts represent a response rate of 68% from the hurricane Bonnie survey respondents and an overall 49% response rate from the hurricane Bonnie sample. We focus on respondents to both surveys (n = 533).

We performed a multivariate regression analysis to determine who was more likely to respond to the follow-up survey. The major factor explaining whether recontacted respondents agreed to participate in the follow-up survey is whether the respondent evacuated for hurricane Bonnie (p = .01). The overall evacuation rate for hurricane Bonnie was 26%. The hurricane Bonnie evacuation rate for the subsample of resurvey respondents is 22% and

31% for the subsample of nonrespondents. Length of time at residence increases the probability of resurvey response (p = .01), a typical result for panel surveys. Non-respondents had lived at their residence an average of 9 years at the time of the Bonnie survey. Respondents had lived at their residence an average of 12 years.

## **DEPENDENT VARIABLES**

**Risk Perceptions** 

Following the conceptual model two sets of dependent variables are considered. In the first stage of the model the dependent variables are hurricane wind and risk perceptions. Hurricane wind risk perception is elicited with the question:

"Now consider the risks from wind damage from a hurricane. In your opinion, was the home where you lived at the time of Hurricane [name] located in a high-risk site, a medium risk site, or a low-risk site?"

A similar question was asked to elicit flood risk perceptions:

"Now consider the risks from coastal flooding from a hurricane. In your opinion, was the home where you lived at the time of Hurricane [name] located in a high-risk site, a medium risk site, or a low-risk site?"

Respondents were allowed to interpret the three-response categories, high, medium, and low, in their own way.

After hurricane Bonnie (1998), 46% and 34% of respondents perceived that they lived in a medium and high wind risk home. Following hurricanes Dennis and Floyd (1999), respondents revised their wind risk perceptions downward with 45% and 30% considering their home medium and high risk. The difference across surveys is statistically significant at the p = .10 level ( $\chi^2$ =5.00[2 df]). After hurricane Bonnie, 26% and 14% of respondents thought they lived in a medium and high flood risk home. Following hurricanes Dennis and Floyd, respondents revised their flood risk perceptions downward with 23% and 12% considering their home low, medium and high risk. However, the differences across surveys are not statistically significant ( $\chi^2$ =2.55[2 df]). Flood risk perceptions are significantly lower than wind risk perceptions in the first ( $\chi^2$ =90.64[4 df]) and second ( $\chi^2$ =80.34[4 df]) surveys.

Considering the interpersonal changes in risk perception, 49% and 41% of respondents changed their wind and flood risk perception from the first to the second survey. Nineteen percent of all respondents revised their wind risk perception up (i.e., from low to medium or high, from medium to high) and 29% revised their wind risk perception down. Seventeen percent of all respondents revised their flood risk perception up and 23% revised their flood risk perception down. When considering the interpersonal differences, changes in risk perceptions are significantly different at the p = .01 level for wind risk ( $\chi^2$ =73.37[4 df]) and flood risk ( $\chi^2$ =85.85[4 df]).

#### Hurricane Preparedness

In the second stage of the conceptual model the dependent variables are measures of hurricane preparedness. Preparedness is measured by whether the respondent follows the five steps of the North Carolina Division of Emergency Management's recommended steps of hurricane storm preparation: whether the household (1) has an evacuation plan, (2) rehearses the evacuation plan, (3) prepares their property in advance of a storm, (4) stores supplies, and (5) watches weather reports closely. The existence of an evacuation plan is elicited with the question:

"Which of the following recommended steps of Hurricane Storm Preparation does your household follow? Does your family have a Hurricane Evacuation Plan?"

Respondents who answered yes were then asked: "Do you rehearse your Hurricane Evacuation Plan?" Other questions are "Do you prepare your property in advance of a storm?" "Do you store up to 2 weeks of supplies in your home?" and "Do you watch or listen to weather reports closely?"

Forty-six percent of the sample had an evacuation plan after hurricane Bonnie (1998) and 56% had an evacuation plan after hurricanes Dennis and Floyd (1999). The increase in evacuation plans is statistically significant at the p = .01 level ( $\chi^2$ =10.94[1 df]). Of those respondents who had an evacuation plan, 39% (1998) and 34% (1999) rehearsed their plan. For those who had an evacuation plan in both years (n = 172), 43% did not rehearse in either year, 20% rehearsed in both years, 19% rehearsed only after hurricane Bonnie (1998), and 17% rehearsed only after hurricanes Dennis and Floyd (1999).

Seventy-one respondents had an evacuation plan after hurricane Bonnie (1998) but not after hurricanes Dennis and Floyd (1999). One hundred twenty-five respondents had an evacuation plan in 1999 but not in 1998. Of those who had an evacuation plan in either year, 26% (1998) and 27% (1999) rehearsed the plan. There is no statistical difference in rehearsals across time ( $\chi^2=0.34[1 \text{ df}]$ ).

Most respondents store supplies and prepare their property in advance of a storm while almost all respondents watch weather reports closely. Sixty-nine percent (1998) and 61% (1999) store two weeks of supplies. This difference is statistically significant at the p = .01 level ( $\chi^2$ =7.97[1 df]). Ninety-two percent (1998) and 89% (1999) of respondents prepare their property. This difference is statistically significant at the p = .10 level ( $\chi^2$ =2.83[1 df]). Ninety-eight percent (1998) and 96% (1999) of respondents watch weather reports closely. This difference is not statistically significant ( $\chi^2$ =2.53[1 df]).

#### **INDEPENDENT VARIABLES**

#### **Objective Risk**

We consider several objective measures of risk as factors potentially affecting risk perceptions and hurricane preparedness. Variables that measure objective risk are *hurricane experience*, *property damage*, *mobile home*, *shutters*, *island residence*, *storm surge zone*, and

*flood plain.* Property damage is equal to one if the respondent suffered property damage and zero otherwise. Property damage is expected to increase both wind and flood risk perceptions. Mobile home is equal to one if the respondent lived in a mobile home and zero otherwise. Mobile home is expected to increase wind risk perceptions.

Shutters is equal to one if the respondent had installed storm shutters on their home and zero otherwise. The shutters variable is expected to increase wind risk perception if the respondents in high-risk areas are those who install the shutters. However, shutters could have the opposite effect if respondents feel that shutters completely mitigate against risk. Storm surge zone is equal to one if the respondent lives in a storm surge zone and zero otherwise. Flood plain is equal to one if the respondent lives in the 100-year flood plain and zero otherwise. Storm surge zone and flood plain are expected to increase flood risk perception.

*Hurricane experience* is a measure of objective risk and is captured by a dummy variable equal to zero for the hurricane Bonnie (1998) data and one for the hurricanes Dennis and Floyd (1999) data. The effect of hurricane experience on perceived risk may be either positive or negative. If experience with hurricanes suggests that hurricanes are mild weather events, hurricane experience might lead respondents to adopt low perceptions of risk. If experience suggests that hurricanes are severe weather events, hurricane experience might increase risk perceptions.

#### Demographics

Demographic variables include *pets*, *homeowner*, *apartment*, *white*, *female*, *age*, *education*, and *income*. Pets is equal to one if the respondent owns at least one pet and zero otherwise. Similarly, homeowner is equal to one if the respondent owns their home, apartment is equal to one if the respondent lives in an apartment, townhouse, or condominium, white is equal to one if the respondent is white, and female is equal to one if the respondent is female. Age is the age of the respondent. Education is the number of years of schooling. Income is the 1998 household income of the respondent (in thousands).

County dummy variables are also included. *Brunswick, Carteret, Onslow,* and *Pender* are equal to one if the respondent lives in those counties and zero otherwise. *Outer Banks* is equal to one if the respondent lives in the "Outer Banks" counties of Currituck, Dare, or Hyde Counties and zero otherwise. New Hanover County is the omitted variable. We have no a priori expectations concerning the county dummy variables.

#### Data Summary

Most of the variables are similar in the two surveys (Table 1). Thirty-four percent of the sample suffered property damage during hurricane Bonnie (1998) and 33% suffered property damage during hurricanes Dennis and Floyd (1999). However, these numbers mask significant differences in the households that suffered property damage. Forty-eight percent of the sample did not suffer damage in either storm season. Eighteen percent suffered property damage in 1999 but not 1998. Nineteen percent suffered property damage in 1998 but not 1999. Only 15% suffered damage in both storms.

	Bonni	<b>Bonnie (1998)</b>		loyd (1999)
	Mean	Std.Dev.	Mean	Std.Dev.
Property Damage	0.34	0.47	0.33	0.47
Mobile Home	0.13	0.34	0.13	0.33
Shutters	0.17	0.38	0.16	0.36
Island	0.14	0.35	0.12	0.32
Storm surge zone	0.19	0.39	0.19	0.39
Pets	0.66	0.47	0.66	0.47
Homeowner	0.81	0.39	0.83	0.38
Apartment	0.08	0.27	0.08	0.26
White	0.85	0.35	0.85	0.35
Female	0.60	0.49	0.60	0.49
Age	47.85	16.45	47.85	16.45
Education	13.71	2.17	13.71	2.17
Income	42.91	30.53	42.91	30.53
Outer Banks	0.11	0.32	0.11	0.32
Brunswick	0.12	0.32	0.12	0.32
Carteret	0.15	0.36	0.16	0.36
Onslow	0.28	0.45	0.28	0.45
Pender	0.09	0.29	0.09	0.29

#### **Table 1. Independent Variables**

From the hurricane Bonnie (1998) survey, 13% of the sample lived in mobile homes, 17% have storm shutters, 14% lived on an island, 19% lived in a storm surge zone, and 22% lived on a flood plain. Each of these characteristics is slightly lower after hurricanes Dennis and Floyd (1999). Almost two-thirds of the sample owns a pet in both surveys. Eighty-one percent and 83% of the sample are homeowners in the first and second surveys. Eight percent of the sample lives in an apartment, townhouse, or condominium in both surveys.

Eighty-five percent of the sample is white and 60% is female. The average age is 48 years, the average education level is 14 years, and the average household income is \$43 thousand.

#### **EMPIRICAL MODEL**

We estimate the factors that affect the wind and flood risk perceptions using the ordered probit model (Greene, 1999)

$$y_i^* = \beta' X_i + e_i \tag{1}$$

where  $y_i^*$  is the unobserved latent risk perception variables,  $i = 1, ..., n, \beta$  is a vector of parameters,  $X_i$  is a vector of independent variables and the error term,  $e_i$ , is distributed normally with zero mean and variance equal to  $\sigma^2$ . The latent dependent variable is measured by the ordinal dependent variables. With m = 3 risk choices the dependent variables are coded y = 0, ..., m-1. The ordered probit model uses the same structure as the basic probit model with a censoring parameter,  $\mu$ ,

$$y = 0 \quad if \ y^* \le 0$$
  

$$y = 1 \quad if \ 0 < y^* \le \mu$$
  

$$y = 2 \quad if \ \mu \le y^*$$
(2)

where y = 0 is low risk, y = 1 is medium risk, and y = 2 is high risk. The ordered probit model estimates the probability of the outcome variable using the normal distribution

$$\Pi(y=0) = \Phi(-\beta' X_i)$$

$$\Pi(y=1) = \Phi(\mu - \beta' X_i) - \Phi(-\beta' X_i)$$

$$\Pi(y=2) = 1 - \Phi(\mu - \beta' X_i)$$
(3)

where  $\Phi$  is the standard normal distribution function.

Since we have multiple observations on both risk variables (i.e., measures of risk perceptions in the first and second surveys) we treat the data as a panel. The random effects ordered probit model is a panel data extension of the ordered probit model where the error term accounts for the correlation across respondents

$$y_{it}^* = \beta' X_{it} + e_{it} \tag{4}$$

where t = 1, 2 time periods (i.e., observations for each respondent). The error term,  $e_{it}$ , is distributed normally and is composed of two parts,  $v_{it} + u_i$ , where  $v_{it}$  is the normally distributed random error with mean zero and variance,  $\sigma_v^2$ ,  $u_i$  is the error common to each individual with mean zero and variance,  $\sigma_u^2$ , and  $\sigma_e^2 = \sigma_v^2 + \sigma_u^2$ . The correlation in error terms,  $\rho = \sigma_u^2/\sigma_e^2$ , is increasing in the contribution of the individual error to the total error and is a measure of the appropriateness of the random effects specification.

The determinants of hurricane preparedness are estimated with the binary probit model (Greene, 1999). The probit model is similar to the random effects probit, equation (1), except the latent dependent variable is measured by the dummy variable

$$y = 0 \quad if \ y^* \le 0$$
  
$$y = 1 \quad if \ y^* > 0$$
 (5)

where y = 1 if the respondent engaged in evacuation preparation and y = 0 otherwise. The probit model estimates the probability of the outcome variable with the normal distribution

$$\Pi(y=1) = \Phi(\beta' X_i) \tag{6}$$

The random effects probit model, which is similar to equation (4) but with only two outcomes, is used for the panel data. Each of the probit models is estimated with the LIMDEP statistical software (Greene, 1998).

## **EMPIRICAL RESULTS**

#### **Risk Perceptions**

The hurricane Bonnie (1998) and hurricanes Dennis and Floyd (1999) wind risk models are statistically significant according to the model  $\chi^2$  value (Table 2). In the hurricane Bonnie model we find that property damage, mobile home and island residence have positive effects on perceived wind risk. Residents who live in a storm surge zone and the 100-year flood plain perceive higher wind risks. These variables likely distinguish between those who live near the water and are exposed to hurricane wind damage and those who live in the inland portions of each county. Those who live in the Outer Banks counties, Onslow and Pender counties perceive higher wind risks than New Hanover County residents.

	Bonnie	e (1998)	Dennis/Fl	Dennis/Floyd (1999)		Panel	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	
Constant	-0.207	-0.47	0.009	0.02	-0.029	-0.07	
Property Damage	0.392	3.67	0.259	2.34	0.344	4.04	
Mobile Home	0.392	2.49	0.084	0.53	0.279	2.01	
Shutters	-0.015	-0.12	-0.022	-0.16	-0.032	-0.28	
Island	0.377	2.29	0.295	1.77	0.387	2.66	
Storm surge zone	0.403	2.83	0.212	1.58	0.348	2.93	
Flood Plain	0.222	1.75	0.234	1.87	0.217	2.12	
Pets	-0.062	-0.54	0.186	1.66	0.089	0.85	
Homeowner	0.147	0.97	-0.208	-1.35	0.011	0.08	
Apartment	0.215	1.05	-0.163	-0.83	0.110	0.65	
White	0.059	0.38	0.038	0.25	0.035	0.24	
Female	-0.079	-0.76	-0.148	-1.45	-0.138	-1.41	
Age	-0.002	-0.67	-0.000	-0.04	-0.001	-0.35	
Education	0.032	1.23	0.025	1.02	0.033	1.33	
Income	0.003	1.56	0.003	1.55	0.003	2.10	
Outer Banks	0.380	1.89	0.176	0.93	0.337	1.74	
Brunswick	0.154	0.88	0.190	1.08	0.239	1.53	
Carteret	0.227	1.37	0.203	1.24	0.268	1.73	
Onslow	0.301	2.13	0.021	0.15	0.187	1.42	
Pender	0.430	2.32	0.398	2.00	0.497	2.54	
Property Damage (1998)			0.077	0.71			
Hurricane Experience					-0.168	-2.29	
μ	1.350	18.70	1.252	18.24	1.552	22.43	
ρ					0.67	8.92	
χ2	65.18		46.25		135.58		
Sample Size	533		533		533		
Time Periods	1		1		2		

#### Table 2. Ordered Probit Models of Wind Risk Perception

We find different determinants of wind risk perception in the hurricane Dennis and Floyd (1999) model. The determinants of risk are the property damages, island, flood plain, pets, and Pender County variables. We also include a lagged property damage variable, the coefficient of which is not statistically significant. Surprisingly, those who live in mobile homes perceive the same wind risks as those who live in a fixed structure. Closer inspection of the relationship between wind risk and mobile home supports this result. A significantly

greater number of mobile home residents perceive high wind risks than others after hurricane Bonnie ( $\chi^2$ =6.06[2 df]). However, there is no statistically significant difference in wind risk perceptions between the two groups after hurricanes Dennis and Floyd ( $\chi^2$ =0.11[2 df]).

The random effects probit model constrains all coefficients to be equal across the years of the survey. The measure of correlation across time periods,  $\rho$ , is significantly different from zero, indicating that the random effects specification is appropriate. In this model all of the determinants of wind risk perception are the same as those in the hurricane Bonnie (1998) model except for the Onslow dummy variable. In addition, those with higher incomes and those in Carteret County perceive higher wind risks. The hurricane experience dummy variable measures the additional experience acquired from hurricanes Dennis and Floyd. The effect of hurricane experience is negative indicating that wind risk perceptions are significantly lower after hurricanes Dennis and Floyd when holding the other independent variables constant.

Each of the flood risk models are statistically significant according to the model  $\chi^2$  value (Table 3). In the hurricane Bonnie (1998) model we find that property damage, storm shutters, storm surge zone, flood plain residence, and Outer Banks variables have positive effects on perceived risk. Older respondents have lower perceived flood risks.

	Bonnie	Bonnie (1998)		Dennis/Floyd (1999)		Panel	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	
Constant	-0.003	-0.01	0.239	0.47	0.251	0.56	
Property Damage	0.264	2.24	0.157	1.25	0.215	2.17	
Mobile Home	-0.083	-0.51	-0.067	-0.36	-0.067	-0.43	
Shutters	0.237	1.70	-0.169	-1.10	0.094	0.72	
Island	0.242	1.46	0.440	2.39	0.360	2.49	
Storm surge zone	0.906	6.31	0.766	5.20	0.914	7.38	
Flood Plain	0.339	2.46	0.718	5.06	0.561	5.10	
Homeowner	-0.075	-0.59	-0.231	-1.74	-0.110	-0.95	
Apartment	-0.035	-0.21	-0.403	-2.19	-0.217	-1.44	
Pets	0.099	0.40	-0.374	-1.30	-0.073	-0.35	
White	-0.066	-0.41	0.163	1.00	0.018	0.12	
Female	0.119	1.00	-0.003	-0.02	0.057	0.52	
Age	-0.007	-2.13	-0.009	-2.38	-0.009	-2.63	
Education	-0.020	-0.72	-0.020	-0.68	-0.024	-0.94	
Income	-0.001	-0.40	-0.002	-1.07	-0.001	-0.83	
Outer Banks	0.319	1.66	0.179	0.81	0.318	1.70	
Brunswick	-0.157	-0.73	0.005	0.02	-0.119	-0.63	
Carteret	0.209	1.10	0.085	0.44	0.151	0.87	
Onslow	0.036	0.22	-0.034	-0.20	-0.013	-0.08	
Pender	0.090	0.42	0.120	0.51	0.059	0.29	
Property Damage (1998)			0.451	3.66			
Hurricane Experience					-0.127	-1.53	
μ	0.963	12.31	0.981	12.06	1.141	14.99	
ρ					0.67	6.45	
$\chi^2$	105.39		135.94		233.92		
Sample Size	533		533		533		
Periods	1		1		2		

#### **Table 3. Ordered Probit Models of Flood Risk Perception**

In the hurricanes Dennis and Floyd (1999) model we find that lagged, but not current, property damage has a positive effect on flood risk. The current property damage coefficient

is also insignificant in a model that does not include the lagged property damage variable. As in the hurricane Bonnie model, the storm surge zone, flood plain and age variables have statistically significant effects on flood risk. In addition, island residents perceive higher flood risks. Homeowners perceive lower flood risks than renters and apartment dwellers perceive lower flood risks than those in single-family homes.

In the panel data model, the correlation in errors across time periods is significantly different from zero, indicating that the random effects specification is appropriate. Property damage, island, storm surge zone, flood plain and Outer Banks residents perceive higher flood risks than others. Older respondents perceive lower risks. When the determinants of flood risks are held constant there is no difference in risk perceptions across time.

#### Hurricane Preparedness

The probit models for four of the hurricane preparedness activities are presented in Tables 4-7. Since almost all respondents watch weather reports closely, determinants of this measure of preparedness are not analyzed. In the individual models we test for whether perceived wind and flood risk are predictors of evacuation preparedness and include the demographic variables as control variables. In the random effects probit models we also test the effects of hurricane experience. Each of the probit models, except one, is statistically significant. In the random effects probit models the  $\rho$  statistics are statistically significant indicating that the panel data specification is appropriate.

In the hurricane Bonnie (1998) evacuation plan model, wind risk perception increases the likelihood that a household will have an evacuation plan (Table 4). Flood risk perception has no effect. Other results are that white and female respondents are more likely to have an evacuation plan. In the hurricanes Dennis and Floyd (1999) model, wind risk has a positive effect on whether the respondent has an evacuation plan. In the panel model, wind risk, white, female, and education each contribute to explaining whether the household has an evacuation plan. With hurricane experience, respondents are more likely to have an evacuation plan.

	Bonnie (1998)		Dennis/Fl	oyd (1999)	Panel	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	-1.348	-3.12	-0.710	-1.67	-1.469	-3.31
Wind Risk	0.162	1.96	0.222	2.78	0.223	3.09
Flood Risk	0.088	1.08	0.056	0.66	0.076	1.03
White	0.269	1.68	0.152	0.96	0.267	1.70
Female	0.220	1.93	0.156	1.36	0.238	2.03
Age	0.003	0.84	-0.005	-1.47	-0.001	-0.39
Education	0.032	1.18	0.044	1.62	0.049	1.77
Income	0.001	0.74	0.0002	0.09	0.001	0.57
Hurricane Experience					0.362	4.10
ρ					0.396	6.39
χ2	17.71		19.40		78.74	
Sample Size	533		533		533	
Periods	1		1		2	

Table 4. Probit Models of Hurricane Preparedness: Evacuation Plan

In the evacuation plan rehearsal models we include only respondents who had an evacuation plan in either year (Table 5). In the hurricane Bonnie model, flood risk perception increases the likelihood that a household will rehearse their evacuation plan and wind risk perception has no effect. Older respondents and those with more education are more likely to rehearse. None of the coefficients in the hurricanes Dennis and Floyd (1999) model are statistically significant. In the panel model, increasing perceived flood risk increases the likelihood that a household will rehearse their evacuation plan and wind risk has no effect. Other results are that older respondents are more likely to rehearse their evacuation plan. Those with greater income are less likely to rehearse their plan.

	<b>Bonnie (1998)</b>		Dennis/Floyd (1999)		Panel	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	-2.401	-3.94	-0.017	-0.03	-1.307	-2.50
Wind Risk	-0.013	-0.12	0.028	0.28	0.004	0.04
Flood Risk	0.190	1.84	0.084	0.82	0.147	1.71
White	-0.035	-0.16	-0.282	-1.38	-0.185	-1.04
Female	-0.094	-0.63	-0.113	-0.77	-0.114	-0.92
Age	0.015	3.17	0.001	0.23	0.008	2.07
Education	0.087	2.32	-0.013	-0.35	0.038	1.18
Income	-0.004	-1.59	-0.005	-1.90	-0.005	-2.19
Hurricane Experience					0.083	0.78
ρ					0.192	2.05
χ2	18.78		8.30		22.20	
Sample Size	368		368		368	
Periods	1		1		2	

#### Table 5. Probit Models of Hurricane Preparedness: Rehearse the Plan

Wind and flood risk perceptions have no effect on whether respondents store two weeks of supplies (Table 6). Older respondents are more likely to store supplies in the hurricane Bonnie (1998) model. Respondents with higher incomes are less likely to store supplies in the hurricanes Dennis and Floyd (1999) model. In the panel model, older respondents are more likely to store supplies. Respondents with higher educations and incomes are less likely to store supplies. With hurricane experience respondents are less likely to store supplies.

	Bonnie (1998)		Dennis/Floyd (1999)		Panel	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	0.825	1.847	0.740	1.734	1.119	2.48
Wind Risk	0.023	0.266	0.066	0.820	0.074	0.97
Flood Risk	-0.007	-0.079	-0.010	-0.122	-0.016	-0.21
White	-0.078	-0.465	0.070	0.436	0.002	0.01
Female	0.132	1.119	-0.155	-1.342	-0.023	-0.19
Age	0.007	1.922	0.004	1.167	0.007	1.95
Education	-0.042	-1.492	-0.036	-1.325	-0.050	-1.80
Income	-0.002	-1.136	-0.004	-2.169	-0.004	-2.18
Hurricane Experience					-0.285	-3.16
ρ					0.393	6.06
$\chi^2$	12.65		12.86		60.66	
Sample Size	533		533		533	
Periods	1		1		2	

In the hurricane Bonnie (1998) prepare property model, white and higher income respondents are more likely to prepare their property (Table 7). In the hurricanes Dennis and Floyd (1999) model, those who perceive higher wind risks and white respondents are more likely to prepare their property. In the panel model respondents who perceive higher wind risks are more likely to prepare their property in advance of a storm. White and higher income respondents are also more likely to prepare property. With hurricane experience respondents are less likely to prepare their property.

	Bonnie	<b>Bonnie (1998)</b>		oyd (1999)	Panel	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Constant	0.773	1.21	0.137	0.246	0.774	1.07
Wind Risk	0.132	1.06	0.255	2.221	0.261	1.78
Flood Risk	-0.025	-0.20	-0.041	-0.342	-0.072	-0.45
White	0.522	2.65	0.790	4.365	0.992	3.98
Female	0.166	0.96	-0.201	-1.228	-0.073	-0.35
Age	0.002	0.40	-0.003	-0.746	-0.002	-0.32
Education	-0.043	-1.00	0.039	1.037	0.013	0.26
Income	0.014	2.95	0.0004	0.159	0.007	1.94
Time (=1 if 1999)					-0.257	-1.65
ρ					0.547	5.86
χ2	23.57		32.18		73.80	
Sample Size	533		533		533	
Periods	1		1		2	

### **CONCLUSIONS**

We find that respondents perceive slightly higher wind and flood risks after hurricane Bonnie (1998) than after hurricanes Dennis and Floyd (1999). The reduction in risk perceptions may be due to the misleading effects of hurricane Floyd, which approached North Carolina as a category 4 storm but landed as a category 2 storm on the ocean-bordering counties producing less wind than expected. The rain-induced flooding from hurricanes Dennis and Floyd occurred primarily in the inland coastal plain counties. This experience may have led coastal North Carolina residents to revise their perceived wind and flood risks downward.

Some of the determinants of risk are similar across both types of risk. Those who face the highest risks from hurricanes, residents of islands, storm surge zones, and flood plains, are more likely to perceive greater wind and flood risks. Those who have incurred property damages perceive greater wind and flood risks. We also find differences in risk perception across the type of risk. Mobile home owners perceive higher wind risks but not flood risks. Females perceive lower wind risks. Those who are older perceive lower flood risks. Those with higher incomes perceive greater wind risk but not flood risk. This result may be due to the higher potential economic damages from wind in high-income neighborhoods. We find some geographic differences in wind and flood risk perception.

We find that risk perceptions differ across storm season. Demographic and geographic factors influence risk perceptions after hurricane Bonnie (1998) but not after hurricanes

Dennis and Floyd (1999). One area of concern is that mobile home owners perceive the same risks from wind as those in fixed structures after experiencing hurricanes Dennis and Floyd. Again, this downward reduction in risk perceptions may be due to the misleading effects of hurricane Floyd. Homeowners and apartment dwellers perceive lower risks in 1999 but not in 1998. Another finding is that the property damage that occurred in 1998 increases flood risk perceptions in 1999, but current property damage does not. This result perhaps reflects the greater flooding in the ocean-bordering counties after hurricane Bonnie relative to flooding after hurricanes Dennis and Floyd.

In terms of evacuation planning respondents are better prepared for evacuation after hurricanes Dennis and Floyd (1999) than before. Somewhat surprisingly, respondents are less well prepared for a hurricane in terms of preparing property and storing supplies after hurricanes Dennis and Floyd than before. Perceived wind and flood risks are factors that are considered when preparing for a hurricane. However, these results are not stable over time. Also, the effect of demographic variables on hurricane preparedness changes over time.

Several of these findings should be of concern to emergency managers. The first is that risk perceptions are not stable over time. Perceived hurricane risks are lower after the experiences of two hurricanes in 1999. Perhaps, hurricane risk perceptions decrease when negative impacts from hurricanes are less than expected. The converse may also be true. Risk perceptions may increase when negative impacts are greater than expected.

Second, hurricane preparedness is not stable over time. Respondents are better prepared for evacuation but less prepared for weathering a storm in their home and protecting their home from a storm after experiencing hurricanes Dennis and Floyd. These contrasting results may be consistent. If respondents are more likely to evacuate after hurricanes Dennis and Floyd they may consider property preparation and supply storage to be less important mitigation activities.

Third, objective measures of risk, those that do not change over time, are taken into account when forming hurricane risk perceptions. Those at the most risk, those in mobile homes and those who suffer property damage, perceive lower risks than are warranted.

These results have policy implications. Hurricane preparedness is important for effective risk mitigation. Since reductions in perceived risk lead to lower levels of hurricane preparedness, better risk communication methods are needed to improve preparedness. Coastal residents may experience storms that weaken as they approach the coast or upon landing where hurricane damage and mortality is less than expected. At these times, residents should be better informed that not all major storms weaken as they approach the coast. As in many risk communication problems, however, the information is readily available to those who choose to seek it. Therefore, incentives that lead to the consumption of the risk information are needed.

### REFERENCES

Avila, Lixion A. (1998) Preliminary Report, Hurricane Bonnie 19-30 August 1998, National Hurricane Center, http://www.nhc.noaa.gov/1998bonnie.html, October.

Baker, Earl J. (1991), and "Hurricane Evacuation Behavior," International Journal of Mass Emergencies and Disasters, Vol.9: 287-310.

- Beven, Jack (2000), Preliminary Report: Hurricane Dennis 24 August 7 September 1999, National Hurricane Center, http://www.nhc.noaa.gov/1999dennis\_text.html, January.
- Dow, Kirstin, and Susan L. Cutter (1997), "Crying Wolf: Repeat Responses to Hurricane Evacuation Orders," *Coastal Management*, Vol.26: 237-251.
- Greene, William H. (1998), *LIMDEP Version 7.0 User's Manual*, Revised Edition, Econometric Software, Inc. Plainview, New York.
- Greene, William H. (1999), *Econometric Analysis*, Third Edition, Prentice Hall, Upper Saddle River, New Jersey.
- Ozdemir, Ozlem and Jamie Brown Kruse, (2000), Relationship between Risk Perception and Willingness-to-Pay for Low Probability, High Consequence Risk: A Survey, Discussion Paper 2000-04, Department of Economics and Geography, Texas Tech University.
- Pasch, Richard J., Todd B. Kimberlain, and Stacy R. Stewart (2000), Preliminary Report: Hurricane Floyd, 7–17 September 1999. National Hurricane Center, http://www.nhc. noaa.gov/1999floyd\_text.html, January.
- Riad, Jasmin K., Fran H. Norris, and R. Barry Ruback, (1999), "Predicting Evacuation in Two Major Disasters: Risk Perception, Social Influence, and Access to Resources," *Journal of Applied Social Psychology*, Vol.29: 918-934.
- Simmons, Kevin M. and Jamie Brown Kruse (2000), "Market Value of Mitigation and Perceived Risk: Empirical Results," *Journal of Economics*, Vol.26: 41-51.
- Tropical Prediction Center (1999), *The Saffir-Simpson Hurricane Scale*. National Hurricane Center, *http://www.nhc.noaa.gov/aboutsshs.html*, July 14.
- Whitehead, John C., Bob Edwards, Marieke Van Willigen, John Maiolo, Ken Wilson, and Kevin Smith (2000), "Heading for Higher Ground: Factors Affecting Real and Hypothetical Hurricane Evacuation Behavior," *Environmental Hazards*, Vol.2: 133-142.

Chapter 37

## **RISK PERCEPTION AND THE VALUE OF SAFE ROOMS AS A PROTECTIVE MEASURE FROM TORNADOES: A SURVEY METHOD**\*

## **Ozlem Ozdemir**

Department of Economics, Yeditepe University, Istanbul, Turkey

## **INTRODUCTION**

Over the past thirty years, researchers have developed a theoretical framework concerning risk and the protective mechanisms chosen by individuals (e.g. insurance, seat belts, storm shelters) against disasters (Kunruether, 1978, 1996; Quiggin, 1992; Shogren and Crocker, 1999; Ehrlich and Becker, 1972; Lewis and Nickerson, 1989; Shogren 1990; Viscusi, 1992; Cook and Graham, 1979; Hirshleifer, 1966, Dong and Wong, 1996; Dixit, 1990; Arrow, 1972, 1996; Slovic, 1977). In most of these models, buying insurance is studied as the protective mechanism and monetary loss is taken as the damage from the hazard.

Since the theoretical models are not fully adequate to describe people's reactions under some uncertain and risky situations, particularly low-probability events (Camerer and Kunreuther, 1989), empirical investigations become crucial. As McClelland, Shulze, and Coursey (1993) stated, "both field and lab studies provide complementary information that will lead to a fuller understanding of insurance purchase and other protective behaviors for low-probability risks."

Numerous survey studies have examined risk perception for low-probability, highconsequence (LPHC) hazards (Smith and Devousges, 1987; Fischoff, Watson, and Hope, 1984; Slovic, 1987; McDaniels, Kamlet and Fischer, 1992; Slovic, Fischoff, and Lichtenstein, 1980; Kunreuther, 1976, 1978; Camerer and Kunreuther, 1989; Kunreuther, 1996; McClelland, Schulze, and Hurd, 1990; Kunreuther, Onculer, and Slovic, 1998). Most found

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

divergence in risk perceptions and mitigation actions taken by individuals. While some individuals focus on the probability of the risky event, others act by looking at the extreme loss. These studies call for further investigation of the relationship between the valuation of risk and perception of risk.

The purpose of the current study is to investigate the relationship between risk perception and willingness-to-pay for increased safety in LPHC risk situations. A tornado is chosen as the LPHC risk and willingness-to-pay for an in-residence storm shelter (safe room) is specified as the mitigation measure. The probability that a damaging tornado will strike a given location is miniscule, even in "high" tornado risk areas. However, the prospect of injury or loss of life due to a tornado is frightening. Brown, Kruger and Bos (2000) place tornadoes "among the most violent and lethal of all natural disasters."

An in-residence storm shelter or "safe room" has been proposed to reduce injuries and loss of life in the event of a tornado. One such safe room design by the Wind Engineering Research Center at Texas Tech University has been promoted by The Federal Emergency Management Agency (FEMA)(2000). A safe room is an integral part of the interior of a house and located in the basement or in an interior room on the first floor of a house without a basement. The room is constructed to endure strong winds, windborne missiles, and falling objects.

Brown, Kruger and Bos (2000) report the results of on-site interviews of persons in communities that were severely damaged during the May 3, 1999, Oklahoma City tornado outbreak. 78% of the 614 persons interviewed were inside their homes when the tornado struck. The safe room provides a haven that is consistent with human behavior because it is accessible from within the home.

The question remains whether perception of risk is sufficient to prompt a several thousand-dollar investment in this protective device if it were commercially available. In that respect, first, a theoretical model is developed that uses the severity of risk (magnitude), the probability of occurrence and the protective ability of safe rooms to explain the maximum amount people are willing-to-pay. Different from previous models, the effect on health status as well as monetary loss outcomes are considered in the utility function. In addition, the independent effect of probability and magnitude of risk on WTP are investigated.

Then, a mail survey was conducted to test the conceptual model that explores the relationship between individual tornado risk perception and willingness to pay (WTP) for safe rooms. Different from other empirical research, I examine the willingness-to-pay as two independent components (whether people are willing to buy or not with an undisclosed price and the maximum amount they are willing to pay) that makes a distinction between the choice and payment decision. For perceived risk measurement, perceived exposure, perceived severity, perceived control over risk, and perceived responsibility are chosen as the components of perceived risk (Kunreuther et.al, 1990; Slovic et.al., 1980; Smith, 1986).

The rest of the paper is organized as follows. In the next section I describe prior studies related to this paper. I follow with a description of the theoretical model, the survey instrument and the data collection procedures and relate them to the conceptual model. Results and regression analysis follow with conclusions and suggestions for future research in the last section.

## **RELATED RESEARCH**

The theoretical models used to explain behavioral responses to risk are subject to much debate. Some scholars support the use of expected utility theory to describe the judgments and choices made in LPHC risk situations (Brookshire, Thayer, Tschirhart, Schulze, 1985; Gould, 1989), while others do not (Tversky, Sattath, and Slovic, 1988; Schoemaker, 1982; Robertson, 1974; Kunreuther, 1978; Slovic, 1987; Harless and Camerer, 1994). Thus, researchers have offered alternative theories to expected utility (extended expected utility theories, generalized expected utility and prospect theory) that incorporate a behavioral tendency to overweight low-probabilities (Machina, 1982; Kahneman and Tversky, 1979).

Answering the need for empirical investigation of choice under LPHC situations, many survey studies have investigated the value of mitigation i.e. WTP for reducing the loss from hazards. Each study either has taken different risks (natural disasters, hazardous waste, accidents etc.) or different methodologies to investigate valuation of risk i.e. WTP (e.g., Randall, 1993; Bishop and Heberlein, 1990; Mansfield, 1998; Hammitt, 1990; Smith, Desvousges, and Freeman III, 1985; Brookshire and Coursey, 1987; Crocker and Shogren, 1991; Viscusi and Evans, 1990).

Willingness-to-pay for mitigation measures is highly dependent on perceptions about the uncertain hazard. For many reasons (e.g., the probabilistic information may be a challenge for humans), there has been a gap between public perceptions of risks and the real risks. For instance, members of the League of Women Voters and students were asked to estimate the number of deaths per year in U.S. from automobile accidents. Women estimate 28,000 deaths while students' estimation is 10,500 deaths. In fact, the real number was 50,000 deaths (Slovic, Fischhoff, and Lichtenstein, 1979).

The reasons for biased judgments of risk have always been an interesting research question in many fields (Fischer, Morgan, Fischhoff, Nair, and Lave, 1991; Slovic, Kraus, Covello, 1990, and many others). Slovic (1987) states that risk perception is affected by the length of delay between cause and effect, the degree of control over the risk, the potential magnitude of its effect, and whether the risk is faced voluntarily. Additionally, biased evaluations of risks may be because of unfamiliarity with an event, overconfidence about judgments, and the divergence of opinions about risk (Slovic, et. al, 1980). Whatever the reasons behind the gap between the statistical risk and perceived risk, we are compelled to accept it and must search for an appropriate measurement of perceived risk.

Previous studies determine the perceived risk by stating different components. In Kunreuther, Easterling, Desvousges, and Slovic (1990), four components of perceived risk are: perceived seriousness, control, dread, and trust of the federal government. Slovic, Fischoff, and Lichtenstein (1980) not only claim" perceived risk is quantifiable", but also they list five determinants of perceived risk: frequency of death, subjective fatality estimates, disaster potential, qualitative characteristics, judged seriousness of death. According to Smith (1986), perceived attributes of hazards can be identified as: the extent of knowledge about a hazard, whether it is a source of dread, and the perceived exposure to the hazard.

Alternatively, a "risk ladder," a numerical base comparing risks in terms of their annual mortality probabilities is used in some studies to detect perceived risk (Hammitt, 1990; Mitchell and Carson, 1989; Smith, Desvousges, and Freeman III, 1985; McClelland, Schulze, and Hurd, 1990). A "risk ladder" is a numerical base comparing risks in terms of their annual

mortality probabilities (Hammitt, 1990; Mitchell and Carson, 1989; Smith, Desvousges, and Freeman III, 1985; McClelland, Schulze, and Hurd, 1990). Respondents estimate the annual number of deaths per million from tornadoes [the basic premise is the same as the "determinants of perceived risk" from Slovic, Fischoff, and Lichtenstein (1980, page 190)]. Providing the information to respondents in the form of a "risk ladder" has not been examined precisely by previous studies. However, Camerer and Kunreuther (1989) argue that the degree of uncertainty associated with these risk estimates should be explained; otherwise, people may distrust the entire analysis. Slovic, Kraus, and Covello (1990, p.389) offer the following comment on the usage of a "risk ladder":

.... simplicity and intuitive appeal of comparisons of unrelated risks may be highly deceptive. Many factors appear to play a role in determining whether such comparisons will be useful. Whether these kinds of comparisons ultimately generate more light than heat will depend on the degree to which both the context of risk communication and the content of the messages are sensitive to those factors.

In order to see whether using a "risk ladder" fits the current study, I included a risk ladder question in the survey and concluded that it is statistically inappropriate for this case. One possible reason behind this may be the high variation in the values stated for perceived number of deaths per million per year for tornadoes compared to other risks. Further, the stated number on the risk ladder may capture the combination of exposure and severity; however, the inability to separate exposure from severity seriously limits the explanatory power of this measure.

Previous studies of low-probability, high-consequence risks find divergence in risk perceptions and mitigation actions taken by individuals. While some individuals focus on the probability of the risky event, others act by looking at the extreme loss. For example, Indians near Santa Barbara, California, still live without any fear next to a liquefied natural gas terminal that is shown to be dangerous to health by experts (Kasper, 1980). Individuals tend to buy insurance only when the probability is above some threshold (Camerer and Kunreuther, 1989; Kunreuther, 1996). On the other hand, people "perceive other technological risks, such as a nuclear plant accident, to be more likely than do experts" according to Camerer and Kunreuther (1989). Slovic, Fischoff, and Lichtenstein (1980) explain the overreaction to low-probability risks by emphasizing the importance of the amount of loss rather than the frequency. Consistent with their result, Slovic, et. al (1980, p. 184) states tornadoes to be among the most overestimated risks in terms of the gap between judged frequency and actual number of deaths.

The inconsistent perception of low-probability risks becomes more remarkable in McClelland, Shulze, Hurd (1990). They find some people to perceive the risk of a hazardous waste site as if no hazard exists, while others rate a low-probability risk equal to more frequent risk exposure. These studies call for further investigation of the relationship between WTP and perceived risk.

#### THEORETICAL AND CONCEPTUAL MODELS

For this study, a theoretical model is developed in the spirit of Ehrlich and Becker (1972) to generate testable hypotheses regarding the value of mitigation measures (willingness to pay for increased safety) in risky situations. The theoretical model can be found in Appendix I. Adoption of a safe room is taken as self-insurance type of mitigation, since having a safe room does not affect the probability, but can influence the consequences of the risk. Monetary compensation would not necessarily restore health status in the same way other physical assets can be replaced, thus, market insurance is not a perfect substitute for self-insurance. The theoretical model yields three testable implications: (1) maximum willingness to pay for the protective measure is increasing in the protective measure is increasing in the magnitude (severity) of the loss and, (3) maximum willingness to pay is increasing in the protective ability of the measure (safe room).

For the empirical investigation, first a conceptual model is built to test in our hypotheses.

 $WTP = \alpha + \sum_{n=1}^{n} [B_{1n} Exposure + B_{2n} Severity + B_{3n} Experience + B_{4n} Control + B_{5n} Re sponsibility + B_{6n} Age + B_{7n} Sex + B_{8n} Education + B_{9n} Children + B_{10n} Riskaversion + B_{11n} Timeinresidence + B_{12n} Protective bility + B_{13n} Pr ecaution + B_{14n} Possess + \varepsilon$ 

In this model the maximum amount that people are willing to pay for a safe room is described as a function of perceived risk of a tornado and some socioeconomic variables. Then a survey described in the following section is constructed.

## SURVEY DESIGN AND MEASUREMENT

In this section, important variables used in the survey are classified by topic and discussed with their associated survey questions.

Do you have a smoke alarm?	YES	NO
a burglar alarm?	YES	NO
a car alarm?	YES	NO
natural disaster insurance?	YES	NO
life insurance?	YES	NO
emergency items/ food	YES	NO

## **DEGREE OF RISK AVERSION**

This question is a modified version of the one used for a "risk index variable" in Singh and Thayer (1992). The index is calculated by summing "yes" responses to yes/no questions for six different "revealed preferences about risk behavior." These are listed precautions *already taken* against potential risks that a person/family may be exposed to in daily lives. According to their "relative risk aversion" measurement, the individuals that possess 5 or 6 items are considered "relatively risk averse" and those who have 0,1, or 2 index values are risk lovers.

In the present risk index version, I first modify the items that an individual possesses for protection. For that reason, I chose natural disaster insurance vs. earthquake home insurance, then add life insurance for health concern, and finally include the possession of emergency items/food as the last category. This measurement of risk aversion differs from the previous study in the sense that it uses an ordinal scale. In addition, each individual is placed on a risk attitude scale in terms of the relative degree of risk aversion. For example, a person who has 5 items is considered to be more risk averse than people who scored 0, 1, 2, 3 or 4.

#### **PERCEIVED EXPOSURE TO TORNADO RISK**

In your view, how likely is it that your house will be hit by a tornado?

	Not Very Lil	Not Very Likely			Very Likely		
This Year	1	2	3	4	5		
Within the next 5 years	1	2	3	4	5		
Within the next 10 years	1	2	3	4	5		

Personal exposure is considered rather than public exposure for two reasons; first, it is a common sense that "the individual is primarily concerned with its own exposure to danger" (Bohnenblust and Pretre, 1990). This is different from McDaniels, Kamlet, and Fischer's study (1992). The authors use seven-point psychometric scale reflecting exposure to measure subjective probability. It is also important to note that a similar exposure question was used in "South Florida Mitigation baseline Survey" (1998) by DCA. International Hurricane Center, Institute for Public Opinion Research. Second, it is more appropriate to detect the effect of perceived probability of occurrence on the willingness-to-pay for in-residence shelter, because purchasing a shelter is considered to be an individual decision. Factor analysis (factor scores for each period) is used to determine the linear combination of three periods. Factor scores are not found to be significantly different from each other (and are not shown here). Accordingly, the direct average of three periods is deemed appropriate for the regression analysis.

#### **PERCEIVED SEVERITY OF RISK**

Please rank the following activities in terms of the threat (risk) they pose on human life in general.

(1=highest risk, 5=lowest risk)

Motor vehicle accident

 Airline crash
 \_\_\_\_\_

 Floods
 \_\_\_\_\_

 Tornadoes
 \_\_\_\_\_

 Home accident
 \_\_\_\_\_

In Slovic, Fischoff, and Lichtenstein (1980), 110 people were asked to rank 30 different risks ("activities, substances, and technologies") from each activity/technology from least to most risky. The ranking is by the death rate in the U.S. population. Risks included in their list were unrelated in terms of the exposed setting/ cause of the risk. The ordinal scale question in the present study focuses on the individual's perceived severity of tornado generally compared to the other selected risks from the "safety" category. The risk category "safety" is one of the risk classifications in Fischer, Morgan, Fischoff, Nair, and Lave (1991). In their study, subjects were asked to rank five risks "of greatest concern" ranging from "1=greatest concern to 5=least concern" and their categories were: health, safety, environment, society, and other. I use two natural hazards, including tornadoes, and three other selected risks from "safety category." The reason for selecting the risks from the same category is to eliminate the fact that "the simplicity and intuitive appeal of comparisons of unrelated risks may be highly deceptive" (Slovic, Kraus, and Covello, 1990). Note that rank is coded as (6 – the number stated for tornado) for the analysis.

# **CONTROLLABILITY OF TORNADO RISK**

To what extent do you feel that you can *do something (anything)* to protect yourself and your family from a possible tornado?

I can't do	much		I can	do a lot
1	2	3	4	5

This question is a reworded version of the question used to measure perceived "personal efficacy" in Fischer, Morgan, Fischoff, Nair, and Lave (1991).

#### **PERCEIVED RESPONSIBILITY**

In your opinion, what degree of responsibility each individual/institution below has to undertake to protect you and your family from the damage of a possible tornado?

	Not Responsible At All			Very	Very Responsible Myself		
Local/state government	1	2	3	4	5		
Federal government	1	2	3	4	5		
Media (newspaper, TV)	1	2	3	4	5		
Homebuilders/owners	1	2	3	4	5		
Others (Please specify)	1	2	3	4	5		

Ozlen Ozdemir

I not only modified the alternatives of responsible entities, but also measured the degree of responsibility each individual/institution has to undertake from Fischer, Morgan, Fischoff, Nair, and Lave (1991) study.

# PERCEIVED PROTECTIVE ABILITY OF SAFE ROOM

Do you think this shelter can protect you and your family from a possible tornado? (Please circle one).

Not at all Very much 1 2 3 4 5

Individuals that have a response below 3 are assumed to be *weak believers* of a shelter's protective ability, above 3 to be *strong believers*, and the ones who circled 3 are considered to be *neutral believers*.

# WILLINGNESS-TO-PAY (WTP) MEASURES

Would you be willing to purchase the safe room (in-residence shelter) described above?

Certainly	y Not		Certa	ainly
1	2	3	4	5

How much, at most, would you be willing to pay for a safe room inside your house?

- Less than \$1000
- \$1000-\$2000
- \$2000-\$3000
- \$3000-\$4000
- \$4000-\$5000
- \$5000-\$6000
- \$6000-\$7000
- \$7000-\$8000
- □ \$8000-\$9000
- □ \$9000-\$10,000
- □ More than \$10,000

Consistent with the previous survey studies, contingent valuation method is used for the value of WTP. Different from the studies that used an open-ended question (e.g., McDaniels, Kamlet, and Fischer, 1992; Brookshire and Coursey, 1987) or a percent stated as a portion of current income or a dichotomous type of question (e.g., Bishop and Heberlein, 1990) to measure WTP, for this study, the payment-scale is found to be more appropriate. One reason is that respondents do not have any idea about ranges of safe room price (safe rooms were not available in the market when this study was done). In addition, an open-ended question can be

appropriate for a well-known good, but the range of the values that individuals state may become so diverse for a new good that the researcher may not be able to interpret the result (Donaldson, Thomas, and Torgerson, 1997).

In addition to the maximum willingness to pay question, before giving monetary information about the safe room, subjects are asked to indicate whether they were likely to purchase this kind of good regardless of its price (with an undisclosed price). While the former is considered to be a choice decision, the latter is a payment decision.

#### **SAMPLING PROCEDURE**

A sample of one thousand homeowners from Lubbock was selected to be the potential respondents to the mail survey. The addresses were randomly selected from the population of homeowners by a research company called "InfoUSA". A tornado hit Lubbock on May 11, 1970, killing 27 people. The city is considered by FEMA to be in a high-risk wind zone where shelter is suggested as an effective protective mechanism for tornadoes.

The response rate is 20% with no compensation provided to respondents and with no follow-up method used. According to Dillman (1978), this rate is acceptable as an initial response In order to assess sample representative ability, I conducted two different tests. First, I compared the demographics of the homeowners in the general sampling frame (homeowners in the city of Lubbock) with those of my respondents. This test resulted in no significant differences. Second, in order to evaluate the degree of non-response bias in my data, I compared the early (n1=125) and late (n2=75) respondents in the sample in terms of the mean values of the variables in the research. This test is also suggested that non-response bias may not be a critical problem in the data since I found no statistically significant differences.

# **DESCRIPTIVE STATISTICS**

There are 138 male and 62 female respondents in the data pool. The median age of a respondent is 38. Almost half of the respondents have a college education (42.33%). The median family income is \$45,000. Thus, the demographic description of the respondents is very similar to the general makeup of Lubbock's population. It is interesting to note that less than 15% of the respondents have any kind of shelter. However, 85% report that they would choose to go to safe place *inside* their houses after notice of an approaching tornado. This is consistent with the evidence reported by Brown, Kruger and Bos (2000). Individuals responding to the survey perceive themselves as the most responsible entity to protect themselves and their family from the damage of a possible tornado (68.5% of respondents). The average maximum willingness to pay in the sample is \$2449. This amount is lower than the estimated average cost of building a safe room (\$4000).

The correlation matrix (not shown) gives the strength of the association between variables. We find a highly positive correlation between willingness to purchase and maximum willingness to pay (p-value of 0.000). Relative severity and perceived exposure are uncorrelated and thus useable for the two-component measure of risk perception. Numerous survey and experimental studies suggest that women are more risk-averse than men in

financial decisions (for a review see Schubert, Brown, Gysler, and Brachinger, 1999). The correlation of the degree of risk aversion and gender variables is -0.066 with p-value of 0.360 indicating no strong gender component of risk attitude here. However, the presence of children has a significant positive relationship to willingness to purchase and maximum willingness to pay (Max WTP).

# **REGRESSION ANALYSIS**

The focal variable of the present study is risk perception; however, in order to conduct a stronger test of the effects of risk perception on willingness to pay, several control variables are included. The analysis is based on estimated linear regressions after checking the OLS assumptions. Two regression equations are estimated with dependent variables, willingness to purchase and maximum willingness to pay. The effect of perceived severity and perceived exposure that are the two main components of perceived risk out of four, are examined both as a multiplicative effect and independent effect of each separately.

The multiplicative effect of perceived exposure and severity is found to have a positive effect on willingness to purchase a safe room however no statistical significance is detected for maximum willingness to pay. McDaniels, Kamlet, and Fischer (1992) note that empirical studies have shown that it is applicable to use additive and multiplicative functions for eliciting multi-attribute utility and value. Slovic (1987) describes the similarity between the multiplicative model and analytical representation of risk perceptions.

The results of the regression analysis are summarized in Table 1. Perceived exposure has a significant positive effect on willingness to purchase the safe room, but not on the maximum willingness to pay (p-values are 0.012 and 0.354 respectively). Perceived severity has a significant positive influence on the maximum amount respondents are willing to pay, but not their willingness to purchase the shelter (p-values are 0.136 and 0.038 respectively).

For the analysis on the control variables, the degree of risk aversion has no significant effect on either of the WTP measurements. The belief that individuals can exert some control over the losses from a tornado has a significant positive effect on their maximum willingness to pay. If people believe that safe room protects them from a tornado, then they respond with higher values for both WTP measurements. Income does not have a significant effect on the WTP measurements. The presence of dependent children is a significant factor with a positive influence on the maximum WTP (p-value 0.018). Contrary to many previous studies but consistent with Fischer, Morgan, Fischoff, Nair, and Lave (1991), male respondents stated higher values for WTP measurements. The maximum amount male respondents are willing to pay for the shelter is \$500 higher on average than the amount stated by females. Finally, Max WTP is positively related to the level of education (p-value 0.010).

#### CONCLUSIONS

The present study examines the relationship between individual risk perceptions and willingness to pay for a safe room for tornado protection. The theoretical model based on expected utility predicts that the probability of loss, the severity of loss, the protective ability

of the safe room have a positive effect on the willingness to pay for a safe room. The empirical results are consistent with the theoretical predictions in some dimensions but not in others. The results of the survey study are based on a small segment of the population and so may not generalize. However, they are remarkable in many aspects.

	WTP	MaxWTP
	(Choice Decision)	(Payment Decision)
Intercept	1.560	-3535.0
Expected time to live in residence	-0.019	-7.6
	(0.058)	(0.668)
Risk Aversion	-0.093	51.0
	(0.231)	(0.711)
Severity	0.1510	373.0
	(0.136)	(0.038)
Experience	-0.026	-328.0
	(0.879)	(0.277)
Controllability	-0.206	98.0
	(0.013)	(0.503)
Precautions	0.188	42.0
	(0.011)	(0.744)
Protectivity	0.411	315.0
	(0.000)	(0.093)
Have any shelter	-1.006	-68.0
	(0.000)	(0.878)
Income	-0.000	0.009
	(0.232)	(0.110)
Have Children	0.272	995.6
	(0.249)	(0.018)
Age	0.003	22.9
-	(0.688)	(0.088)
Sex	-0.178	-506.6
	(0.311)	(0.104)
Education	0.020	475.0
	(0.844)	(0.010)
Exposure	0.217	140.0
-	(0.012)	(0.354)
R-square	27.5%	18.0%
F-statistic	4.62	2.67
p-value	0.000	0.001

#### **Table 1. Regression Results**

First, theory predicts a positive relationship between perceived protective ability of the safe room and willingness to pay. Empirically, individuals who have a stronger belief about the protective ability of the shelter i.e. strong believers stated significantly higher values for both WTP measures. Second, the theory predicts a positive relationship between the probability of the harmful event and willingness to pay. Accordingly, willingness to purchase the shelter (choice decision) is positively related to perception of the hazard's probability (perceived exposure). Third, theory predicts that willingness to pay is increasing in the size of

the loss (severity of the hazard). Empirically, there exists a significant positive relationship between perceived severity and maximum willingness to pay (payment decision). Lastly, the presence of dependent children has important explanatory power on both WTP measures. On average people with dependent children in the home state a maximum willingness to pay nearly \$1,000 higher than other respondents.

One of the most interesting facets of this study is the distinction evidently made between the action (or choice) decision and the valuation (or payment) decision embodied in the different descriptions of WTP. Economic theory does not distinguish the WTP measures along these lines. However, respondents do distinguish and react to the two components of risk differently depending on whether they are making a choice or a payment decision. Further research is necessary in many aspects. First, the determinants of perceived risk needs deeper investigation for a precise and more exploratory identification of risk. Second, the questionnaire can be conducted on different samples with a better follow-up procedure. In addition an analysis of different types of disasters (e.g. hurricanes) would be useful.

# APPENDIX 1 Theoretical Model

Let  $U(H_i, W_i)$  represent preferences in state i, where U(., .) is an increasing concave function of health status,  $H_i$ , and wealth level,  $W_i$ . For the following analysis, there are two states of the world, 0 and 1. In state 0, a damaging tornado occurs. The probability of state 0 is *p*. The maximum loss of health status due to tornado as loss of life is denoted as  $H_0=0$ . Further, assume  $H_1 \ge H_0$ ,  $W_1 \ge W_0$  and U(0,W)=0.

Expected utility of the tornado risk is:

$$pU(H_0, W_0) + (1 - p)U(H_1, W_1)$$
(1)

 $H_{0s}$  is defined as health status in the event that a tornado occurs with access to a safe room. The safe room protects health status, thus,  $H_{0s} > H_0$ . The maximum willingness-to-pay (WTP), *c* solves

$$pU(H_0, W_0) + (1-p)U(H_1, W_1) = pU(H_{0S}, W_0 - c) + (1-p)U(H_1, W_1 - c)$$
(2)

For the simplest case, assume that wealth is unaffected by the tornado. Further, suppose that health status is zero in the event of the tornado (i.e. the health outcome is death). The safe room, however, provides complete protection leaving health status the same as in the notornado state. Mathematically,  $W_1=W_0$ ,  $H_0=0$ , and  $H_{0s}=H_1$ . With these simplifications, equation (2) becomes:

$$(1-p)U(H_1, W_1) = U(H_1, W_1 - c)$$
(3)

$$F(p,c) = (1-p)U(H_1, W_1) - U(H_1, W_1 - c) = 0.$$
(4)

Using the implicit function theorem, we can derive the relationship between maximum willingness-to-pay and the probability of a tornado. So,

$$\frac{\partial c}{\partial p} = -\frac{F_p}{F_c} = \frac{U(H_1, W_1)}{U_W(H_1, W_1 - c)} > 0.$$
(5)

Thus, maximum willingness-to-pay for a safe room is increasing in probability, p. Now, suppose the perceived reduction in health status due to tornado is some level of harm, but not necessarily death, thus,  $H_1 \ge H_{0s} > H_0 \ge 0$ . The maximum willingness to pay for this case,  $c^*$ , solves

$$pU(H_0, W_0) + (1-p)U(H_1, W_1) = pU(H_{0S}, W_0 - c^*) + (1-p)U(H_1, W_1 - c^*)$$
  
$$p[U(H_0, W_0) - U(H_{0S}, W_0 - c^*)] + (1-p)[U(H_1, W_1) - U(H_1, W_1 - c^*)] = 0$$
(6)

Again, by the implicit function theorem,

$$\frac{\partial c^*}{\partial p} = \frac{-[U(H_0, W_0) - U(H_{0S}, W_0 - c^*)] + [U(H_1, W_1) - U(H_1, W_1 - c^*)]}{pU_W(H_{0S}, W_0 - c^*) + (1 - p)U_W(H_1, W_1 - c^*)}$$
(7)

$$=\frac{[U(H_1, W_1) - U(H_1, W_1 - c^*)] + [U(H_{0S}, W_0 - c^*) - U(H_0, W_{01})]}{pU_W(H_{0S}, W_0 - c^*) + (1 - p)U_W(H_1, W_1 - c^*)} > 0$$
(8)

By concavity of U and where U (0,w) = 0, thus, a positive relationship is illustrated between the probability of tornado risk and the maximum willingness to pay. H<sub>1</sub>-H<sub>0</sub> is defined as unprotected health loss. In the event of a tornado, the safe room will improve health status up to, but not exceeding the health status in state 1. Further, let  $\theta$  be the level of protection afforded by the safe room against tornado, where  $0 \le \theta \le 1$ .

$$H_{0s} = (1 - \theta) H_0 + \theta H_1. \tag{9}$$

Trivially, 
$$\partial H_{0s} / \partial H_0 = (1 - \theta) \ge 0$$
 and  $1 > (1 - \theta) \ge 0$ . (10)

A larger  $\theta$  implies greater protection. Recall the simplified framework described by equation (3). This is the case where  $\theta = 1$ . We can examine changes in H<sub>0</sub> for a sense of the level of harm or severity of the tornado. Thus,

$$\frac{\partial c^{*}}{\partial H_{0}} = \frac{-p[U_{H}(H_{0},W_{0}) - U_{H}(H_{0S},W_{0} - c^{*})\partial H_{0S}/\partial H_{0}]}{pU_{W}(H_{0S},W_{0} - c^{*}) + (1 - p)U_{W}(H_{1},W_{1} - c^{*})} \\
= \frac{-p[U_{H}(H_{0},W_{0}) - (1 - \theta)U_{H}(H_{0S1},W_{0} - c^{*})]}{pU_{W}(H_{0S},W_{0} - c^{*}) + (1 - p)U_{W}(H_{1},W_{1} - c^{*})} < 0$$
(11)

This means that as unprotected health status increases (decreases) within the bounds,  $0 \le H_0 \le H_1$ , willingness-to-pay decreases (increases). Decreasing  $H_0$  relative to  $H_1$  implies greater severity of the loss. Greater severity implies a higher maximum willingness-to-pay. Lastly, substituting (1- $\theta$ )  $H_0 + \theta H_1$  for  $H_{0s}$  into equation (6), we have

$$p[U(H_0, W_0) - U[(1 - \theta)H_0 + \theta H_1, W_0 - c^*]] + (1 - p)[U(H_1, W_1) - U(H_1, W_1 - c^*)] = 0.$$
  

$$\frac{\partial c^*}{\partial \theta} = \frac{p[U_H(H_{0S}, W_0 - c^*)](H_1 - H_0)}{pU_W(H_{0S}, W_0 - c^*) + (1 - p)U_W(H_1, W_1 - c^*)} > 0$$
(12)

Therefore as the ability of the safe room to protect individuals from a possible tornado increases, the higher the maximum willingness to pay for a safe room. To summarize, the simple theoretical model yields three testable implications. They are maximum willingness to pay is:

- 1. increasing in the probability of the harmful event,
- 2. increasing in the perceived severity of the loss and
- 3. increasing in the protective ability of the safe room.

#### REFERENCES

- Bishop, R.C., T.A. Heberlein (1990), *Economic valuation of natural resources: Issues, theory, and applications*, edited by R.L. Johnson and G. V. Johnson, Boulder CO, Westview Press.
- Brookshire, D. S., D. L. Coursey (1987), "Measuring the value of a public good: an empirical comparison of elicitation procedures", *American Economic Review*, Vol.77: 554-6.
- Brown, S., E. Kruger, and J. Bos (2000), *Injury Update: A Report to Oklahoma Injury Surveillance Participants*, Injury Prevention Service, Oklahoma Department of Health, Oklahoma City, OK.
- Camerer C. F., H. Kunreuther (1989), "Decision Processes for Low Probability Events: Policy Implications", *Journal of Policy Analysis and Management*, Vol.8 (4): 565-92.
- Churchill, G.A. Jr. (1991), *Marketing Research Methodological Foundations*, Fifth edition, The Dryden Press.
- Crocker, T.D., J. F. Shogren (1991), "Preference Learning and Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-public Goods", *Land Economics*, Vol.72: 80-90.
- Crouch, E.A.C., R. Wilson (1982), Risk/Benefit Analysis, Cambridge, Balinger.
- Dillman, D.A. (1978), Mail and Telephone Surveys The Total Design Method. New York, Wiley.
- Donaldson, C., R. Thomas, and D. Togerson (1997), "Validity of Open-ended and Payment Scale Approaches to Eliciting Willingness to Pay", *Applied Economics*, Vol.29: 79-84.
- East Bay Seismic Retrofit Survey (1998), University of California, Berkley and University of Pennsylvania.

- Ehrlich, I, G. S. Becker (1972), "Market Insurance, Self-insurance and Self-protection." *Journal of Political Economy*, Vol. 80: 623-48.
- FEMA Mitigation Directorate (1999), Taking Shelter from the Storm, Building a Safe Room Inside Your House.
- Fischer, G.W., M.G. Morgan, B. Fischoff, I. Nair, and L.B. Lave (1991), "What Risks Are People Concerned About?", *Risk Analysis*, Vol.11 (2): 303-14.
- Fischoff, B., S. Watson, and C. Hope (1984), "Defining Risk", *Policy Sciences*, Vol.17: 123-139.
- Hammitt, J. K (1990), "Risk Perceptions and Food Choice: An Exploratory Analysis of Organic-versus Conventional-produce Buyers", *Risk Analysis*, Vol.10 (3): 367-74.
- Harless, D., C. Camerer (1994), "The Predictive Utility of Generalized Expected Utility Theories", *Econometrica* Vol.62: 1251-89.
- Hasher, L., R.T. Zacks (1984), Automatic Processing of Fundamental Information, American Psychologist.
- Kunreuther, H. (1976), "Limited Knowledge and Insurance Protection", *Public Policy* Vol.29 (2): 227-61.
- Kunreuther, H. (1978), *Disaster Insurance Protection*, Public Policy Lessons, John Wiley, New York.
- Kunreuther, H. (1996), "Mitigating Disaster Losses Through Insurance", Journal of Risk and Uncertainty, Vol.12 (2/3): 171-87.
- Mansfield, C. (1998), "A Consistent Method for Calibrating Contingent Value Survey Data", *Southern Economic Journal*, Vol.64 (3): 665-81.
- McClelland, G.H., W.D. Schulze, and B. Hurd (1990), "The Effect of Risk Beliefs on Property Values: A Case Study of a Hazardous Waste Site", *Risk Analysis*, Vol.10 (4): 485-97.
- McDaniels, T.L., M.S. Kamlet, and G.W. Fischer (1992), "Risk Perception and the Value of Safety", *Risk Analysis*, Vol.12 (4): 495-503.
- Mitchell, R. and R.T. Carson (1989), Using Surveys to Value Public Goods: The Contingent Valuation Method, Resources for the Future, Washington, DC.
- Randall, A. (1993), "Passive-use values and Contingent Valuation Valid for Damage Assessment", *Choices*, Vol.2: 12-15.
- Schubert, R., M. Brown, M. Gysler, and H.W. Brachinger (1999), "Financial Decisionmaking: Are Women Really More Risk-averse?", *American Economic Review*, Vol.89 (2): 381-86.
- Singh, H., M. Thayer (1992), "Impact of Seat Belt Use on Driving Behavior", *Economic Inquiry*, Vol.30: 649-658.
- Slovic, P., N. Kraus, and V.T. Covello (1990), "What Should We Know About Making Risk Comparisons?", *Risk Analysis*, Vol.10 (3): 389-92.
- Slovic, P. (1987), "Perception of Risk", Science, Vol.236: 280-85.
- Slovic, P., B. Fischoff, and S. Lichtenstein (1980), Facts and Fears: Understanding Perceived Risk. Societal Risk Assessments: How Safe is Safe Enough? Edited by Richard C. Schwing and Walter Albers, Jr. Plenum Press, New York.
- Smith, V.K. (1986), "Benefit Analysis for Natural Hazards", Risk Analysis, Vol.6: 325-34.
- Smith, V.K., and W. Devousges (1987), "An Empirical Analysis of the Economic Value of Risk Changes, *Journal of Political Economy*, Vol.95: 1.

- Smith, V.K., W.H. Desvousges, and A.M. Freeman III (1985), Valuing Changes in Hazardous Waste Risks: A Contingent Valuation Analysis, Research Triangle Institute, Research Triangle Park, North Carolina.
- South Florida Mitigation Baseline Survey (1998), DCA, International Hurricane Center, Institute for Public Opinion Research.
- Underwood, B.J. (1971), "Recognition Memory" In H.H. Kendler and J.T. Spence, 313-335, Appleton-Century-Crofts, New York.
- Viscusi, W.K., W. N. Evans (1990), "Utility Functions that Depend on Health Status: Estimates and Economic Implications", *American Economic Review*, Vol.80 (3): 353-74.

#### **ACKNOWLEDGMENTS**

This work was performed under the Department of Commerce NIST/TTU Cooperative Agreement Award 70NANB8H0059.

Chapter 38

# DRIVING DANGEROUSLY: MOTIVES UNDERLYING YOUNG MEN'S DECISION TO ENGAGE IN RISKY DRIVING<sup>\*</sup>

# **Robert Ho<sup>1,†</sup> and Rita Yong Gee** <sup>1</sup>School of Psychology and Sociology,Central Queensland University Rockhampton, QLD 4702,Australia

# ABSTRACT

Two studies were conducted to: (1) identify the motives underlying dangerous driving among young males, and (2) evaluate the hypothesized structural relations (both direct and indirect) between the personality construct of sensation seeking, perception of danger, and the identified motives in representing the way risky driving decisions are made. In study 1, exploratory factor analysis (N = 200) yielded a three-factor structure representing three major motives for risky driving – *driving fast/risk taking, confidence in one's driving skills, disrespect for traffic laws*. Confirmatory factor analysis (N = 264) confirmed and further clarified this factor structure in representing the motives underlying young males' driving behavior. In study 2, path analysis (N = 384) provided overall support for the 'fit' of the hypothesized model for dangerous driving. The implications of the findings with regard to the development of effective intervention strategies for dangerous driving among young males are discussed.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: <sup>1</sup>School of Psychology and Sociology, Central Queensland University, Rockhampton, QLD 4702, Australia, Telephone No:+61+7+49 30 9105, Fax No:+61+7+49+30 6460, Email:r.ho@cqu.edu.au

# INTRODUCTION

The recently published *World Report on Road Traffic Injury Prevention* (Peden, 2004) identified road traffic injuries as a major cause of morbidity and mortality worldwide, with an estimated 1.2 million people killed in road traffic crashes each year and as many as 50 million injured or disabled. Projections indicate that these figures will increase by about 65% over the next 20 years unless there is new commitment to prevention. More specifically, without appropriate action, by 2020, road traffic injuries are predicted to be the third leading contributor to the global burden of disease and injury (Murray & Lopez, 1996). Such statistics are no less appalling in Australia, where deaths due to transportation-related crashes now rank eighth among the ten leading causes of death (Australian Bureau of Statistics, 2003).

While motor vehicle crashes result from a variety of factors, studies examining demographic factors relating to traffic fatalities show that age and gender are major predictors of involvement in accidents. Road trauma is most likely to be the leading cause of death for young people in industrialized countries, with young drivers being injured or killed more often than older drivers (Arnett, 1990; Evans, 1991; Lourens, 1992; Levy, 1990). In terms of gender differences, research findings have consistently shown that males are significantly over-represented in road crashes compared to females. Death rates for road injury in Australia are around three times higher for males than females and this differential has been consistent for the last decade or more (Australian Transport Safety Bureau, 2004; Kreisfeld, Newson, and Harrison, 2004).

While age and gender have been identified as important predictors of involvement in road accidents, their interactive effect clearly points to young male drivers as a high risk group in regard to accident involvement. While males made up 72% of motor vehicle traffic deaths in Australia in 2002, males in the 15-29 year age range accounted for 29% of such fatalities in the same period. Statistics show that within the young driver group, males have a higher risk of being involved in a road crash leading to injury or death due to risky driving (Groeger & Brown, 1989), aggressive driving (Simon & Corbett, 1996), and excessive speed (Federal Office of Road Safety, 1997a). These causal factors suggest that the high crash rates among young male drivers may not be due simply to their relative inexperience as drivers or possible exposure to particularly hazardous driving conditions, but rather that there is something problematic about the judgments or decisions that they make when driving (Harré, 2000). In particular, the plethora of research evidence suggests that it is the propensity to take risks by young male drivers that explains their tendency to be involved in dangerous driving practices (DiBlasio, 1986; Evans, 1991; Hurrelmann, 1990).

A considerable body of literature has examined the underlying factors that may contribute to risky driving by young males. These factors include personality, attitude, risk acceptance, overconfidence, and mood.

#### Personality

Of particular importance is the link between the personality trait of *sensation seeking* and risk taking behavior which is often cited as a characteristic of young male drivers that explains their higher crash risk. Sensation seeking is defined as the need for varied, novel and

complex sensations and experiences and the willingness to take physical, social, legal, and financial risks for the sake of such experiences (Zuckerman, 1994). In his review and synthesis of the literature relating to sensation seeking and risky driving, Jonah (1997) reported that (1) sensation seeking is higher in males than females and is highest in the 16 to 19 year age group, and (2) risky driving behaviors such as speeding, impulsivity, driving under the influence of alcohol, and failure to wear set belts all correlated highly with thrill seeking and boredom susceptibility. Together, these findings suggest that young male drivers who rate highly on sensation seeking tend to drive dangerously.

#### Attitude

Driving safely or dangerously is a choice a driver makes. Young drivers have been shown in a number of studies to have different attitudes toward safe driving compared to older drivers. For example, while older drivers have been found to drive with the aim of getting from point A to point B, young male drivers tend to drive for recreational purposes as it provides them with a sense of freedom and control over their lives (Zimbardo, Keough, & Boyd, 1997). Their attitudes toward law enforcement and in particular, traffic laws also differ from other groups. Young male drivers expect less negative outcomes as a result of committing traffic offences, perceive such offences as socially acceptable, and experience less self-control when conducting such behavior (Parker, Manstead, Stradling, Reason, & Baxter, 1992). This display of non-compliance with law enforcement, lack of self-control, and perceived social acceptability, appears also to be related to their attitudes toward speed and a desire for danger. Speeding has been found to be three times more common among male drivers who had a high number of traffic offences compared to other drivers with no violations. This desire for danger appears to reflect their lower perception of risk, as well as a higher evaluation of their driving ability compared to older drivers (Donovan, Marlatt, & Salzberg, 1983).

#### **Risk Acceptance**

Risk acceptance is the individual's level of perceived risk, or risk threshold in which they are willing to accept and consequently act upon. This level of risk acceptance provides a number of explanations as to why an individual engages in dangerous driving, such as continuing to speed when a pedestrian walks onto the road. First, such behavior may reflect their poor risk perception and/or their overconfidence, resulting in the driver misjudging the distance to the pedestrian and the time to brake effectively. Second, the behavior may reflect poor driving skills, which may not allow the driver enough time to respond and slow down. Finally, a high level of risk acceptance, combined with overconfidence in their driving ability may motivate the young driver to accept risks in order to minimize time delays (Deery, 1999).

#### **Overconfidence/Optimism Bias**

A number of studies have demonstrated higher levels of confidence in young drivers compared to older drivers (Dejoy, 1992; Guerin, 1994; Guppy, 1993). However, this bias appears to be particularly strong in young male drivers who have consistently been found to perceive themselves as at less risk of having an accident than their peers of the same age and sex (Bragg & Finn, 1982; Matthews & Moran, 1986). McKenna, Stanier, and Lewis (1991) also reported that individuals tend to exhibit either a positive self-judgment and/or negative other-judgment. They found that when drivers rated their driving skill on a 1 to 10 scale, ranging from very poor to very good, they tended to rate the skill of other drivers as average, while rating themselves as above average. These findings suggest that while drivers do not have a negative view of others, they do see themselves as better than the average driver.

#### Mood

There is evidence to suggest that mood factors such as a high intensity of anger and aggression, emotional instability and impulsiveness, and emotionally involved driving are more characteristic of drivers (both adolescent and adult) who have a history of crashes than those who have not. In particular, the evidence suggests that younger drivers who have less control of their risk taking impulses were more likely to drive for emotional release (Arnett, Offer, & Fine, 1997; Dukes, Clayton, Jenkins, Miller, & Rodgers, 2001). Grey, Triggs, and Haworth (1989) found that young male drivers scored significantly higher on measures of impulsivity and were more likely to use their motor vehicles to express impulses. Lajunen and Parker (2001) found that younger drivers tend to engage in aggressive risk- taking driving more than do older drivers. They proposed that this might be due to the fact that younger drivers are more prone to being annoyed and react accordingly through more violent acts than older drivers do.

It is clear from the above review that considerable work has already been done in identifying the background factors that predispose some young male drivers to engage risky driving choices. While there is clearly much value in this approach, the mere identification of such factors however, does not provide a framework for understanding the judgments made by this group of drivers in different driving situations and the psychological processes that underlie these judgments. That is, the mere identification of these factors and the measurement of their relationship with the incidence of accidents say very little about how young drivers think and why they think the way they do. In order to gain a better understanding of the decision-making process that underlies their driving practices, theoretical models are needed to move the level of discourse from identification towards an explanation of the decision making process. Such models can be employed to integrate the diverse empirical findings that already exist as well as to guide research on both the motives and the subjective processes involved in risky driving choices.

The present study was designed to achieve two aims. First, the study was designed to examine the motives underlying dangerous driving among young males. Although a number of factors have been identified that may predispose some young male drivers to engage in risky driving, a review of the literature showed that there are presently no assessment tools that tap into the motives for dangerous driving. Although there are a number of measures such

as Zuckerman's (1994) Sensation Seeking Scale, which measures sensation-seeking and risk taking behaviors, and the Driver Anger Scale, which measures a driver's level of anger in a variety of driving situations (Deffenbacher, Oetting, & Lynch, 1994), these scales are essentially unidimensional and each identifies only one major facet of driving practices. As there is no one parsimonious motive for young males to drive dangerously, it would be advantageous to have available a multidimensional measurement tool that can reliably tap into different motives simultaneously. The second aim of the study was to investigate the association between the personality construct of sensation seeking, perception of danger, and the identified motives in influencing young male drivers' decision to engage in risky driving. Such investigation is important because the decision-making process underlying risky driving choices (which may involve sensation seeking, danger perception and specific motives to some degree) is likely to serve as a cognitive mechanism that links sensation seeking, the perception of danger and specific motives to actual risky-driving practices. The present study reports on two studies that were designed and conducted to achieve the above two aims.

# STUDY 1: DEVELOPMENT OF THE MOTIVES FOR DANGEROUS DRIVING SCALE

The primary aim of study 1 was to develop a comprehensive, multidimensional instrument that can tap into the motives for risky driving. It is hoped that the development of such a tool can serve firstly, to identify the motives underlying dangerous driving by young males and secondly, to contribute to the development of a theoretical model that can aid understanding of the psychological processes underlying such motives in dangerous driving.

#### Step 1

A focus group consisting of 15 males between the ages of 18 to 28 years and who held a current driver's license were involved in this stage of the study. The participants, who volunteered for the study, were recruited from a university campus in the Brisbane metropolitan area in Queensland, Australia. At the initial stage of this study, the participants took part in a group discussion of the types of dangerous behaviors young males engage in while driving. After approximately 15 minutes of discussion, the participants were asked to write down on a piece of paper as many reasons as they could think of as to why young males engage in these types of dangerous driving behavior. A total of 94 reasons for driving dangerously were generated and recorded.

The 94 responses were initially grouped together by two judges on the basis of similarity of phrasing. For example, "they get excited about driving fast" was grouped with "speed excites them". A third independent judge resolved any disparities. This reduced the number of responses to 72. The responses were then content-analyzed, based on an arbitrary frequency criterion in which responses listed at least four times were grouped. For example, based on the frequency criterion, responses reflecting motives toward speed were grouped together. This further reduced the number of responses to 54.

A subsequent content analysis (on the basis of similarity of meaning) was carried out on these responses. From the analysis, four thematic categories of responses were identified as motives for dangerous driving. These categories were labeled 'risk-taking', 'mood', 'attitude', and 'skill'. Finally, in order to reduce the number of responses to a more manageable unit, a proportionate number of statements were written by the authors to reflect the meaning-content of each of the 4 categories. A total of 40 statements were written (10 statements for each category/motive) and these were included in a questionnaire for final scale construction and item analysis.

#### Step 2: Exploratory Factor Analysis (EFA)

#### Participants and Procedure

A total of 200 male participants from the Brisbane metropolitan area, Australia volunteered to fill the in the study's questionnaire. Their ages ranged from 18 to 24 years, with a mean age of 21 years. The participants held a current driver's license for an average of two years and three months. The majority of the participants (47%) was employed at the time of the study, and had a mean income ranging from \$10,001 to \$20,000 per year.

#### Materials

Participants responded to a questionnaire consisting of two sections. Section 1 consisted of the 40 statements written to reflect the 4 categories/motives for dangerous driving identified in step 1 of the study. Each statement was to be rated on a 6-point Likert scale with high scores indicating strong endorsement of the driving motives: 1=strongly disagree, 2=moderately disagree, 3=barely disagree, 4=barely agree, 5=moderately agree, 6=strongly agree. Section 2 was designed to elicit demographic information relating to the participant's age, level of education, personal income, employment status, and how long they have held a driver's license.

#### Results

Participants' responses to the 40-item questionnaire were subjected to a principal components analysis, followed by oblique rotation. Inspection of the results revealed that seven factors had eigen-values greater than 1.00. However, examination of the items that loaded on these seven factors indicated that only three factors were interpretable, as well as containing the fewest number of cross-correlated items. In conjunction with results obtained from the scree-plot, these findings suggested a three factor solution. These three factors accounted for 41.12, 7.53, and 5.01% of the total variance respectively, for a combined total of 53.67%. Since the factor correlation matrix showed that the factors were correlated (0.29 to 0.35), oblique rotation, limited to three factors was then conducted.

From the obtained pattern matrix, a total of 29 items were retained, using the criteria of selecting items with factor structure coefficients greater than or equal to 0.33 and no significant cross-correlations. The use of the 0.33 value as a criterion for selecting items is based on the logic that squaring the correlation coefficient ( $0.33^2$ ) yields approximately 10% of the variance explained. Of the 29 items, 15 correlated with Factor 1, eight correlated with Factor 2, and six correlated with Factor 3. Examination of the items that correlated with these

three factors indicated that Factor 1 consisted of items that reflected a desire to drive fast and/or to take risks while driving (e.g., driving fast calms me down; I often overtake on the solid line on my side of the lane). Factor 2 consisted of items that reflected confidence in one's driving skills (e.g., I am a skillful driver and am always in control of my driving; my driving skills allow me to negotiate traffic hazards safely). Factor 3 comprised of items that reflected a negative attitude (disrespect) toward traffic laws (e.g., the present traffic laws are too harsh; it is okay to violate traffic laws).

In order to maximize the internal consistency of the derived factor solution, the items representing each of the three factors were item analyzed. Two criteria were used to eliminate items from these factors. First, an item was eliminated if the inclusion of that item resulted in a substantial lowering of Cronbach's Alpha (Walsh & Betz, 1985). Second, an item was considered to have an acceptable level of internal consistency if its corrected item-total (IT) correlation was equal to or greater than 0.33 (Hair, Anderson, Tatham, & Black, 1997). Examination of the Cronbach's Alphas for the three factors and their items' IT correlations showed that all items were acceptable based on the above two criteria. As such, all 29 items were retained to represent their respective factors. Table 1 presents the three-factor multidimensional Motives for Dangerous Driving Scale (MDDS), together with the factor loadings and corrected item-total correlations for the 29 items.

#### Table 1. Factor Loadings and Corrected Item-Total (IT) Correlations for the Motives for Dangerous Driving Scale

Factor loadings IT correlations

Factor 1: Driving fast/risk taking		
Factor 1: Driving fast/risk taking		
Driving fast calms me down (m34)*	0.77	0.79
I like to drive close behind slower	0.74	0.80
drivers (m8)		
I often sound my horn or make obscene	0.73	0.62
gestures at other drivers if they cut		
in front of me (m7)		
I often overtake on the solid line on my	0.72	0.63
side of the lane (m25)		
When driving at night, it is okay to drive	0.71	0.72
through red lights or stop signs as long as		
I am careful (m33)		
I have no problems exceeding the speed	0.67	0.82
limit because I know I am a good		
driver (m31)		
Driving fast is one way of showing my	0.66	0.80
friends that I am a skillful driver (m15)		

# Table 1. Continued

# Factor loadings IT correlations

I often pull out into oncoming traffic (m39)	0.64	0.68
I like to "race" other cars at the traffic lights (m32)	0.61	0.71
I try my best not to violate traffic laws (m38*1)	0.56	0.50
Playing loud music in the car makes me drive faster (m9)	0.55	0.70
I often like to change lanes even in heavy traffic (m6)	0.52	0.73
I take out my frustrations by driving fast (m2)	0.49	0.73
I tend to drive fast so I can get to my destination sooner (m13)	0.45	0.79
I often drive through traffic lights when the light is amber (m26)	0.42	0.58
Cronbach coefficient = $0.94$		
Factor 2: Confidence in one's driving skills		
<i>Factor 2: Confidence in one's driving skills</i> I react quickly when faced with unexpected traffic hazards (m40)	0.81	0.66
I react quickly when faced with unexpected traffic hazards (m40) I am fluent in changing lanes in heavy traffic (m37)	0.75	0.79
I react quickly when faced with unexpected traffic hazards (m40) I am fluent in changing lanes in heavy traffic (m37) I am a skillful driver and am always in control of my driving (m4)	0.75 0.60	0.79 0.62
I react quickly when faced with unexpected traffic hazards (m40) I am fluent in changing lanes in heavy traffic (m37) I am a skillful driver and am always in	0.75	0.79
I react quickly when faced with unexpected traffic hazards (m40) I am fluent in changing lanes in heavy traffic (m37) I am a skillful driver and am always in control of my driving (m4) My driving skills allow me to negotiate traffic hazards safely (m21) I am able to judge accurately the speed of an oncoming car (m36)	0.75 0.60	0.79 0.62
I react quickly when faced with unexpected traffic hazards (m40) I am fluent in changing lanes in heavy traffic (m37) I am a skillful driver and am always in control of my driving (m4) My driving skills allow me to negotiate traffic hazards safely (m21) I am able to judge accurately the speed of an oncoming car (m36) I am a more skillful driver than most other drivers on the road (m24)	0.75 0.60 0.60 0.57 0.57	0.79 0.62 0.62 0.64 0.67
I react quickly when faced with unexpected traffic hazards (m40) I am fluent in changing lanes in heavy traffic (m37) I am a skillful driver and am always in control of my driving (m4) My driving skills allow me to negotiate traffic hazards safely (m21) I am able to judge accurately the speed of an oncoming car (m36) I am a more skillful driver than most other	0.75 0.60 0.60 0.57	0.79 0.62 0.62 0.64

Cronbach coefficient = 0.86

#### **Table 1. Continued**

#### Factor loadings IT correlations

Factor 3: Disrespect for traffic laws		
It is okay to drink and drive as long as	0.77	0.54
I know I am in control of my car (m18)		
It is highly unlikely that my driving will		
ever cause an accident (m30)	0.69	0.40
The present traffic laws are too harsh	0.62	0.54
(m28)		
Exceeding the speed limit by 10 km per	0.59	0.71
hour is no big deal (m11)		
It is okay to violate traffic laws (m3)	0.55	0.67
I would rather drive a car that is powerful	0.53	0.67
than one that is comfortable (m10)		

Cronbach coefficient = 0.82

\* Order of items in questionnaire: 1= strongly disagree; 6=strongly agree

\*1 Reverse-scored

#### Step 3: Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) was carried out to evaluate the adequacy of the factor structure identified in the exploratory factor analysis. CFA, unlike exploratory factor analysis, allows the researcher to explicitly posit an *a priori* model (e.g., on the basis of the factors identified through exploratory factor analysis) and to assess the fit of this model to the observed data. Based on the factor structure identified through exploratory factor analysis, a three-factor model representing the three motives for dangerous driving was posited. For this measurement model, the three latent constructs of 'driving fast/risk taking', 'confidence in one's driving skills', and 'disrespect for traffic laws' were represented by 15, 8, and 6 indicator items respectively (generated from EFA in step 2). While it can be argued that a greater number of indicators per latent construct will represent that latent construct to a higher degree than fewer indicators, in practice however, too many indicators make it difficult if not impossible to fit a model to data (Bentler, 1980). Based on Hair et al.'s (1997) suggestion that three is the preferred minimum number of indicators to represent a construct, it was decided to limit the number of indicators to three for each of the model's latent construct. This was achieved by using item parcels to represent the original number of items for each latent construct.

*Item parcels.* This technique involves summing responses to individual items and then using scores on these summed parcels in the latent variable analysis. For example, on the basis of a reliability analysis of the 15 items representing the latent driving motive of 'driving fast/risk taking', the items were divided into three parcels, and the items in each parcel were then summed to form three measured variables to operationalize the latent construct.

Adapting the procedure described by Russell, Kahn, Spoth, and Altmaier (1998), the development of these item parcels involved the following steps:

- 1. A reliability analysis on the 15 items assessing 'driving fast/risk taking' was conducted.
- 2. The items were rank-ordered on the basis of their corrected item-total (I-T) correlation coefficients.
- 3. Items were assigned to parcels in a way that equated the average I-T coefficient of each parcel of items with the factor.

Specifically, items ranked 1, 2, 7, 14 and 15 were assigned to parcel 1; items ranked 3, 4, 8, 12, and 13 were assigned to parcel 2; and items ranked 5, 6, 9, 10, and 11 were assigned to parcel 3. This procedure ensured that the resulting item parcels reflected the underlying latent driving motive of 'driving fast/risk taking' to an equal degree.

Figure 1 presents the three-factor measurement model representing the three motives for driving dangerously (driving fast/risk taking; confidence in one's driving skills; disrespect for traffic laws). Each latent driving motive was represented by three computed indicator variables (item parcels). For this model, all factor loadings were freed, indicators were allowed to correlate with only one factor, and the three factors were allowed to correlate (equivalent to oblique rotation).

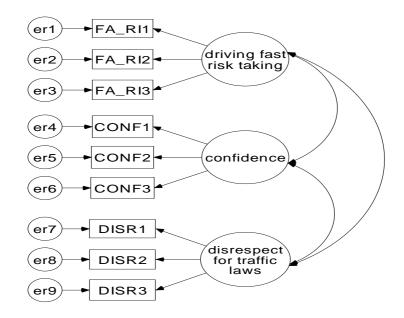


Figure 1. Confirmatory Factor Analysis Model for Dangerous Driving Motives.

#### Participants and Procedure

The sample consisted of 264 male participants recruited from the Rockhampton and Brisbane metropolitan areas, Australia by the researcher and fourth year psychology students from Central Queensland University. None of these participants were part of the exploratory factor analysis stage of the study. Their ages ranged from 18 to 28 years, with a mean age of 21 years. The participants held a current driver's license for an average of four years. The majority of the participants (48%) was employed at the time of the study, and had a mean income ranging from \$10,001 to \$30,000 per year.

#### Materials

Participants responded to a questionnaire consisting of four sections. Section 1 consisted of 5 items written to tap the participant's age, level of education, personal income, employment status, and how long they have held a driver's license.

Section 2 consisted of the 29-item Motives for Dangerous Driving Scale (MDDS), representing the three identified motives of driving fast/risk taking, confidence in one's driving skills, and disrespect for traffic laws. The items were to be rated on 6-point Likert scales with high scores indicating strong endorsement of the driving motives.

Section 3 consisted of Zuckerman's (1994) Sensation Seeking Scale (SSS) (Form V). The 40 forced choice items on this scale require participants to choose between a statement which reflects a desire for sensation ("I like wild and uninhibited parties") and one that reflects a more cautious predilection (I prefer quiet parties with good conversation"). The forced choice items were to be scored with '0' and '1' responses, with high scores indicating high sensation seeking. It is important to note that none of the items refer to driving behavior. The SSS yields four sub-scales: Thrill and Adventure Seeking (TAS), Experience Seeking ES), Disinhibition (DIS) and Boredom Susceptibility (BS). The four subscales, when summed together, provide and overall index of the sensation seeking trait.

Section 4 consisted of the 19-item Danger Assessment Questionnaire (Franken, Gibson, & Rowland, 1992). This measurement tool measures the extent to which a variety of activities are considered to be dangerous. Each item was to be rated on a 6-point Likert scale from 1 (not at all dangerous) to 6 (very dangerous), with high scores indicating strong endorsement of that activity as being dangerous.

#### Results

The purpose of this phase of the study was to evaluate the posited *a priori* model of dangerous driving motives (Figure 1). A  $\chi^2$  goodness-of-fit test (via the statistical program AMOS 5.0; SPSS, Inc. 1997) was employed to test the null hypothesis that the sample covariance matrix was obtained from a population that has the proposed model structure. Table 2 presents the goodness-of-fit indices for this model.

Although the overall chi-square value was significant,  $\chi^2$  (df = 24, N = 264) = 110.92, *p* <.001, the incremental fit indices ((NFI, IFI, TLI, CFI) are all above 0.90 (range: 0.93 – 0.95). These fit indices indicated that the model provided a good fit relative to a null or independence model (i.e., the posited model represented over 90% improvement in fit over the null or independence model), and support the hypothesized structure of the posited three-factor model for dangerous driving. The RMSEA value of 0.07 is also within the range

suggested by Browne and Cudeck (1993) and indicates that the model fits the population covariance matrix reasonably well.

# Table 2. χ<sup>2</sup> Goodness-of-Fit Value, Normed Fit Index (NFI), Incremental Fit Index (IFI), Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), and Root Mean Square Error of Approximation (RMSEA)

Model	χ2 (N=264	) df	` p	NFI	IFI	TLI	CFI R	MSEA
Null Model	1883.87	36	<.001	0.00	0.00	0.00	0.00	0.37
Three-factor	110.92	24	<.001	0.94	0.95	0.93	0.95	0.07
Model								

While the above fit indices can be used to evaluate the adequacy of fit in CFA, it must be noted that this is only one aspect of model evaluation. As pointed out by Marsh and his colleagues (e.g. Marsh, 1996; Marsh & Balla, 1994; Marsh, Hau, & Wen, 2004), model evaluation should be based on a subjective combination of substantive or theoretical issues, inspection of parameter estimates, goodness-of-fit, and interpretability. Table 3 presents the standardized regression weights, residuals, and explained variances for the three-factor model.

	Standardised	Residual	Explained	
Parameter	Regression Weights	Variances	Variances	
Driving fast-risk taking	g → FA_RI1	0.83	0.32	0.68
Driving fast-risk taking	$g \rightarrow FA_{R12}$	0.84	0.29	0.71
Driving fast-risk taking	g → FA_RI3	0.95	0.11	0.89
Confidence $\rightarrow$ CONF	l	0.77	0.41	0.59
Confidence $\rightarrow$ CONF2	2	0.75	0.43	0.57
Confidence $\rightarrow$ CONF3	3	0.80	0.37	0.63
Disrespect for traffic la	aws $\rightarrow$ DISR1	0.68	0.54	0.46
Disrespect for traffic la	aws $\rightarrow$ DISR2	0.65	0.58	0.42
Disrespect for traffic la	aws $\rightarrow$ DISR3	0.81	0.34	0.66

# Table 3. Standardized Regression Weights, Residual Variances, and Explained Variances for the Dangerous Driving Motives Indicator Variables

The standardized regression coefficients (factor loadings) for the measurement indicators were all positive and significant by the critical ratio test, p<.001. Standardized loadings ranged from 0.65 to 0.95 (M = 0.78). These values indicated that the indicator variables hypothesized to represent their respective latent driving motives did so in a reliable manner. The percentage of residual (unexplained) variances for the 9 indicator variables ranged from 11% (i.e. 89% of the variance explained) (FA\_RI3) to 58% (i.e. 42% of the variance explained) (DISR2).

While CFA has confirmed the fit of the three-factor model, an evaluation of the factor correlations yielded by EFA, CFA, and the raw scale scores would be useful in identifying the extent of the overlap between the factors. Table 4 presents these factor correlations.

# Table 4. Factor Correlations Generated from Exploratory Factor Analysis, Confirmatory Factor Analysis, and Raw Scale Scores

Driving fast- Risk taking	Driving fast- Risk taking	Confidence	Disrespect for traffic laws
Confidence	0.21a 0.56b 0.49c		
Disrespect for Traffic laws	0.33a 0.78b 0.63c	0.31a 0.58b 0.47c	
a Exploratory factor and	5		

b Confirmatory factor analysis

c Scale scores

The results indicated that the three factors were moderately correlated (range: 0.21 - 0.78; M = 0.48) and suggest some overlapping between the three motives for dangerous driving. These correlations are not unexpected given that all three motives reflect reasons to engage in risky driving. Indeed, the correlations between these motives suggest that they are jointly implicated, either directly or indirectly, in the decision-making process of young males when they get behind the steering wheels of their cars.

#### Step 4: Test of Convergent Validity

Convergent validity refers to the degree the developed scale is correlated with other concepts in a theoretically based model. That is, theoretically supported relationships from prior research or accepted principles are identified and then the scale is assessed as to whether it has corresponding relationships. High correlations indicate that such relationships, based on theory and/or prior research, do exist. Test of convergent validity for the developed Motives for Dangerous Driving Scale (MDDS) was demonstrated by correlating the summated scales for the three identified motives of driving fast/risk taking, confidence in one's driving skills, and disrespect for traffic laws with the summated scores obtained from the Sensation Seeking Scale (Zuckerman, 1994) and the Danger Assessment Questionnaire (Franken et al., 1992). It is hypothesized that the three identified motives of driving fast/risk taking, confidence in one's driving skills, and disrespect for traffic laws will be (1) positively correlated with sensation seeking, and (2) negatively correlated with perception of danger.

#### Results

The items representing the three motives for dangerous driving were summed across their respective factors and their means computed. Similarly, the 40 forced choice items on the sensation seeking scale and the 19 items on the Danger Assessment Questionnaire were summed across their respective scales and their means computed. Pearson's product-moment correlation analysis was then conducted to investigate the direction and strength of the relationships between the three driving motives and sensation seeking and perception of danger. The results of this analysis are presented in Table 5.

# Table 5. Correlations Between the Driving Motives of Driving Fast/Risk Taking, Confidence in One's Driving Skills, and Disrespect for Traffic Laws with Sensation Seeking and the Perception Of Danger

Driving fast/	Confidence	Disrespect	Sensation Perception
Risk taking		traffic laws see	eking of danger
Driving fast/ Risk taking Confidence .49*** Disrespect .63*** Traffic laws Sensation .60*** Seeking Perception58*** Of danger	.47*** .40*** 33***	.59*** 44***	47***

\*\*\* *p* <.001

The results indicated that all three motives for dangerous driving were highly and significantly correlated with sensation seeking and perception of danger (p < .001). More specifically, the three identified motives of driving fast/risk taking, confidence in one's driving skills, and disrespect for traffic laws were found to be significantly and positively related to sensation seeking and significantly and negatively related to the perception of danger. Thus, the stronger the participants' desire to drive fast/take risks, the stronger their confidence in their driving skills, and the greater their disrespect for traffic laws, the stronger is their need for sensation seeking and the lower their perception of danger. These findings are in line with the study's hypotheses and offer support for the developed Motives for Dangerous Driving Scale's convergent validity.

#### Discussion

The substantive purpose of study 1 was to identify motives underlying the dangerous driving behaviors of young males. Initial exploratory factor analysis of responses derived from qualitative analysis identified a three-factor structure representing three major reasons

for risky driving. Reliability analysis indicated good internal consistency for all three factors. Confirmatory factor analysis confirmed and further clarified the adequacy of this factor structure in representing the motives underlying young males' driving behavior. Finally, correlation analysis offered support for the developed Motives for Dangerous Driving Scale's convergent validity. Together, these findings suggest that the decision by young males to engage in risky driving is a joint function of their desire to drive fast and to take risks, an inflated sense of confidence in their driving ability, and a negative attitude (disrespect) toward traffic laws. While these findings point to the motives underlying the decision to engage in dangerous driving, what is still unclear is what roles these motives play in influencing the decision-making process. That is, do they operate directly and/or indirectly, being mediated by other motives, in influencing the way young male drivers think? And what other factors may interact with these motives in influencing their decision-making process? These questions were investigated in study 2.

# **STUDY 2: MOTIVATIONAL MODEL FOR DANGEROUS DRIVING**

In order to understand how young drivers think and why they think the way they do when driving, this study posits a motivational model that integrates the identified motives for risky driving with the personality characteristics of sensation seeking and risk/danger perception - two characteristics that have been shown to be strongly related to the propensity to take risks by young males. Figure 2 presents the motivational model for dangerous driving, representing the structural relations hypothesized to exist between the personality characteristic of sensation seeking and risk/danger perception, and the endogenous driving motives (driving fast/risk taking, confidence in one's driving skills, and disrespect for traffic laws) identified in study 1. The model is fully specified.

# Relations Between Sensation Seeking, Perception of Danger and Driving Fast/Risk Taking

Several studies have noted that high sensation seekers perceive less risk (i.e. lower perception of danger) in various driving situations and that danger perception and risky driving are negatively correlated (Arnett, 1990; Yu & Williford, 1993). This suggests that danger perception may mediate the relationship between sensation seeking and risky driving. High sensation seekers may not perceive certain driving behaviors as dangerous because they feel that they can speed, follow closely, or drive after drinking and still drive safely as a result of their perceived superior driving skills. Alternatively, high sensation seekers may not only perceive less risk in such activities, they may judge these risks to be desirable. That is, they may initially perceive their behavior as being risky but accept the risk in order to experience the thrill of engaging in it. As pointed out by Jonah (1997), once high sensation seekers have experienced risky driving behavior which has not resulted in negative consequences, they may lower their perceived level of risk, and engage in the behavior more often in the future.

Relations Between Overconfidence, Risk/Danger Perception, and Driving Fast/Risk Taking

When young drivers have reduced risk/danger perception, they underestimate the level of risk in a situation. The considerable evidence that young drivers may tend toward reduced risk/danger perception points to two major reasons. First, there is the tendency for young male drivers, compared to older drivers, to underestimate the risks involved in driving. Past studies that have investigated differences between younger and older drivers in their perception of how risky they consider various driving situations or behaviors to be, have generally found younger drivers to perceive less risk in dangerous situations than older drivers (Bragg & Finn, 1982; Matthews & Moran, 1986; Tränkle, Gelau, & Metker, 1990). Second, young driver crashes are often attributed to overconfidence or excessive optimism about driving skills. The existence of an optimism bias, in which subjects believe they are more skilled and safer in comparison to their peers, has been found in all age groups of drivers (Forsyth, 1992; Guerin, 1994; Guppy, 1993). However, this bias appears to be particularly strong in young male drivers who have consistently been found to perceive themselves as at less risk of having an accident than their peers of the same age and sex (Bragg & Finn, 1982; Finn & Bragg, 1986; Matthews & Moran, 1986).

# Relations Between Sensation Seeking, Attitudes Toward Traffic Laws, and Driving Fast/Risk Taking

Young male drivers may recognize that violations of traffic laws increase their risk of a crash, and judge this undesirable, but are prepared to accept this risk because they consider it necessary to achieve another aim. This acceptance of traffic law violations appears to be particularly true for high sensation seekers. Horvath and Zuckerman's (1993) examination of their Risk Appraisal Scale found that high sensation seekers scored higher than low sensation seekers on the Crime factor and Minor Violations factor (which included involvement in collisions as a result of running red lights or speeding). Moreover, they found that perceived personal risk of engaging in risky driving was negatively correlated with both sensation seeking and the Minor Violations scale, such that high sensation seekers believed their risk was lower than that of low sensation seekers, and that people who perceived less risk were more likely to engage in risky driving behaviors.

Based on the above rationale underlying the posited model for dangerous driving (Figure 2), it is hypothesized that the structural relation between sensation seeking and the decision to drive fast/take risk is both direct and indirect, being mediated by risk/danger perception and the driving motives of confidence and attitudes (disrespect) toward traffic laws. The hypothesized direct relationship reflects the argument that the personality characteristic of sensation seeking is a significant indicator of overall driving intention, and operates directly to influence young male drivers' decision to engage in risky driving. The hypothesized indirect relationships suggest that at least part of the hypothesized direct relationship may be mediated by risk/danger perception and the driving motives of confidence and attitudes toward traffic laws. Thus, young male drivers' level of sensation seeking relates directly to the way they evaluate the risk/danger associated with their risky driving, their level of

confidence in their driving abilities, and their attitude toward traffic laws. These responses, in turn, are related to their decision to engage in speeding and risky driving in general.

# **PARTICIPANTS AND PROCEDURE**

The participants were recruited from the Brisbane and Rockhampton metropolitan areas. The sample consisted of 384 males aged between18 to 24 years with a mean age of 21 years and who held a current driver's license for an average of two years and three months. The majority of the participants (46%) was employed at the time of the study, and had a mean income ranging from \$10,001 to \$20,000 per year. None of these participants took part in the exploratory and confirmatory factor analysis stages of the study.

# MATERIALS

Participants responded to the same questionnaire that was employed for the confirmatory factor analysis stage in study 1. That is, the questionnaire consisted of four sections, with section 1 designed to tap the participants' demographic characteristics. Section 2 consisted of the 29-item Motives for Dangerous Driving Scale (MDDS). Section 3 consisted of Zuckerman's (1994) Sensation Seeking Scale (SSS) (Form V), and section 4 consisted of the 19-item Danger Assessment Questionnaire (Franken, Gibson, & Rowland, 1992).

#### RESULTS

Structural equation modeling was employed to test the path model presented in Figure 2. The model incorporates the three previously identified motives underlying dangerous driving behaviors of young males, together with the two latent constructs of sensation seeking and risk/danger perception. The latter two latent constructs were operationalized by three indicator *item parcels* each (the procedure for deriving these item parcels was described in the earlier CFA section). The fit of the path model posited to represent the direct and indirect structural relationships between sensation seeking and the decision to drive fast/take risk was tested using the statistical program AMOS 5.0 (SPSS, Inc., 1997). This program analyzed the covariance matrix generated from the model's measurement variables.

Table 6 presents the goodness-of-fit indices for this model. Although the overall chisquare value was significant,  $\chi^2$  (df = 101, N = 384) = 388.43, p < .001, the incremental fit indices (NFI, IFI, TLI, CFI) are all above 0.90 (range: 0.90 – 0.93). These fit indices indicated that the model provided a good fit relative to a null or independence model (i.e. the posited model represented over 90% improvement in fit over the null or independence model), and support the hypothesized structure of the posited path model for dangerous driving. The RMSEA value of 0.08, while within the range suggested by Browne and Cudeck (1993), suggests that the posited model represents reasonable errors of approximation in the population.

#### Table 6. χ<sup>2</sup> Goodness-of-Fit Value, Normed Fit Index (NFI), Incremental Fit Index (IFI), Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), and Root Mean Square Error of Approximation (RMSEA)

Model $\chi 2$	(N=384)	df	р	NFI	IFI	TLI	CFI RM	SEA
Null Model	4026.10	136	<.001	0.00	0.00	0.00	0.00	0.27
Path model	388.43	101	<.001	0.90	0.93	0.90	0.93	0.08

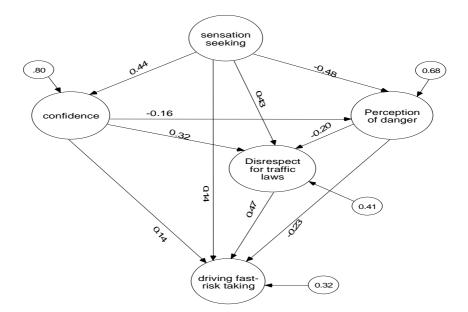


Figure 3. Path Model for Dangerous Driving with Significant Standardized Path Coefficients.

The model's standardized path coefficients are shown in Figure 3. All path coefficients are statistically significant (i.e. p < .05). As shown in Figure 3, sensation seeking was found to be associated directly and indirectly with the motive to drive fast/take risk. Specifically, for the young male respondents, the higher their sensation seeking scores, the stronger was their motive to engage in speeding and risky driving ( $\beta = 0.14$ ). High sensation seeking also increased their level of confidence in their driving abilities ( $\beta = 0.44$ ) and decreased their perception of danger ( $\beta = -0.48$ ), which subsequently increased their motive to drive fast/take risk ( $\beta = 0.14$  and  $\beta = -0.23$  respectively). Increase in their level of confidence also increased their level of disrespect for traffic laws ( $\beta = 0.32$ ) and decreased their perception of danger ( $\beta = -0.16$ ). Disrespecting traffic laws in turn, increased their motive to engage in speeding and risky driving ( $\beta = 0.47$ ).

Figure 3 also reports the standardized residual for each endogenous variable in the model. These coefficients provide an estimate of the proportion of variance in each endogenous variable not predicted by the model. Alternatively, subtracting these values from 1.00 indicates the proportion of variance predicted by the model. These coefficients indicated that the posited model accounted for 20% of the variance in respondents' confidence, 32% of the variance in their perception of danger, 59% of the variance in their attitude toward traffic laws, and 68% of the variance in their motive to drive fast and to take risk.

#### DISCUSSION

The finding of a direct relationship between sensation seeking and the motive to drive fast/take risk fits well with past studies that have linked sensation seeking to a range of risky driving practices including speeding, racing with another vehicle, and illegal passes (Arnett, 1996; Jonah, 1997). While young male drivers are highly aware that such activities are dangerous, the growing body of literature is consistent in showing that they are willing to take the physical and social risks for the sake of such experiences (Arnett, 1992a; McMillen, Smith, & Wells-Parker, 1989; Yu & Williford, 1993). Young males have also been found to score more highly than other groups with regard to "lethality", a related concept that includes an orientation toward danger and violence, bravery and adventure, and thrill seeking and fast driving (Thorsen & Powell, 1987). Together, these findings suggest that the young male sensation seeker feels at risk but judges this to be desirable, with the challenge posed by the risk of driving dangerously overwhelming the fear of the risk and the desire to be cautious.

The findings that sensation seeking is associated with an increase in driving confidence and a decrease in the perception of danger are also in line with past studies that have shown that young driver crashes are often attributed to overconfidence or excessive optimism about driving skills and reduced risk/danger perception. While it has been well established that most drivers consider themselves above average in terms of skill (e.g., Sivac, Soler, & Tränkle, 1989; Svenson, 1981), a number of studies have demonstrated higher levels of confidence in young drivers compared to older drivers, particularly for young males (e.g., Dejoy, 1992). Other studies have argued that the problem for young drivers is underestimation of the risks involved in driving rather than a tendency to over-rate their ability to cope with these situations (e.g., Tränkle, Gelau, & Metker, 1990). The research of Dejoy (1992) demonstrated that the problem seems to be related to both aspects, with young males being characterized by an exaggerated view of their own driving competency and lower perceptions of risk in hazardous on-road situations. When adolescent drivers possess both an inflated sense of their own driving ability and a reduced risk perception, they underestimate the level of risk in a situation, and ultimately strengthening their motivation to drive dangerously and to take risks.

For the young male drivers in the present study, their level of sensation seeking was found to be associated with an increase in negative attitudes (disrespect) toward traffic laws. This finding is supported by social norms relating to both gender and age. For example, women (who tend to score lower on sensation seeking) evaluate traffic violations more seriously than do men (Agostinelli & Miller, 1994; Moyano, 1997), whereas men are more angered than are women by the presence of police (Deffenbacher, Oetting, & Lynch, 1994). Furthermore, research by Rienzi, McMillin, Dickson, and Crauthers (1996) found that adolescents considered drink-driving to be more acceptable for boys than for girls. With regard to age, findings from past studies showed that, young drivers, compared with older drivers, give a lower evaluation of the risk involved in the commission of traffic violations (Dejoy, 1992; Finn & Bragg, 1986; Tränkle et al., 1990). These findings are consistent with those obtained from the present study and suggest that men and younger drivers not only have a higher disrespect for traffic laws, but also expect less negative outcomes as a result of committing traffic violations, perceive more social approval of such behaviors, and experience less control over such behaviors, compared to women and older drivers (Parker et al., 1992).

# **GENERAL DISCUSSION**

The development of the multidimensional Motives for Dangerous Driving Scale (MDDS: in study 1) represents an important contribution to the identification, measurement and ultimately, the understanding of the motives underlying the dynamics of risk-taking behaviors among young male drivers. Such an understanding provides the basis for developing tools and strategies that can be employed to predict at-risk drivers as well as to evaluate and guide responses to them. For example, as young males over-represent crash statistics, it is imperative that driver-training and traffic-safety programs are effective at tapping into what motivates them to engage in high-risk driving practices. Through the development of both reliable and valid assessment tools, researchers and program planners may be able to focus on specific motives for dangerous driving practices. Given the ability of the MDDS to discriminate between motives for dangerous driving, the scale may be used as a screening tool for identifying possible at-risk individuals. By identifying sub-groups of high-risk drivers, interventions or training programs may be tailored specifically to that group.

The MDDS may also be useful in the evaluation of driver-safety programs, particularly where young male drivers may be required to enter court-ordered driver-safety programs as a result of traffic violations. The effectiveness of such programs can be evaluated by applying the MDDS prior to and at the completion of these programs and examining any changes in the `sub-scales' scores. Similarly, the MDDS may also be utilized in the evaluation of traffic safety campaigns such as those that focus on peer intervention programs. These programs are aimed at motivating the young driver's peers to intervene when high-risk behaviors are likely in a given situation, such as drink-driving after a party.

The overall findings from study 1 and study 2 fit well with the growing body of literature that characterizes those at greatest driver risk as: high risk takers, sensation seeking, overconfident in their driving ability, low in danger perception, disrespectful of traffic laws, and male. In particular, past research has shown young male drivers' overconfidence in their driving ability to be primarily responsible for the way they assess risk and danger when driving. The relationship is clearly demonstrated in study 2 where path analysis showed that the young male drivers' confidence in their driving abilities lowered their perception of danger while increasing their negative attitude (disrespect) toward traffic laws as well as their willingness to drive fast and take risks. The problem of overconfidence may, paradoxically lie with the very driver training courses that have been and are still used to train novice drivers the skills to handle and to control their vehicle. While such skills are necessary to be able to even begin to drive, a by-product of such skill-based training is an overestimation of the young drivers' skills. A more effective training strategy may be one that moves the emphasis on training new drivers in basic driving skills to one that helps them to have some insight into their own limits as drivers. The rationale underlying "insight training" is that by making young drivers more aware of the limits to their ability to handle the driving situation, their overconfidence will be reduced (Gregersen, 1996).

A review of the dangerous driving literature will show that there has been more relevant research on risk seeking than on any other factors. Yet, it is probably the most difficult state to shift. This is largely because it is propped up with an entire social system of norms and media images that equate fast driving and 'skillful' maneuvers with masculinity, adulthood, and peer group approval (Harré, 2000). At the individual level, intervention strategies are

unlikely to succeed if they fail to acknowledge the youthful imperative to increase social status by courting danger to demonstrate courage. As pointed out by Nell (2002), driving represents the most common form of sensation seeking in young men because "it bypasses the genetic endowments of strength and speed and makes the demonstration of courage available to all young men, including the slow and the weak" (p. 78). At the social level, risk taking is a highly prized social virtue. One only has to look to certain groups of people – soldiers, police, paramedics, firefighters – to see that risk taking is not only highly valued but is also entrenched in our social establishment where young men learn and exploit the value of risk taking.

In conclusion, the present study provides evidence in support of the two aims posited for this study. The development of the MDDS provides future researchers with an instrument that can act as a quick screening tool to evaluate driving behaviors in young males. Understanding how young male drivers think and why they think the way they do provides directions for the development of effective interventions, as well as the identification of high risk individuals and situations. While these findings contribute to the understanding of the decision-making process underlying risky driving choices, continued investigation of this area is crucial if effective intervention programs are to be developed that can effectively lower the high road injury rate of this group of drivers.

# REFERENCES

- Agostinelli, G., & Miller, W. (1994). Drinking and thinking: How does personaldrinking affect judgments of prevalence and risk? *Journal of Studies on Alcohol*, 55, 327-337.
- Arnett, J. J. (1990). Drunk driving, sensation seeking, and egocentrism amongadolescents. *Personality and Individual Differences*, 11, 541-546.
- Arnett, J. J. (1992a). Reckless behavior in adolescence: A developmental perspective. Developmental Review, 12, 339-373.
- Arnett, J. J. (1996). Sensation seeking, aggressiveness and adolescent reckless behavior. *Personality and Individual Differences*, 20, 693-702.
- Arnett, J. J., Offer, D., & Fine, M. A. (1997). Reckless driving in adolescence: 'state' and 'trait' factors. Accident Analysis and Prevention, 29, 57-63.
- Australian Bureau of Statistics (2003). Deaths. ABS Catalogue No. 3302.0
- Australian Transport Safety Bureau (2004). *Road fatalities Australia: monthly bulletin (June 2004)*. ACT.
- Bentler, P. M. (1980). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bragg, B. W., & Finn, P. (1982). Young driver risk-taking research: Technical report of experimental study (Report No. DOT HS-806-375). Washington, DC: U.S. Dept. of Transportation.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A.
- Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 445-455). Newbury Park, CA: Sage.
- Deery, H. A. (1999). Hazard and risk perception among young novice drivers. *Journal of Safety Research*, 30, 225-236.

- Deffenbacher, J. L., Oetting, E. R., & Lynch, R. S. (1994). Development of a driving anger scale. *Psychological Reports*, 74, 83-91.
- Dejoy, D. M. (1992). An examination of gender differences in traffic accident risk perception. Accident Analysis and Prevention, 24(3). 237-246.
- DiBlasio, F. A. (1986). Drinking adolescents on the roads. *Journal of Youth and Adolescence*, 15(2), 173-188.
- Donovan, D. M., Marlatt, G. A., & Salzberg, P. M. (1983). Drinking behavior, personality factors and high-risk-driving. *Journal of Studies on Alcohol*, 44, 395-423.
- Dukes, R. L., Clayton, S. L., Jenkins, L. T., Miller, T. L., & Rodgers, S. E. (2001).
- Effects of aggressive driving and driver characteristics on road rage. *Social Science*, 38, 323-332.
- Evans, L. (1991). Traffic safety and the driver. New York: Van-Nostrand Reinhold.
- Federal Office of Road Safety (1997a). Young Driver Research Program Mass crash data analyses: Overview of results from Australian and USA mass crash database analyses. CR 131 (11). Retrieved May 12, 2000 on http://www.general.monash.edu.au/muarc/ rptsum/escr131.htm.
- Finn, P., & Bragg, B. W. (1986). Perception of risk of an accident by young and older drivers. *Accident Analysis and Prevention*, 18(4), 289-298.
- Forsyth, E. (1992). Cohort study of learner and novice drivers, Part 2: Attitudes, opinions and the development of driving skills in the first 2 years. TRL Report No: RR372. Crowthorne, England: Dept. of Transport.
- Franken, R. E., Gibson, K. J., & Rowland, G. L. (1992). Sensation seeking and the tendency to view the world as threatening. *Personality and Individual Differences*, 13(1), 31-38.
- Gregersen, N. P. (1996). Young drivers' overestimation of their own skill An experiment on the relation between training strategy and skill. *Accident Analysis and Prevention*, 20(2), 243-250.
- Grey, E. m., Triggs, T. J., & Haworth, N. L. (1989, March). Driver aggression: The role of personality, social characteristics, risk and motivation. *Transport and Communications*. Federal Office of Road Safety.
- Groeger, J. A., & Brown, I. D. (1989). Assessing one's own and other's driving ability: influence of sex, age, and experience. *Accident Analysis and Prevention*, 21, 155-168.
- Guerin, B. (1994). What do people think about the risks of driving? *Journal of Applied Psychology*, 24(11), 994-1021.
- Guppy, A. (1993). Subjective probability of accident and apprehension in relation to selfother bias, age, and reported behavior. Accident Analysis and Prevention, 25(4), 375-382.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1997). *Multivariate data analysis with readings* (4<sup>th</sup> ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Harré, N. (2000). Risk evaluation, driving, and adolescents: A typology. *Developmental Review*, 20, 206-226.
- Horvath, P., & Zuckerman, M. (1993). Sensation seeking, risk appraisal, and risky behavior. *Personality and Individual Differences*, 14, 41-52.
- Hurrelmann, K. (1990). Health promotion for adolescents: Preventive and corrective strategies against problem behaviour. *Journal of Adolescence*, 13, 231-250.
- Jonah, B. A. (1997). Sensation seeking and risky driving: A review and synthesis of the literature. Accident Analysis and Prevention, 29(5), 651-665.

- Kreisfeld, R., Newson, R., & Harrison, J. (2004). *Injury deaths, Australia 2002*. Australian Institute of Health and Welfare: Canberra.
- Lajunen, T., & Parker, D. (2001). Are aggressive people aggressive drivers? A study of the relationship between self-reported general aggressiveness, driver anger and aggressive driving. Accident Analysis and Prevention, 33, 243-255.
- Levy, D. M. (1990). Youth and traffic safety: the effects of driving age, experience and education. *Accident Analysis and Prevention*, 11, 125-127.
- Lourens, P. F. (1992). Young drivers in the Hague. International Journal of Adolescent Medicine and Health, 5 (3-4), 257-267.
- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology*, 70, 810-819.
- Marsh, H. W., & Balla, J. (1994). Goodness of fit in confirmatory factor analysis: The effects of sample size and model parsimony. *Quality and Quantity*, 28, 185-217.
- Marsh, H. W., Hau, K-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11(3), 320-341.
- Matthews, M. L., & Moran, A. R. (1986). Age differences in male drivers perception of accident risk: The role of perceived driving ability. *Accident Analysis and Prevention*, 18(4), 299-313.
- McKenna, F. P., Stanier, R. A., & Lewis, C. (1991). Factor underlying illusory selfassessment of driving skill in males and females. *Accident Analysis and Prevention*, 23, 45-52.
- McMillen, D. L., Smith, S. M., & Wells-Parker, E. (1989). The effects of alcohol, expectancy and sensation seeking on driving risk-taking. *Addictive Behaviors*, 14, 477-483.
- Moyano, D. E. (1997). Evaluation of traffic violation behaviors and the causal attribution of accidents in Chile. *Environment and Behavior*, 29, 264-282.
- Murray, C. J. L., & Lopez, A. D. (1996). The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Boston, MA, Harvard School of Public Health.
- Nell, V. (2002). Why young men drive dangerously: Implications for injury prevention. *Current Directions in Psychological Science*, 11(2), 75-79.
- Parker, D., Manstead, A. S. R., Stradling, S. G., Reason, J. T., & Baxter, J. T. (1992). Determinants of intention to commit driving violations. *Accident Analysis and Prevention*, 24, 117-134.
- Peden, M. (2004). *The world report on road traffic injury prevention*. Geneva, World Health Organization.
- Rienzi, B. M., McMillin, J. D., Dickson, C. J., & Crauthers, D. (1996). Gender differences regarding peer influence and attitude toward substance abuse. *Journal of Drug Education*, 26, 339-347.
- Russell, D. W., Kahn, J. H., Spoth, R., & Altmaier, E. M. (1998). Analyzing data from experimental studies: A latent variable structural equation modelling approach. *Journal* of Counseling Psychology, 45 (1), 18-29.
- Simon, F., & Corbett, C. (1996). Road traffic offending, stress, age, and accident history among male and female drivers. *Eregonomics*, 39, 757-780.

- Sivac, M., Soler, J., & Trankle, U. (1989). Cross-cultural differences in driver selfassessment. Accident Analysis and Prevention, 21, 371-375.
- SPSS Inc. (1997). Statistical Package for the Social Sciences. Chicago, U.S.A.
- Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologia*, 47, 143-148.
- Thorsen, J. A., & Powell, F. C. (1987). Factor structure of a lethal behaviors scale. *Psychological Reports*, 61, 807-810.
- Tränkle, U., Gelau, C., & Metker, T. (1990). Risk perception and age-specific accidents of young drivers. *Accident Analysis and Prevention*, 22(2), 119-125.
- Walsh, W., & Betz, N. (1985). Tests and assessments. New Jersey: Prentice-Hall Inc.
- Yu, J., & Williford, W. R. (1993). Alcohol and risk/sensation seeking: Specifying a causal model on high-risk driving. *Journal of Addictive Diseases*, 12(1), 79-96.
- Zimbardo, P., Keough, K. A., & Boyd, J. N. (1997). Present time perspective as a predictor of risky driving. *Personality and Individual Differences*, 23, 1007-1023.
- Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking*. Cambridge University press, New York.

#### **Driving fast/risk taking**

FA\_RI1

I have no problems exceeding the speed limit because I know I am a good driver

Driving fast is one way of showing my friends that I am a skillful driver

I often like to change lanes even in heavy traffic

I often drive through traffic lights when the light is amber

I try my best not to violate traffic laws

FA\_RI2

I like to drive close behind slower drivers

Driving fast calms me down

When driving at night, it is okay to drive through red lights or stop signs as long as I am careful

I often overtake on the solid line on my side of the lane

I often sound my horn or make obscene gestures at other drivers if they cut in front of me FA\_RI3

I tend to drive fast so I can get to my destination sooner

I take out my frustrations by driving fast

I like to "race" other cars at the traffic lights

Playing loud music in the car makes me drive faster

I often pull out into oncoming traffic

#### Confidence

CONF1

I am fluent in changing lanes in heavy traffic

I am a more skillful driver than most other drivers on the road

I often pay attention to other road users

CONF2

I can easily re-gain control of my car if it skids in wet weather

I am a skillful driver and am always in control of my driving I react quickly when faced with unexpected traffic hazards CONF3 I am able to judge accurately the speed of an oncoming car

My driving skills allow me to negotiate traffic hazards safely

## **Disrespect for traffic laws**

# DISR1

Exceeding the speed limit by 10 km per hour is no big deal It is highly unlikely that my driving will ever cause an accident

# DISR2

It is okay to violate traffic laws The present traffic laws are too harsh

## DISR3

I would rather drive a car that is powerful than one that is comfortable It is okay to drink and drive as long as I know I am in control of my car

Chapter 39

# RISK MANAGEMENT DECISION-MAKING IN A HIGH RISK RECREATIONAL ACTIVITY: LESSONS FROM MOUNTAINEERING<sup>\*</sup>

*Robert C. Lee*<sup> $1,\dagger$ </sup> *and Linda S. Cook*<sup>2</sup>

<sup>1</sup> University of New Mexico, Department of Emergency Medicine, Albuquerque, New Mexico, USA
<sup>2</sup> University of New Mexico, Department of Internal Medicine, Albuquerque, New Mexico, USA

### ABSTRACT

Mountaineers can be viewed as "edgeworkers" who carefully manage risks in a voluntary activity that has the potential for serious injury or death. We maintain that the act of managing risks contributes to a sense of "flow" or transcendence that is a major psychological motivation for mountaineers. We also maintain that decisions associated with engaging in mountaineering as an activity, choices of particular types of mountaineering, locations, specific mountains, etc., and decisions associated with the act of mountaineering itself are conducted, implicitly or explicitly, within a rational multi-objective risk management framework. There are numerous risk-risk, risk-benefit, and benefit-cost tradeoffs throughout this risk management hierarchy. Furthermore, the risk management process is dynamic on a number of different levels. This rational process must be conducted within a context of uncertainty and fear, balanced with the sense of flow, and survival is sometimes at stake. Based on a review of the peer-reviewed, mountaineering association, and popular mountaineering literature, interviews with mountaineers and professional guides, and personal experience, we were unable to find any explicit exploration of dynamic hierarchical multi-objective risk management in

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: Department of Emergency Medicine, MSC 10 5560, 1 University of New Mexico, Albuquerque NM 87131-0001, USA, Telephone: 505-272-1429, Fax: 505-272-6503, Email: rclee@salud.unm.edu

voluntary, recreational risky activities. The psychology of decision-making in this context is also largely unexplored. Thus, we have developed a qualitative framework, based on Hammond, Raiffa, and Keeney's PrOACT framework, that helps elucidate and inform the decision-making processes associated with mountaineering. This approach may help mountaineers and society understand the psychology and tradeoffs associated with this activity, and may be useful for risk management of other types of risky recreational activities.

#### INTRODUCTION

"There are only three sports: bullfighting, motor racing, and mountaineering; all the rest are merely games": attributed to Ernest Hemingway

"... you have to be good when you take nasty risks, or you'll lose it, and then you're in serious trouble": Hunter S. Thompson

"Risky", "extreme", or "adventure" recreation (i.e. voluntary activities involving the potential for serious injury or death) are increasingly popular in developed countries. Such activities may include skydiving, SCUBA diving, surfing, mountain biking, skiing, hang-gliding, and the like. Mountaineering is commonly viewed as one of these risky activities.

Risk is defined here in the formal sense; i.e. as a function of probability and consequence [Cox 2002]. A hazard is a potential risk. Mountaineering is defined here as mountain climbing or travel activities that routinely involve specialized techniques and equipment to manage risks (i.e. to reduce probability and/or consequences of adverse events), as opposed to, say, walking, hiking, or ski touring in non-glaciated terrain. Such equipment may include ropes, harnesses, fall protection (i.e. devices placed in rock, snow, or ice that are clipped to the person or rope and thus theoretically limit falls), helmets, and so forth. Thus, mountaineering may include rock climbing, waterfall ice climbing, glacier travel, "technical" mountain climbing, and ski mountaineering involving glacier travel. Mountaineering may be conducted as a solitary or group activity, and can be performed by amateurs or professionals.

The question of why individuals engage in mountaineering, which can be not only risky, but physically miserable and costly, has been discussed in the social science literature [Loewenstein 1999, Heywood 1994], as well as numerous popular books and articles. From a psychological perspective, there may be a popular perception among some non-mountaineers that mountaineers are risk-loving, and that pursuit of risk *per se* is a main driver for their activities. This may be true some individuals, but there is ample evidence that this is not the case for many mountaineers and similar recreationalists [Creyer et al. 2003, Delle Fave et al. 2003, Olivier 2006]. Further, mountaineering has been discussed as a means of escape from a rationalized world [Heywood 1994, Kiewa 2002], but the phenomenon is likely more deepseated and complex than this. Lyng [1990, 2005], employing a term first used by the psychoactive-drug using popular writer Hunter S. Thompson [1974], has used the term "edgework" to describe seeking of gratification, self-actualization, or transcendence (what Csikszentmihalyi and colleagues [1991] refer to as "flow") via risky activities. However, edgeworkers must carefully manage the risks in order for the activities to be sustainable; hence the Thompson quote at the beginning of this chapter. We prefer the term "manage"

rather than "control" (which is used by many authors), in that "control" may imply that the risk can be reduced to zero. This is never the case.

Lyng and colleagues (1990, 2005) draw many parallels between diverse edgeworkers such as drug addicts, high-stakes businessmen, and risky recreationalists. A consistent attribute is management or control of a risky situation, from which mountaineers and other edgeworkers derive satisfaction [Loewenstein 1999, Celsi et al. 1993]. Professional mountain guides exemplify this factor, in that they derive satisfaction from not only managing their risk, but that of their clients [Beedie 2003]. The explosion of "adventure" recreation and tourism in recent years may be symptomatic of a change in culture that may be spawning more edgeworkers, although in many cases this sort of "risky" adventure occurs in a carefully controlled environment, and is in reality quite safe. Examples include guided whitewater rafting [e.g. Holyfield et al. 2005] and expansion-bolt protected "sport" climbing. In these cases, the participants may essentially delude themselves into thinking that the activity is risky, similar to participants on an amusement park ride. The net psychological effect may be the same, although as discussed below the consequences of failure to manage risk may be vastly different between, say, sport climbing and high-altitude mountaineering.

Given that pursuit of flow along with management of risks are characteristic of edgeworkers in general, and mountaineers in particular, we propose that risk management in mountaineering can and must be conducted in a rational fashion, and thus can be framed within utility theory. This is independent of whether one considers the act of mountaineering per se to be rational. Although it has been argued that utility theory has its limitations as applied to recreation [McNamee et al 2001], and indeed proposed that mountaineering is a means of escape from a rationalized world [Heywood 1994, Kiewa 2002], Lowenstein [1999] compellingly argues that mountaineering decision-making can fit within a broad definition of utility, and indeed uses the example of mountaineering to call for a return from narrow, economics based views of utility to the original broad utility concepts of Bentham. Lowenstein, however, does not explore the complex dimensions of risk management associated with the activity. If we assume that the pursuit of flow or some similar objective is a strong motivator of behaviour and is strongly related to the value structure of the mountaineer, in order to attain that objective in a sustainable fashion we argue that risks should and indeed must be managed in a rational framework. The lucky, overly-romantic mountaineer may be able to blunder irrationally through many climbs and survive (and many have, according to the popular literature. . .), but the chances of survival are obviously increased if systematic risk management occurs.

Thus, interesting conflicts exists in the mind of the mountaineer. She seeks flow via management of risky situations, and deliberately puts herself in those situations. These situations may elicit a range of emotions from exalted joy to mild apprehension to crushing fear. A rational thought process must nonetheless occur throughout to maximize survival. This dynamic psychological state can present extreme challenges. Indeed, the physical challenges of mountaineering often pale in comparison.

We further propose that not only decisions associated with the acts or process of mountaineering, but also those associated with engaging in mountaineering itself as an activity, as well as choices of particular types and locations of mountaineering, can be framed within a multi-objective hierarchy. Multi-objective decision-making is an approach to informing risk management decision-making that explicitly addresses the objectives of the decision maker and the tradeoffs across multiple objectives [Keeney 1992]. There are

numerous risk-risk, risk-benefit, and benefit-cost tradeoffs throughout the mountaineering decision hierarchy (examples are provided in Table 1). Additionally, the decision-making process is dynamic on a number of different levels. Multi-objective decision-making provides an informative structure in which to frame these complex decisions, and thus potentially refine the decision-making process of the mountaineer.

# Table 1. E xamples of Risks, Benefits, and Costs Associated with Mountaineering; and Examples of Associated Tradeoffs

- *Risks:* Human failure, equipment failure, foreign travel, ice conditions causing falls, ice fall, rock conditions causing falls, rock fall, avalanches, weather, getting disoriented/lost, cold temperatures, altitude, unstable political environments in developing countries, insect-, food- and water-borne illness.
- *Benefits:* Physical fitness, mental fitness, exposure to wild natural environments (beauty, connection with "nature", etc.), exposure to different ecosystems and cultures, changes in brain chemistry (e.g. endorphin release), sense of "flow", sense of accomplishment.
- Costs: Equipment, travel (air fare, vehicle, hotels, etc.), house/child sitting.
- *Risk-risk tradeoffs:* Risk of proceeding up a climb with bad rock conditions vs. risk of descending with poor anchors, risk of climbing slowly in adverse conditions using much safety equipment vs. risk of inadequate safety equipment but climbing quickly, risk associated with climbing vs. risk of mental stress associated with not climbing.

Risk-benefit tradeoffs: Risk of injury vs. benefit of physical fitness, risks of illness vs. benefits of travel to beautiful and interesting foreign environments

There is a rich literature on the psychology and nature of voluntary versus involuntary risk; e.g. Slovic [1986], McDaniels et al. [1992]; and a psychological/sociological literature on edgework, as previously mentioned. However, based on a review of the peer-reviewed, mountaineering association, and popular mountaineering literature, we were unable to find any exploration of dynamic hierarchical multi-objective risk management in voluntary, recreational risky activities. We conducted interviews with approximately 20 mountaineers and professional mountain guides to build on our personal experience (approximately 50 years of combined mountaineering experience) in developing a formal, qualitative description of an approach that informs the risk management decision-making processes inherent in mountaineering, and which adds to the edgework and similar literature. We based this approach on the PrOACT framework (Hammond et al. 1999). PrOACT incorporates consideration of 5 major elements; Problem, Objectives, Alternatives, Consequences, and Tradeoffs; and also uncertainty, risk tolerance, and linked decisions. This framework is based on 40 years of utility theory and application of decision science [e.g. Raiffa 1968] in a wide variety of contexts. We then apply this approach to a personal case study. Our intent is to not present a normative, prescriptive decision-making process, but rather an analytical framework to inform risky decisions that can aid complex mountaineering decisions that are often fraught with complex emotions. Our approach may be useful for informing risk management decision-making in other types of risky recreational activities.

#### **Risks Associated with Mountaineering**

The quote at the beginning of the chapter attributed to Mr. Hemingway implies that risk to life and limb is a necessary constituent of "sport". Mountaineering entails some risk, because people are routinely injured or die while they are doing it, but estimating the risk is difficult for three major reasons.

First, there is a wide variability across different types of mountaineering. For example, climbing a large, technically difficult wilderness mountain peak entails a much different set of risks than those associated with, say, fixed expansion-bolt protected sport rock climbing at the local cliff (Table 2).

Technical mountaineering risks may involve "objective" risks such as avalanches, storms, rock-fall, etc., as opposed to "subjective" risks such as misjudging the difficulty or length of a climb. Popular perception of risks associated with mountaineering is no doubt influenced by the literature and media coverage of high altitude (e.g. over 7000 metres) difficult mountaineering. This type of mountaineering entails a much higher degree of risk and tradeoffs than other types. However, this type of mountaineering is constrained to, at most, a few hundred individuals worldwide at any one time, and extrapolating these experiences to the vast majority of mountaineering activities results in a biased view of the risks.

Secondly, there is wide variability in the skill, training, experience, and judgement of mountaineers. Intuition and analysis of accidents [e.g. American Alpine Club 2004] indicate an approximate inverse relationship between these factors and the probability of injury and necessary rescue. Of course, there may be a positive correlation between the training and skill of mountaineers and the degree of difficulty of their objectives, cancelling out the previous effect. However, the degree of risk aversion in a particular climber must be accounted for, and sometimes even the best, most careful mountaineers are subject to uncontrollable risks, and simply suffer "bad luck". A special situation exists in the form of guided mountaineering trips; i.e. where the mountaineers are under the management of professional guides, who are essentially professional risk managers. There is a general lack of statistics regarding numbers of different types of mountaineers, and the skill level of those mountaineers. This makes estimation of denominators (i.e. populations at risk) difficult. Typical statistics, such as British surveys of climbing popularity [Crowe et al. 2004], do not account for these differences.

Lastly, there is a great deal of uncertainty regarding the magnitude of risk itself, even within any particular arena of mountaineering. Reliable statistics on injury or death rates for most mountain areas or types of mountaineering simply do not exist. A recent article [Athern 2004] discussed the accident related death rates associated with two mountains for which statistics do exist (Mt. Rainier in Washington and Denali in Alaska, USA). The author noted a drop in the total number of accident related deaths in recent years, but that the risk of death associated with climbing these mountains was lower than other risks such as dying of heart disease, and comparable to driving a car. However, Huey and Hornbein [2004], in another analysis which adjusted for exposure duration, estimated that the annual death rate associated with climbing Mt. Rainier is 115 times the age-standardized rate for death from heart disease. Huey and Eguskitza [2000] estimate that the death rate for climbing and descending Mt. Everest is 1 in 29, and for K2 (the second highest peak in the world and much more technically difficult than Mt. Everest) as 1 in 7, which is risky indeed (i.e. on a par with

Russian roulette). It is obviously important for mountaineers who attempt such peaks to realize what they are getting into.

Туре	Hazards	Cost <sup>2</sup>	Popularity <sup>3</sup>	Degree of risk⁴
Bolted sport climbing <sup>1</sup>	Human failure, equipment failure, travel, rock fall	Low	High	Low
Multi-pitch (i.e. more than one rope length) rock climbing	Human failure, equipment failure, travel, unstable rock conditions causing falls, rock fall, weather, getting disoriented/lost	Medium	Medium	Medium
Ski mountaineering, glacier climbing	Human failure, equipment failure, travel, suboptimal snow conditions causing falls, crevasse falls, avalanches, weather, getting disoriented/lost, cold temperatures	High	Medium	Medium to high
Waterfall ice climbing	Human failure, equipment failure, travel, unstable ice conditions causing falls, ice fall, avalanches, weather, getting disoriented/lost, cold temperatures	High	Medium	High
Alpine mountain climbing (i.e. involving technical rock, snow, ice)	Human failure, equipment failure, travel, ice conditions causing falls, ice fall, rock conditions causing falls, rock fall, avalanches, weather, getting disoriented/lost, cold temperatures	High	Low	High
High-altitude (e.g. over 7000 metres) alpine mountain climbing (i.e. involving technical rock, snow, ice)	Human failure, equipment failure, foreign travel, ice conditions causing falls, ice fall, rock conditions causing falls, rock fall, avalanches, weather, getting disoriented/lost, cold temperatures, altitude, unstable political environments in developing countries, insect-, food- and water- borne illness <sup>5</sup>	Very high	Very low	Very high

# Table 2: Comparative Aspects of Different Types of Mountaineering (note that risk associated with normal travel to climbs (e.g. driving a car) is associated with all levels)

1:Baseline comparator used is bolted sport climbing, as this is likely the safest type of outdoor climbing. Bolted climbing involves relatively "clean" rock, use of permanent expansion bolts placed in the rock for protection, typically single pitch (i.e. one rope-length) or shorter in length, close to road access.

2: Inferred from personal experience

3: Statistics are difficult to estimate, but relative popularity can be inferred from mountain club surveys such as the British Mountaineering Council (2003).

4: Relative degree of risk is inferred from accident statistics (e.g. American Alpine Club, 2005)

5: These risks are present in any outdoor activity; however, the risks are much higher in many developing countries where this type of climbing takes place.

Semi-quantitative information from organizations such as the American Alpine Club [2004], although subject to many biases and of limited utility for estimating the degree of risk, allows rough ranking of risks and thus potentially informs risk management (Table 2). For instance, according to these statistics, if one wants to avoid serious injury or death one should avoid high-altitude or waterfall ice climbing, and stick to bolted sport climbing. Relying solely upon such statistics, however, is not optimal for risk management decision-making, as will be discussed below.

Of course, risks to life and limb are only a subset of a larger suite of risks that mountaineers expose themselves to, including financial costs, impacts to careers, impacts to family and other personal relationships, and so on. These risks are harder to objectify, yet the psychological consequences can be severe. For example, a compelling account of the potentially severe impacts on the friends and families of mountaineers is found in Coffey [2005]. Regardless, at a high level in a broad risk management framework these risks should be accounted for and balanced against benefits in decision-making.

#### Benefits Associated with Mountaineering

There are benefits associated with mountaineering, but these are even more difficult to characterize and quantify than the risks. We (via our interviews and personal experience) and others have tried, but it is difficult. Some of the benefits likely and directly relate to the state of being in and managing a risky situation, as previously discussed. Many mountaineers seek some altered mental and physiological state that is a result of exertion, pain, and altitude, and with being exposed to a wide variety of difficulties and hazards. The extreme focus that is required for the simple acts of moving on a mountain and surviving the experience is addicting. In this sense, the benefit is not very different than a psychotropic drug induced state, and in a few cases is just as pathological and self-destructive (i.e. a small proportion of climbers who essentially continue to climb more and more difficult objectives, seeking these stimuli, until they exceed the limits of their abilities and they die climbing). However, for most mountaineers there is a more complex set of benefits that relate to not only the risk, but the experience of being in a wild natural environment, the camaraderie of companions, the fulfilment of a goal, and so on. Additionally, the benefits of attempting a particular mountain or route can be roughly scaled according to the "attractiveness" in terms of the condition of the rock or ice, the purity of the climbing line, the size of the route, the difficulty, and other factors.

One way of characterizing the quality of interaction between a mountaineer and a mountain is whether the mountaineer attains a state of "flow" (according to Csikszentmihalyi et al. [1991]). This transcendent state is likely what many mountaineers seek from their activity, but it can be highly transient and variable in intensity, ranging from non-existent on an unpleasant climb to many hours or even days when there is a good "match" between the person(s) and the environment. The type and degree of flow is highly dependent on the person(s), the degree of skill and training, and so forth. For example, a highly experienced mountaineer may experience a state of flow on a long, highly technical route, and indeed may seek such routes, whereas a novice may simply find the experience frightening and laborious. Of course, if mountaineering is conducted in a pair or team, there must be some degree of concordance of flow, or the partners may not attain it.

Management of risk is likely an important objective for most mountaineers, whether viewed as part of the flow phenomenon, or as a separate benefit (i.e. flow can be attained in a relatively "safe" fashion via distance running, etc.; i.e. activities not requiring the same degree of risk management). Our interviews and the literature have indicated that the continuous act of risk management is indeed a desirable objective for many mountaineers; especially those with a great deal of experience (in which case they may manage risks for apprentice partners) or for professional mountain guides. In our interviews with guides (e.g. "why do you enjoy guiding?"), as well as other studies [Beedie 2003], management of risks for clients is a consistent "reward" for guiding. Of course, there is variability in the degree of risk that guides are willing to expose clients to, ranging from simple scrambles to long difficult routes. We have asked guides "what is the most difficult climb that you would guide?", and the consistent answer was that it depended on both the experience level and risk tolerance of both the guide and the client, often with a premium added to the financial cost of the experience. In any event, the guides approach climbs that do not involve clients quite differently (e.g. they may engage in much more difficult and risky climbs). As several guides have told us, "guiding is not climbing".

It is quite unclear as to how carefully the typical mountaineer weighs benefits versus risks. We do not question the judgement of mountaineers in terms of engaging in a risky activity; rather, assuming that there are compelling reasons for individuals to engage in mountaineering, we propose that the risks can be managed in a rational, systematic fashion in a multi-objective framework. Some of the more important means and the tradeoffs are discussed below.

#### Risk Management in Mountaineering

As previously mentioned, some of the risks associated with mountaineering can be managed via safety equipment, procedures, and training, fitness, and other means (Table 3).

These reflect Reason's "layers of safety" concept that is often used as a model in the safety field [Reason 1998]. Reason's model is represented as a series of filters (e.g. slices of Swiss cheese) that sequentially reduce the probability that a person who is exposed to a hazardous scenario will experience an adverse event.

If one chooses to expose oneself to the mountains, there is limited ability to manage some "objective" risks such as rock-fall, avalanches, hidden crevasses (cracks in glaciers), extreme weather, altitude, and the like. However, careful observation and the experience to judge when to turn around or "bail" can be effective in managing these risks. In some cases, a risk-averse mountaineer may choose to avoid some types of climbing entirely due to an unwillingness to expose herself to these risks; e.g. many mountaineers choose not to climb high altitude peaks or waterfall ice for this reason. The tradeoff here relates to depriving oneself of what might be very rewarding or even life-changing experiences.

Туре	Methods	Cost	Degree of	Overall
			experience required	effectiveness in managing risk
Bolted sport climbing <sup>1</sup>	Rope <sup>2</sup> , permanent bolts, clip protection <sup>3</sup> , belay <sup>4</sup> , helmet (sometimes) <sup>5</sup> , footwear, water and nutrition	Low	Low	High
Multi-pitch (i.e. more than one rope length) rock climbing	Rope, permanent bolts, rock protection <sup>6</sup> , clip protection, belay, helmet, footwear, clothing, water and nutrition, communication <sup>7</sup> , first aid, navigational aids	Medium	Medium	Medium
Ski mountaineering, glacier climbing	Rope, snow protection <sup>8</sup> , ice protection <sup>9</sup> , prussiks <sup>10</sup> , avalanche rescue equipment <sup>11</sup> , crevasse rescue <sup>12</sup> , footwear, clothing, ski equipment, ice ax, water and nutrition, communication, first aid, navigational aids, bivouac equipment <sup>13</sup>	High	High	Medium
Waterfall ice climbing	Rope, ice protection, belay, helmet, footwear, clothing, ice tools, crampons, water and nutrition, communication, first aid, navigational aids	High	High	Medium
Alpine mountain climbing (i.e. involving technical rock, snow, ice)	Rope, snow protection, ice protection, belay, prussiks, rock protection, ice tools, crampons, helmet, crevasse rescue, footwear, clothing, water and nutrition, communication, first aid, navigational aids, bivouac equipment	High	High	Medium to low
High-altitude (e.g. over 7000 metres) alpine mountain climbing (i.e. involving technical rock, snow, ice)	Rope, snow protection, ice protection, belay, prussiks, rock protection, ice tools, crampons, helmet, crevasse rescue, footwear, clothing, water and nutrition, communication, first aid, navigational aids, bivouac equipment, oxygen <sup>14</sup>	Very high	Very high	Low

# Table 3. Equipment, Materials, and Methods for Risk Management in Mountaineering(note that in all cases proper training and experience is assumed, and professionalguiding may also be a means for risk management).

1: Baseline comparator used is bolted sport climbing, as this is likely the safest type of climbing.

2: "Rope" means use of modern dynamic rope(s), a sit harness, and an appropriate method of joining the two.

3: "Clip protection" means use of aluminium carabiners, nylon slings, and/or other appropriate means of joining the rope to protection placed in rock, ice, or snow.

- 4: "Belay" means one partner is feeding the rope through a friction device while the "leader" climbs and places protection, thus controlling the amount of distance the leader will fall. Once the leader reaches an appropriate stance, she can then belay the partner from above. In sport climbing, it is often possible to "toprope" or place a rope so that all climbers are belayed from above.
- 5: Sport climbers often do not wear helmets, as the probability of rock fall is perceived as low. This practice, in the author's opinion and based on accident statistics, is highly questionable. Ski mountaineers usually do not wear helmets unless they plan to ascend a technical peak.
- 6: "Rock protection" means aluminum or steel devices, other than expansion bolts, that are placed in cracks or pockets in the rock. There are typically removed and re-used as the team progresses up the climb.
- 7: "Communication" means radios, cell phones, satellite phones as appropriate
- 8: "Snow protection" involves devices that can be hammered into or buried in snow.
- 9: "Ice protection" means hollow screw-like devices that are screwed or hammered into ice.
- 10: "Prussiks" means devices that are placed on the rope to allow climbing up the rope itself in the event of a fall into a crevasse.
- 11: "Avalanche rescue equipment" involves electronic transceivers, shovels, probes, balloon packs (i.e. personal airbags), etc. that allow finding and retrieving buried victims. These are routinely used only in ski mountaineering. In other forms of mountaineering, the weight of carrying such items is traded off the utility in terms of rescue; thus they are usually not carried.
- 12: "Crevasse rescue" means knowledge of processes to extract a disabled person from a crevasse. This may involve a range of processes from simply pulling someone out to complex pulley systems, rappelling into the crevasse, etc. Ski mountaineers do not typically belay each other, but they are exposed to crevasse hazards on glaciers.
- 13: "Bivouac equipment" means minimal camping equipment such as a sleeping bag, stove, light tent or cover bag, etc., as opposed to full camping equipment used at a base camp. Carrying such equipment depends on the length of the climb.
- 14: The use of oxygen on very high peaks (e.g. over 8000 metres) has historically been a risk management measure. However, in recent years many high-altitude climbers have been eschewing oxygen, under the concept that the use of oxygen is not "fair means" (attributed to the Austrian climber Reinhold Messner, who was the first person to summit Mt. Everest without oxygen). This likely increases risk many-fold.

Use of safety equipment is an important risk management technique. For instance, a rope can prevent major injuries or death by limiting the distance of a fall, and wearing a helmet can prevents the same outcomes via reduction of head trauma in the case of rock or ice fall. All this comes at a price; modern active mountaineers typically spend thousands of dollars per year on equipment and technical clothing. The degree to which equipment can alter the risk profile of an individual is dependent on the degree of training with and thoughtful use of the equipment; accident reports corroborate this as a major factor in many accidents [American Alpine Club 2004]. An odd cognitive phenomenon among some users, especially with regard to wearing a rope, is that mere donning of the equipment confers some level of safety. This is obviously fallacious.

The degree of physical and mental fitness of the climber is paramount in mountain safety. A fit climber is able to move faster and more confidently, and to retain clear thinking in a state of fatigue. Speed is an oft-overlooked risk management strategy; many successful climbers take an approach to safety that involves getting up and down a climb as fast as possible (e.g. Easton [2005]). This is rational because the time of exposure to uncontrollable hazards is reduced if one climbs as fast as possible. Of course, there are tradeoffs; the safest strategy, all other things being equal, is to stay out of the mountains entirely, and if one climbs too hastily one might make fatal errors. Assuming one exposes oneself to the mountains, one must climb often to gain the degree of fitness necessary to climb fast; i.e. for

training there is no substitute for actual climbing. Thus, an odd tradeoff exists in which the safest climber is one who climbs fast, but in order to climb fast one must climb often (thus increasing exposure time). Thus, the two factors may indeed cancel each other out in terms of cumulative risk. Additionally, in order to climb fast, one must generally carry less weight, which means less equipment, food, and water. Many of the risks associated with high altitude mountaineering relate to compromised physical and mental functioning; e.g. mountaineers make mistakes at high altitude that they would never make at lower altitudes. For example, the popular book "Into Thin Air" [Krakauer 1999] is a highly critical exploration of how rational decision-making is compromised at high altitude (in this case, Mt. Everest).

A critical variable in mountaineering safety relates to one's partner(s). In roped climbing one partner leads while the other follows, and a variety of techniques have been developed so that the partners manage each other's safety. Additionally, the state of mind of the partners may influence each other in both positive and negative directions. In many instances, the life of one partner is literally in the hands of the other. A fascinating account of partners who found themselves in an extreme decision to compromise the security of the rope is found in the book (and subsequent documentary) "Touching the Void" [Simpson 1989], in which one partner cuts the rope (believing the other partner to be dead, which turned out to be false) to save himself. We have asked many climbers whether they would do this, and despite the clear mental torture that resulted from this act and is well-communicated in the book, the majority have said yes. This may reflect an over-riding sense of self-preservation, or simple lack of concern for the other party; the spread of psychological reasons behind this is likely to be large. Our personal experience is unusual because we are married climbing partners, and we expect that such a choice would be extremely difficult.

Professional guides manage the risks, to the extent possible, for their clients. They have therefore developed a series of techniques and decision rules for their business. Guides will not generally take clients on highly difficult routes, as the degree to which the guide can manage the risks is reduced as the climb becomes more difficult. The tradeoffs for the client relate to expense and loss of some control over the experience of the climb. The tradeoffs for the guide may relate to having to manage the client, as opposed to simply climbing with a partner, although they are compensated for this by their fee.

Rescue is the ultimate retroactive risk management method in the event of a non-fatal accident. The existence of timely and efficient rescue capability is highly variable, however, ranging from a high degree of capacity in some US national parks and some areas of Europe to non-existence in remote wilderness ranges. The question as to whether existence of rescue capacity (and ready communication in the form of mobile and satellite telephones) has an impact on the behaviour of mountaineers is unanswered. Regardless, the ability of a team to self-rescue or facilitate rescue likely has a large impact on the survival of injured mountaineers in many instances.

Mountaineering is thus a sport of numerous multi-objective tradeoffs. In the next section we explore this in greater detail using a case study, and introduce the concept of dynamic multi-objective risk management under the PrOACT framework.

### **CASE STUDY: CLIMBING MT. ROBSON**

We use the case of Mt. Robson, which at 3954 metres is the highest peak in the Canadian Rockies. It is considered a "classic" mountain in terms of its size, its beauty, and the difficulty of attaining the peak. We present a hierarchical multi-objective structure, based on PrOACT, for disaggregating the decisions involved in climbing this peak. "Macro-" level decisions involve lifestyle decisions around "whether to climb", including the type of mountaineering to pursue. "Meso-" level decisions involve "what and where to climb". "Micro-" level decisions involve "what and where to climb". "Micro-" level decisions involve "to climb", including risk management on a continuous and dynamic basis (Table 4). An objectives hierarchy is presented in Figure 1.

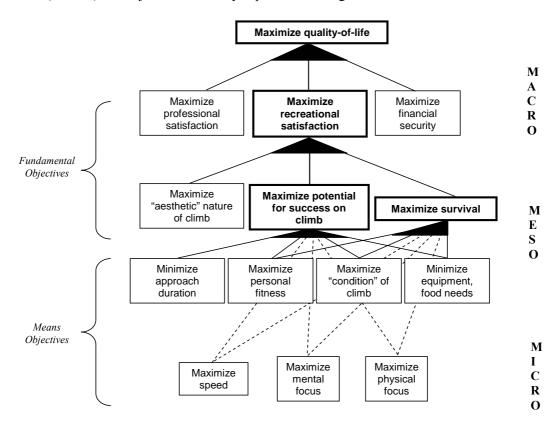


Figure 1: Objectives hierarchy for mountaineering decisions. The bolded boxes are fundamental objectives for which the illustrated means objectives are relevant. Note that for clarity the hierarchy is only partially completed; it focuses on objectives that are directly related to the act of mountaineering itself.

#### Macro Decisions

*Problem:* We want to maximize our quality-of-life. How should mountaineering be part of this?

*Objectives:* We want to maximize our recreational satisfaction, as well as professional satisfaction and financial security (Table 4).

# Table 4. The PrOACT framework as Applied to Multiple Levels of Mountaineering Decision-Making.(note that these are examples only; a complete assessment would involve many more considerations.Uncertainties, risk tolerance, and linked decisions are discussed in the text.

Level	Problem	Objectives	Alternatives	Consequences	Tradeoffs
Macro	How does one maximize quality of life? How does recreational satisfaction fit into this? Should one engage in mountaineering?	Maximize recreational satisfaction (f) <sup>1</sup> ; also maximize professional satisfaction (f), maximize financial security (f)	Multiple; outdoor activity examples include walking, cycling, skiing, mountaineering, etc.	Positive: physical and mental fitness, enjoyment of nature, sense of accomplishment Negative: injury, illness, death	Increase in recreational satisfaction results in potentially limited career choices, increases in cost Choice of mountaineering results in increased physical fitness, but exposes one to risk
Meso	Where/what does one climb?	Maximize "aesthetic" nature of climb (f), maximize potential for success (f), minimize travel time (f), minimize duplication of climbs (f), minimize approach duration (m) <sup>2</sup> , maximize person fitness (m), maximize "condition" of climb (m), minimize equipment and food needs (m)	Multiple; depending on the type of climbing one chooses and where one lives	Positive: enjoyment of climb, feeling of success, travel to new area Negative: unpleasant or adverse conditions, sense of failure	Choosing a large peak such as Mt. Robson maximizes "aesthetic" nature of climb, but limits potential for success Choosing a peak that is difficult provides challenge, but potentially more risk
Micro	How does one climb fast yet safely?	Maximize survival (f), maximize success (f), maximize speed (m), maximize physical and mental fitness/focus (m0, maximize safety equipment (m), maximize nutrition and hydration (m)	Choice of training, equipment (including accounting for rare risks), how much and what kind of nutrition	Positive: completing the climb safely Negative: unintended bivouac, injury or death	Increased training time vs. actual climbing time Too much equipment vs. too little

1: "f" represents a fundamental objective

2: "m" represents a means objective

*Alternatives:* There are innumerable options in terms of recreational activities, ranging from reading books about climbing mountains to actually climbing them. Why a particularly risky and costly choice? The benefits of mountaineering must outweigh the risks and costs.

*Consequences:* The positive consequences of mountaineering, as opposed to other activities, include a high degree of physical and mental fitness, a sense of flow, gratification from managing risks, an increased appreciation for the natural environment, and so on. The negative consequences of mountaineering can be severe, including disability and death. Aside from obvious health and safety risks and impact on families, there are consequences that impact our non-climbing lives. For example, we choose to live near mountains, which can be professionally constraining and stressful. We spend an appreciable amount of money and time on travel and mountaineering, so we must choose jobs that pay enough to be able to do this and which allow flexible time.

*Tradeoffs:* There are large degrees of tradeoffs across the objectives in Table 4. In choosing mountaineering over other activities to maximize recreational satisfaction, we are limited in our jobs and where we live, and we spend a lot of money. If, for example, we chose to read books about mountaineering for recreation, we would be able to live anywhere and perhaps would not have to make as much money. However, we would not enjoy the benefits of mountaineering.

At the time of this writing, considering the tradeoffs, we chose to live near the Canadian Rockies, thus shorter scale decisions centred on that mountain range. For example, a yearly-level decision in 2004 was where to spend our annual holiday. We decided on a number of objectives, based on the appeal of the climbs and our projected ability to complete them.

#### Meso Decisions

Choosing and planning a particular climb represents the next level of decision-making. The PrOACT framework for this level is illustrated in Table 4. We desired to climb an "elegant" route on a major peak that we had not climbed before, and which was well within our capabilities and our available time. The choices were somewhat limited, as we had already climbed many peaks in the Rockies. We chose a relatively straightforward but classic route, the Kain Route on Mt. Robson (named after the person who guided the first ascent in 1913), which involves a long hike, extensive glacier travel to a base camp, a 500 metre 50 degree snow and ice face (the "Kain Face"), and another 300 metres of steep snow and ice ridge to the summit.

Getting to the mountain involved a 5 hour drive from our home, thus includes some degree of risk, cost, and time commitment even before reaching the mountain. Typically, parties allow a week to climb the mountain (to allow for the approach, the climb, and descent, and delays due to bad weather). After discussions with professional guides, we decided to avoid the approach by flying via helicopter to the base of the Kain Face in order to maximize our chance of success (but which of course introduces a new set of risks!). The reasonable expense of flying and the risk of flying were outweighed by the avoidance of a 30 kilometer hike with 25 kilogram backpacks and travel across a heavily crevassed glacier, which can be quite risky.

As we had previously climbed many similar glaciated peaks and were accustomed to this sort of climbing, we did not need to buy any additional equipment or undergo any training, other than climbing to stay fit. We did need to carefully consider our equipment, clothing, and food; for camping in a harsh environment as well as climbing (including a satellite phone for rescue, if necessary). In order to complete the climb safely in a day from a base camp, a team has to move continuously and fast. Also, the mountain becomes more amenable to climbing later in the season, as the amount of arduous snow climbing (as opposed to ice, which is faster) is reduced. Thus, we chose to attempt the mountain at the end of our summer holiday, when we would have been climbing for a couple of weeks and would be as fit as possible, and the mountain would be in good condition. However, this also meant that the length of day was reduced (compared to early summer), thus some amount of climbing in the dark was necessary, which is riskier.

A major source of uncertainty at this level of decision-making regarded the weather. Based on climate statistics, the timing of our trip was marginal; i.e. there was a better-thaneven chance that we would have bad weather. However, we "lucked out", and the weather was good (we had backup plans in case the weather was too bad for the helicopter to fly). We arrived on the mountain, and the next level of decision-making began.

#### **Micro Decisions**

See Table 4 for the PrOACT framework applied at this level. Once on the mountain, the fundamental objectives are maximizing success and survival. There were decisions as to equipment to take on the day of the climb (i.e. as little as possible, while keeping a reasonable margin of safety), food and water needs, timing and "style". Based on experience, we decided to take the "clothes on our backs", a minimum of climbing equipment, and enough food and water to sustain us for approximately 20 hours. We also carried lightweight radios to facilitate communication between us. We planned to leave camp as early as possible in the morning while getting a modicum of sleep (i.e. 2 a.m.). This of course meant that we would have to climb by headlamp for several hours, which obviously can be riskier than climbing in the daylight. We awoke to a clear starry sky (a critical decision-point) the next morning, and the climb began.

Once climbing, the level of decision-making becomes much more focused and microscale. We decided to use the style know as "simul-climbing" in which we both moved continuously on a rope, with ice screws placed into the glacier ice between us at all times to theoretically stop a fall (as opposed to belaying and moving one at a time). There is a risk-risk tradeoff here; belaying would be safer in the event of a fall, but it is slower, and as previously discussed, speed is safety. However, the main form of security on a steep ice face is to ensure that one is connected to the mountain at all times via ice tools for the hands and crampons (spikes) on the feet. There are second-by-second decisions as to placement of hands and feet, and a wrong decision can be deadly. There is no reason or indeed opportunity to consider whether the bills had been paid before we left, how much the new ice tool cost, or whether another mountain would have been more rewarding. The state of rapid risk assessment and management, which indeed occurs at an "automatic" (i.e. without conscious thought) level to some extent, exists for as long as the climber is moving, and can be more draining mentally than physically. However, this is part of the attraction of technical climbing for most people; the focus on moving safely can be rewarding in itself as a source of flow. At its simplest level, climbing becomes a single-objective problem: maximizing survival. Risks are not ignored, but managed on a move-by-move basis and performed at an instinctual, intuitive level that is only gained by experience.

Time is an enemy to a mountaineer. On a long climb like Mt. Robson, survival is maximized by completing the climb as quickly as possible. Neither of us wear watches on our wrists, because we have found that knowing the rate of time passage impedes our abilities to attain a sense of flow. There is also a sense (for us at least, but this is by no means universal as we discovered via interviews) that if we know the actual rate of time passage, that we become apprehensive that we are climbing too slowly. Yet, it is useful (*sans* observations of the sun or other indications) to know if it is getting too late to complete the climb; thus we carry a watch hidden in a pocket until it is needed to inform major decisions.

It becomes more difficult to maintain focus as the climber becomes fatigued, thus careful attention to hydration and blood sugar levels is important. Recent advances in sports nutrition have resulted in products such as energy drinks and glucose packets that allow rapid ingestion and absorption, thus allowing more continuous climbing without the need to stop for extended periods of time.

Additionally, there are minute-by-minute and hour-by-hour choices regarding the route itself in terms of condition and ease of ascent. We chose a line on the Kain Face that was the most direct and icy route to the summit ridge (for the sake of efficiency and speed), yet this line was exposed to serac (section of unstable glacier ice) hazard. This was managed by completing the Face before the sun arose; it would have been much riskier later in the day. On this particular mountain, a major uncontrollable variable was the weather. Mt. Robson, like many big mountains, can create its own weather patterns. A cloud cap can form on the mountain that can result in "white-out" and windy conditions where it is difficult or impossible to see for more than a few yards. Such conditions are highly risky, as it becomes easy to become disoriented and indeed, at worst, to step off a cliff or cornice (overhanging snow) into oblivion. As we climbed the mountain, a cloud cap started to form, thus the "need for speed" intensified. However, by this point the snow was knee-deep (despite our planning) and climbing fast was difficult. Minute-by-minute decisions were made regarding the speed of formation of the cloud cap versus our own speed of ascent. We found ourselves in a conflicted state, as the line between benefit (attaining the summit) and risk became more and more "fuzzy". This corresponds to the mountaineer/writer David Roberts' "moments of doubt" [Roberts 1986]. At such a point, it was advisable to facilitate rational thinking via a brief rest, food, and water. As we were returning the same way, we were able to make the rational decision that we would be able to follow our track, even if it became cloudy. Thus, we were able to attain the summit and enjoy spectacular views. However, we were not able to spend much time enjoying this, as we had to immediately start our descent given the change in the weather.

Descending a mountain is often just as or more risky than ascending. Rappelling (sliding down the rope in a controlled fashion using a friction device) down a climb is one of the most risky aspects of climbing, because one is totally dependant on the integrity of the anchors that are set (as opposed to using one's hands and feet to hold on), plus it is often slower than down-climbing. In our case, we chose to rappel the steepest parts of the descent (the upper part of the Kain Face), but mainly down-climbed for the sake of speed. We made it back to camp 15 hours later, with the upper part of the mountain now immersed in cloud.

As the weather had deteriorated and in order to physically recover, we made the choice to spend the next day at our base camp. However, it is risky to spend time at this elevation on the mountain, as the weather can be extreme. We managed this risk by using a very strong tent and building snow walls around the tent. Despite being ensconced in such a "bombshelter", it is difficult to sleep, and most mountaineers have a sense of apprehension, ranging from mild concern to an overwhelming dread of being blown away (a nontrivial concern on high mountains).

We had previously decided not to fly out, due to the expense, plus we wanted to experience what is commonly considered as one of the most scenic walks in the Canadian Rockies. Unfortunately, the conditions were such that we had to bypass the typical route down the upper Robson Glacier due to its instability (i.e. teetering seracs the size of houses), and take a more circuitous and unpleasant ridge that involved exposed rock climbing with heavy packs. Regardless, a day's worth of glacier travel (which involves continuous, dynamic, and draining decision-making as to choice of route in order to avoid crevasses and detours) and a long day's hike later we completed our climb. The mental relief that is felt by the simple state of being off the mountain and in a hot tub is indescribable.

Additional Proact Considerations: Uncertainty, Risk Tolerance, and Linked Decisions

The PrOACT framework includes consideration of uncertainty as an important step in informed decisions. As described above, mountaineering is rife with uncertainty. In particular, there is always an element of "luck", bad or good, in mountaineering, which can be characterized as random influences that are highly uncertain. For example, a novice mountaineer may cross an avalanche slope unscathed, while a highly skilled and trained mountaineer who has performed multiple stability tests may still be caught in an avalanche. The mountain environment is by its very nature highly complex and uncertain, and risks cannot be entirely managed. Acknowledgement of this fact is necessary for informed decision-making.

Risk tolerance of mountaineers is highly variable, both within an individual and across individuals. Within an individual, it may indeed depend on how well the person feels, how much coffee she has consumed, the "intimidation factor" of a particular climb, and so on. There is some average degree of risk tolerance that may exist, and which for example may determine the type of climbs that are chosen at the meso level, but at the micro-level raw fear may outweigh any sort of rational behaviour if the individual's risk tolerance is exceeded. Being aware of and not exceeding one's risk tolerance level is thus critical for rational decision-making.

Linked decisions are also important at all levels of the mountaineering decision-making process. As we have presented, most of these decisions are hierarchical; the decision to engage in mountaineering influences the decisions on what to climb, which in turn influences the micro-level decisions. The hierarchy is also reversed; ones' experience on a particular climb (positive or negative) may influence macro-level decisions. A wrong decision at the micro-level can change or end one's life in a split second.

## CONCLUSION

Decision making under "risk" has been studied extensively for decades (e.g. Raiffa [1968]). However, the "risk" in the majority of case studies relates to simple opportunity cost, and in most cases is defined for a single strategy or decision [Loewenstein 1999]. This paper has explored dynamic, hierarchical decision making under risk where the risks involve personal injury or death, and where the activity is completely voluntary.

We have not presented a quantitative analysis (e.g. a formal decision analysis) for several reasons. First, the intent of our paper was to explore the psychology and demonstrate a way of thinking about voluntary risk that may clarify what may be, in some instances, life-and-death decisions. In our opinion, the critical aspect of this decision-making exercise is to recognize that there are substantial tradeoffs associated with choosing to engage in and actually engaging in risky recreation. Second, quantitative information associated with different types and aspects of mountaineering is sadly lacking. Indeed, there is little peer-reviewed literature at all related to such activities. Third, as previously discussed, there is difficulty in quantifying or even scaling objectives such as the benefits associated with mountaineering. Fourth, in order to properly model these decisions a complex hierarchical and dynamic probabilistic model would have to be employed. Few (if any) people would actually use such a model in making voluntary risky decisions, yet many may use a qualitative framework for informing their personal risky choices. Indeed, land managers and regulators responsible for areas where such activities take place may use similar frameworks in making programmatic and planning decisions. Last, there are aspects of decision-making throughout the hierarchy that are intuitive as opposed to conscious, particularly at the micro-level, which obviously makes quantification difficult.

However, we argue that qualitative multi-objective thinking is nonetheless useful and informative. Our discussions with professional guides have confirmed this, particularly in terms of educating less-experienced mountaineers about risk management. It would be difficult to estimate, say, the marginal utility of using our framework versus not using it, but as there essentially are no disutilities or downsides associated with using it, the marginal utility is likely to be positive.

We have not explored the myriad heuristic biases inherent in mountaineering decisions. For example, "projection bias" relates to making non-objective judgements about the future or the past, because the risks or benefits are "distant" in time and may not be evaluated properly [Loewenstein et al. 2003]. A common experience among mountaineers is to unfairly weight the benefits of an unpleasant climb as the memory of the unpleasant aspects fades (in fact, without this heuristic bias, many mountaineers would quit!). Group behaviour may obviously influence individual decision-making in either positive or negative directions. Considerations such as age and family may affect individual risk aversion. Although discussions of these issues are common in the popular literature, there has been little primary research, and thus may be fertile ground.

Considering the increasing popularity of "extreme" or "adventure" recreation, it may behoove participants, organizers/guides, and land managers to think more carefully about the risks and management of the risks, or decisions may end up in the hands of risk-averse arbiters such as lawyers or government agencies. For instance, the lack of injury and death statistics in mountaineering, and likely even more so in other risky activities, makes it difficult to conduct risk assessments. Societal risk related decisions in the face of uncertainty are often driven by media attention and reaction or public outcry, rather than proactive and careful consideration. There is ample opportunity for primary research in decision-making under risk and uncertainty in this area.

We have no illusions that mountaineering itself is, in the popular view, a rational activity. However, we have illustrated a framework and process of rational thinking and risk management and how it may contribute to safety for a risky sport. Our multi-objective framework may be useful for both individuals and society in understanding risk management in such activities. We hope to follow-up with studies designed to collect empirical data on how such a framework may contribute to safety in recreational mountaineering.

#### **ACKNOWLEDGMENTS**

This work was funded by the Calgary Health Technology Implementation Unit (University of Calgary, Alberta, Canada). The authors would like to thank the following individuals (all experienced mountaineers with professional and/or academic appointments) for useful discussions and comments: Dr. Bruce Jamieson, Dr. Ian McCammon, Dr. Pascal Haegeli, and Ian Tomm. We would also like to thank several mountain guides and professional and amateur mountaineers who provided useful input, as well as reviewers of initial drafts of this chapter. Some of the ideas presented in this paper were presented to an audience of professional mountain guides at a Canadian Avalanche Association workshop, "Professionalism at a Crossroads", in Penticton, BC in May 2005.

#### REFERENCES

- American Alpine Club (2004). Accidents in North American Mountaineering. Vol. 8, No. 4, Issue 57.
- Athern, L. (2004) The risks of mountaineering put in perspective. *American Alpine Club News*, Vol. 11, No. 246. www.americanalpineclub.org.
- Beedie, P. (2003) Mountain guiding and adventure tourism: Reflections on the choreography of the experience. *Leisure Studies* 22, pp. 147-167.
- Celsi, R.L., Rose, R.L., Leigh, T.W. (1993) An exploration of high-risk leisure consumption through skydiving. *Journal of Consumer Research* 20, pp. 1-23.
- Coffey, M. (2005) Where the Mountain Casts Its Shadow : The Dark Side of Extreme Adventure (St. Martins Griffin).
- Cox, L.A. (2002) *Risk Analysis: Foundations, Models, and Methods* (Boston, Kluwer Academic Publishers).
- Creyer, E.H., Ross, W.T., Evers, D. (2003) Risky recreation: An exploration of factors influencing the likelihood of participation and the effects of experience. *Leisure Studies* 22, pp. 239-253.
- Crowe, L., Mulder, C. (2004) Demand for Outdoor Recreation in the English National Parks (Sheffield Hallam University, Centre for Environmental Conservation and Outdoor Leisure).

- Csikszentmihalyi, M. (1991) *Flow: The Psychology of Optimal Experience* (New York, Harper Collins).
- Delle Fave, A., Bassi, M., Massimini, F. (2003) Quality of experience and risk perception in high-altitue rock climbing. *Journal of Applied Sport Psychology* 15, pp. 82-98.
- Easton, S. (2005) How (not to) climb under seracs. *Rockies Ice Specialists*. http://www.rockies-ice.com/detail\_news.php?nid=17.
- Hammond, J.S., Keeney, R.L., Raiffa, H. (1998) Smart Choices: A Practical Guide to Making Better Decisions (Cambridge, Harvard Business Review).
- Heywood, I. (1994) Urgent dreams: Climbing, rationalization, and ambivalence. *Leisure Studies* 13, pp. 179-194.
- Holyfield, L., Jonas, L., Zajicek, A. (2005). Adventure without risk is like Disneyland. In: Lyng S. (ed.) *Edgework: The Sociology of Risk-Taking* (New York, Routledge).
- Huey, R., Eguskitza, X. (2000) Supplemental oxygen and mountaineer death rates on Everest and K2. *Journal of the American Medical Association* 284(2) p. 181.
- Huey, R., Hornbein, T. (2004) Risks of mountaineering. *American Alpine Club News*, Vol. 11, No. 247. www.americanalpineclub.org.
- Keeney, R.L. (1992) Value-Focused Thinking: A Path to Creative Decisionmaking (Cambridge, Harvard University Press).
- Kiewa, J. (2002) Traditional climbing: Metaphor of resistance or metanarrative of oppression? *Leisure Studies* 21, pp. 145-161.
- Krakauer, J. (1999) Into Thin Air: A Personal Account of the Mt. Everest Disaster (New York, Anchor Publishing).
- Loewenstein, G., O'Donoghue, T., Rabin, M. (2003) Projection bias in predicting future utility. *Quarterly Journal of Economics*, 118, pp. 1209-1248.
- Lowenstein, G. (1999) Because it is there: The challenge of mountaineering for utility theory. *Kyklos* 52, pp. 315-343.
- Lyng, S. (1990) Edgework: A social psychological analysis of voluntary risk taking. American Journal of Sociology 95(4), pp. 851-886.
- Lyng, S. (2005) Edgework: The Sociology of Risk-Taking (New York, Routledge).
- McDaniels, T.L., Kamlet, M.S., Fischer, G.W. (1992) Risk perception and the value of safety. *Risk Analysis*, 12(4), 495-503.
- McNamee, M.J., Sheridan, H., Buswell, J. (2001) The limits of utilitarianism as a professional ethic in public sector leisure policy and provision. *Leisure Studies* 20, pp. 173-197.
- Olivier, S. (2006) Moral dilemmas of participation in dangerous leisure activities. *Leisure Studies* 25, pp. 95-109.
- Raiffa, H. (1968) Decision Analysis: Introductory Lectures on Choices under Uncertainty (Reading, Addison-Wesley).
- Reason, J.T. (1998). *Managing the Risks of Organizational Accidents* (Hampshire, Ashgate Publishing Ltd.).
- Roberts, D. (1986) *Moments of Doubt and Other Mountaineering Writings* (Seattle, Mountaineers Books)
- Simpson, J. (1989) *Touching the Void: The True Story of One Man's Miraculous Survival* (New York, Harper-Collins).
- Slovic, P. (1986) Informing and educating the public about risk. *Risk Analysis*, 6(4), pp. 403-415.

Thompson, H. (1974) Interview by *Playboy Magazine*, November, available at http://www.playboy.com/features/features/hunterthompson/04.html

Chapter 40

# DECIDING ABOUT RISKY PROSPECTS: A PSYCHOLOGICAL DESCRIPTIVE APPROACH<sup>\*</sup>

# Petko Kusev<sup>1</sup> and Paul van Schaik<sup>2,†</sup>

<sup>1</sup> Department of Psychology, School of Social Sciences, CityUniversity, Northampton Square, London, EC1V 0HB, United Kingdom
<sup>2</sup>Psychology subject group, School of Social Sciences and Law, University of Teesside, Borough Road, Middlesbrough, TS1 3BA, United Kingdom

## ABSTRACT

Results from many studies suggest that people violate the principles of rational choice in both the domain of gain and that of loss. People usually treat probabilities nonlinearly by overweighting low and underweighting moderate and large probabilities. The violations of rational choice in human decision-making were disregarded by the normative point of view (von Neumann & Morgenstern's (1947) Expected Utility theory, EUT) until Allais (1953) and Kahneman and Tversky (Kahneman & Tversky, 1979; Tversky & Kahneman, 1981, 1986, 1992) developed a descriptive theoretical approach. In this chapter we investigate what might affect decision-makers' preferences with respect to described real-world protective prospects. People's precautionary ('protective') decision-making in the face of risk implies that they may judge and weight the probability of risky events in characteristic ways that deviate from both normative EUT and psychological descriptive theory of decision-making tested with abstract gambles. The following theoretical frameworks contribute to an explanation of protective decisionmaking: (a) experience-based decision-making models - past and immediate experience affect decision-makers' preferences (Dougherty et al., 1999; Frisch, 1993; Hertwig, Barron, Weber & Erev, 2004; Stewart, Chater & Brown, 2006; Tversky & Koehler, 1994)

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: Psychology subject group, School of Social Sciences and Law, University of Teesside,Middlesbrough,TS1 3BA,United Kingdom e-mail: P.Van-Schaik@tees.ac.uk, Tel.: 01642 342320/342301 (from UK); +44 1642 342320/342301 (from abroad), Fax: 01642 342399 (from UK); +44 1642 342399 (from abroad), Web: http://sss-studnet.tees.ac.uk/psychology/staff/Paul\_vs/index.htm

- and (b) accessibility of information (Higgins, 1996; Kahneman, 2003; Koriat & Levy-Sadot, 2001) - not all available observations of risks are equally accessible in memory.

#### **1. INTRODUCTION**

Existing decision-making research has used formal models that are assumed to apply equally to (a) both rational and human decision-making and (b) both hypothetical and realworld tasks that humans perform. However, research has shown that human-decision making is not necessarily rational and there is increasing evidence that humans perform hypothetical and given real-world tasks differently. In this chapter, we attempt to demonstrate the complexity of people's judgments and decision-making under risk and uncertainty with given real-world tasks, focusing on precautionary ('protective') real-world decision-making tasks and hypothetical decision-making tasks (using abstract gambles). We define protective decisions as those that involve participants deciding whether or not to adopt a specified precaution with a known cost in the face of a described risk. Indeed, precautionary behavior can be described as behavior where people aim to avoid or reduce risks by taking protective actions.

As discussed in this chapter, research findings contradict two main types of theory of decision-making - normative (e.g., von Neumann & Morgenstern, 1947) and descriptive psychological (e.g., Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). Both types assume that people's risk preferences and decisions under risk and uncertainty are taskindependent and can be represented as hypothetical gambles. The nature of the content area being contemplated may influence judgments of the degree of risk and benefit (e.g., Slovic, 1987), but most prominent decision theories assume that, once these judgments are made, decision-making with risky prospects is not influenced by factors associated with this content and is independent of the decision-task. For instance, the decision whether or not to insure my luggage worth £500 for a cost of £5 where the risk of loss is 1% is identical to the decision to pay £5 or take a gamble where I have a 1% chance of losing £500, but the results of decisionmaking in these two scenarios differ (Kusev, Ayton, van Schaik & Chater, 2006, 2007). Therefore, although existing studies of decision-making using hypothetical (abstract) gambles corroborate the need for a distinction between normative and descriptive accounts, evidence suggests they cannot provide all the answers that we need to explain observable behavior in the face of risk. We can assume that given hypothetical tasks do not have a similar experienced 'background' as given real-world protective tasks (e.g., insurance decisions) and this can account for (a) predominantly risk-averse preferences and (b) a lack of loss aversion, where the risk associated with the protective category is exaggerated.

The aim of this chapter is to argue for a differentiation between models developed to account for precautionary decision-making and models of other types of decision-making under risk (Kusev et al., 2006, 2007) and to offer possible psychological mechanisms which might affect people's risk preferences in precautionary-decision-making tasks. We first outline the normative theory of decision-making and its violations by human decision-makers. We then present the psychological descriptive approach that attempts to account for these violations. However, because this approach cannot account for differences between human decision-making in hypothetical decision tasks and that in given real-world decision tasks, we

discuss theoretical approaches that contribute to an explanation of protective decision-making as well as hypothetical decision-making.

## 2. NORMATIVE THEORY OF DECISION-MAKING UNDER RISK AND UNCERTAINTY

#### 2.1 The Normative Approach

The study of decision-making contains both normative and descriptive directions. Normative analysis is concerned with the nature of rationality and the logic of decision-making, while descriptive analysis is concerned with people's beliefs, preferences and imagination as they are, not as they should be (Kahneman & Tversky, 1984). The dominant normative approach to or the theory of rational choice - Expected Utility theory (EUT) (von Neumann & Morgenstern, 1947) - can be interpreted as a theory describing the behavior of an idealized human decision-maker. EUT is an early normative theory of decision-making, originally formulated by Bernoulli (1954), revived and axiomatized by von Neumann and Morgenstern (1947) and then extended by Savage (1954) to non-monetary outcomes and conditions of uncertainty rather than risk. The theory argues that if people are willing to accept a number of axioms or basic principles of behavior (some of which are discussed below), then this logically defines a unique criterion for rational choice.

EUT maintains that, facing uncertainty, people behave (in a descriptive interpretation) or should behave (in a normative interpretation) as if they were maximizing the expectation of some utility function of the possible outcomes. One of the central topics of decision theory is providing a justification for making wise or rational decisions under conditions of risk and uncertainty. EUT provides such a rational criterion, demonstrating that if individuals' preferences satisfy certain basic axioms of rational behavior then their decisions can be described as the maximization (minimization) of expected utility (disutility) (see Figure 1). Maximization of expected utility (EU) appears to be a basic characteristic of rational behavior because it is obtained from axiomatic principles that presumably would be accepted by any rational human (Slovic, 2000).

If we consider a gamble that gives  $P_i$  chance at outcome  $X_i$ , the EU of this gamble is

 $\sum P_i U(X_i)$ , where  $U(X_i)$  measures the *utility* of receiving  $X_i$ . The basic principle of EUT is that decisions under risk are made so as to maximize EU and to minimize the losses (disutility). Traditionally, the process of maximization and minimization is represented by the shape of the utility or disutility function; it has been assumed that the functions are concave for utility of wealth - maximization of wealth, and convex for disutility of loss - minimization of loss, as in Figure 1.

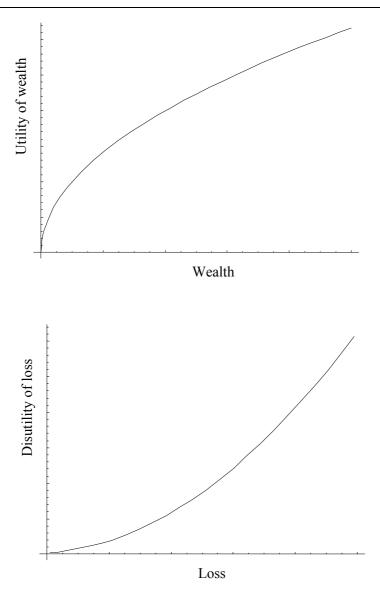


Figure 1. Traditional functions of utility and disutility.

Savage's generalization of EUT (Savage, 1954) allows it to be applied to decision situations where no objective mathematical probabilities are available and where judgments may be no more than expressed individual beliefs about likelihoods or *subjective probabilities* (Ayton, 2005). According to Subjective Expected Utility theory (SEUT), if decision-makers violate one or more of the axioms, then their choices will not maximize EU and will not be normative. Some important principles of EUT related to the psychology of decision-making are *cancellation, dominance* and *invariance*. The cancellation principle argues that with a prospect of two *risky* alternatives, a choice between two alternatives should depend only on those attributes that differ, not on outcomes that are the same for both alternatives, that is common factors should cancel out. The dominance principle proclaims that perfectly rational decision-makers should never choose a dominated option, even if the strategy is only weakly

dominated. For example, according to EUT a decision-maker should never choose product B, if product A strongly dominates product B (e.g., A has better price and quality) or product A weakly dominates B (e.g., A has better quality than product B but is equivalent in cost). The invariance principle assumes that decision-makers are not affected by the format of presented prospects. For example, a rational decision-maker should have no preferences between (a) a compound gamble and (b) a simple gamble to which (a) can be reduced (e.g., one stage lottery) (Plous, 1993).

In the next section, we begin by discussing the systematic and overwhelming evidence of violations of the predictions of EUT and SEUT. We then present two of the leading psychological alternative (descriptive) theories that are vying to replace the normative models.

#### 2.2 Violations of Normative Decision Theory

Despite the intuitive appeal of the rational paradigm, EUT and SEUT models suffer from some serious problems. Most critically, studies using both experimental and *real-world* evidence have repeatedly found that individuals often do not behave in a manner consistent with these models. The paradoxical choices identified by Allais (1953) and Ellsberg (1961) demonstrated the first doubts regarding the still popular rational EUT paradigm. However, the main problem for EUT and supporting economics models is their failure to explain the lack of stability in people's risk preferences - people demonstrate both risk-aversive and risk-seeking preferences with different forms (concave and convex) of the utility and disutility function (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). In particular, there are differences among the different decision domains (gain and loss) in terms of the value function and probability-weighting function, as discussed in this chapter.

EUT does not offer an account for common problems in decision-making and choice phenomena affected by uncertainty and probability - problems, which have dominated decision research over the years. For example, the Allais paradox (also known as common consequence effects and common ratio effects) played a particularly significant role in stimulating and shaping the development of alternatives to EUT. Following the rationality of EUT's independence and cancellation principles, a choice between two alternatives should depend only on how those two alternatives differ and not on any factor that is the same for both alternatives, but human decision-makers typically violate this assumption. As another example, Ellsberg (1961) presented what became famous as the Ellsberg paradox - individuals systematically show a preference for well-defined probabilities and violate EUT's cancellation principle.

Another violation of EUT - a lack of transitivity, has been proposed by Tversky (1969). In a hypothetical (abstract) gamble experiment, Tversky found an effect violating the EUT's *transitivity* axiom. In Tversky's intransitivity phenomenon, when two alternatives have very close (similar) probability of winning, for example:

A. 7/24 chance of winning £5 (EV = £1.46)

B. 8/24 chance of winning £4.75 (EV = £1.58),

participants choose the option with the higher payoff (£5) of winning. In contrast, when the difference in probabilities is large, for example:

A. 7/24 chance of winning £5 (EV = £1.46)

or

B. 11/24 chance of winning £4 (EV = £1.83),

participants chose the option with high probability (11/24) of winning. The intransitivity phenomenon has been an important direction in decision-making science and has high applicability to many hypothetical (descriptive) decision-making applications involving probabilities and values.

Another well-known phenomenon, often interpreted as a failure of the *invariance* principle (stating that preferences over prospects are independent of the method used to elicit them), is the preference reversal phenomenon, observed first by Lichtenstein and Slovic (1971). Reversals of preferences (see also Grether & Plott, 1979) are observed when a so-called \$-bet (offering a high money prize with low probability) is assigned a higher reservation price than a P-bet (offering a lower money prize, but with a higher probability), but is subsequently not chosen in a direct choice between the two. The preference reversal phenomenon is an example of growing evidence in recent years that very minor changes in the presentation or *framing* of hypothetical prospects can have marked impacts upon the choices of decision-makers.

A well-known example of a framing effect, violating the invariance principle, was first reported by Tversky and Kahneman (1981), in which two groups of participants were presented with an Asian-disease story. The choice was between two pairs of medical policy options, which are probabilistically equivalent - one with a certain outcome and one with a risky outcome, having higher potential gain. The only difference was that the information in one experimental condition was presented in terms of lives saved while in the other experimental condition the information was presented in terms of lives lost. Tversky and Kahneman found a strong effect demonstrating the violation of EUT's invariance principle: 72% of participants preferred the first policy when it was described as lives saved, while only 22% of participants preferred this option when it was described in terms of lives lost.

Economic theory traditionally assumes that economic agents (decision-makers) have the ability to use information in the real world and to make rational decisions under risk, without being influenced by their risk own preferences and the domain of the decision problem (i.e., loss or gain). On the other hand, psychologists and social scientists have criticized the use of such strong assumptions, arguing for the existence of individual and contextual specificity. Psychologists became interested in EUT soon after its publication. As an alternative to the maximization hypothesis, Simon (1955, 1956) introduced the principle of *bounded rationality*, stating that cognitive limitations of decision-makers force them to construct a simplified model of the world in order to cope with it. EUT is concerned with probabilities, payoffs and the combination of these factors - people's expectations for a particular outcome (EU). The problem of comparing the worth of one consequence with the worth of another

consequence is solved directly by translating both into a common scale of utility. Simon's concept assumes that the goal of the decision-makers is to reach a *satisficing* rather than a maximizing decision, that is they set an aspiration level which, if achieved, they will be happy enough with; otherwise, they will try to change either their aspiration level or their decision. Decision-makers are constrained by limitations of perception and memory, and bounded rationality forces them to proceed by trial and error, modifying plans that do not ensure *satisficing* outcomes and maintaining those that do until they fail.

Normative decision-making theory assumes that a rational decision-maker maximizes EU for gains and minimizes EU for losses. Decision theory provides a model, based on the maximization (minimization) of EU, which serves as a normative or rational basis for making decisions. This view has been adapted in the light of criticism of EUT, for example, Simon's (1955, 1956) concept of bounded rationality. Such criticisms have led to the development of Prospect Theory, a descriptive model of decision-making under uncertainty proposed by Daniel Kahneman and Amos Tversky (1979). This, in turn, was extended into Cumulative Prospect Theory (Tversky & Kahneman, 1992; Tversky & Fox, 1995).

## **3. DESCRIPTIVE THEORY OF DECISION-MAKING** UNDER RISK AND UNCERTAINTY

#### 3.1 A Psychological Alternative to the Normative Approach

Kahneman and Tversky (1979) proposed Prospect Theory (PT) as a psychological alternative to the dominant EUT. In this theory, based on hypothetical decision problems (experiments with simple hypothetical gambles), choice is modeled as a two-phase process - editing followed by evaluation. In the first phase, prospects are *edited*, using a preliminary analysis of the offered prospects (hypothetical gambles), which results in a simple representation of these prospects. For example, a compound gamble prospect [£300, 30%; £300, 30%] can be reduced to [£300, 60%], using Kahneman and Tversky's *combination* operation.

The editing phase contains six specific operations (principles) that transform the outcomes and probabilities associated with the offered prospects (Kahneman & Tversky, 1979). According to *coding* operation, people normally perceived outcomes from hypothetical prospects as gains or losses rather than as final states of wealth or welfare. Gains and losses are defined relative to some neutral reference point (usually corresponding to the current asset position). Coding may be affected by the formulation of the offered prospects. In the *combination* operation, prospects are simplified by combining probabilities associated with identical outcomes (see above). The *segregation* operation states that some prospects contain a riskless component that is segregated from the risky component in the editing phase. The prospect [£100, 80%; £200, 20%] is naturally decompressed into a sure gain of £200 and a risky prospect [£100, 80%]. Another 'loss' example for a segregation is the prospect [-£400, 40%; -£100, 60%], which might be seen to consist of a sure loss of £100 and of the prospect [-£300, 40%]. The *cancellation* operation states that the components that are shared by the offered prospects are discarded. Participants may ignore the first part of a sequential game. For example, [£200, 20%; £150, 50%; -£100, 30%] and [£200, 20%; £100, 50%; -£50, £30]

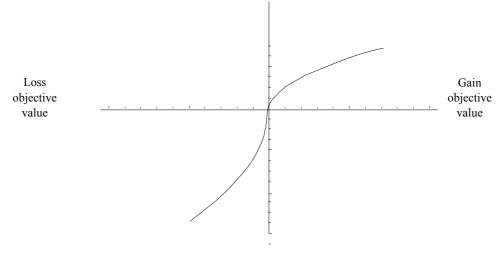
can be reduced by cancellation to a choice between [£150, 50%; -£100, 30%] and [£100, 50%; -£50, 30%]. The *simplification* operation supposes a simplification of prospects by rounding probabilities and outcomes. For example, the prospect [£101, 49%] is likely to be perceived as [£100, 50%]. The *detection of dominance* operation, involves scanning of offered prospects to detect a dominated alternative, which might be rejected without further evaluation.

In the evaluation phase, decision-makers evaluate each of the edited prospects and choose the prospect with the highest value. In the first version of PT (Kahneman & Tversky, 1979), the overall value of an edited prospect is expressed in terms of two functions - the value function and probability-weighting function. The evaluation of positive and negative prospects follows a different rule. In this phase, choices among edited prospects are determined by a reference function (value function), in which outcomes are interpreted as gains and losses relative to a reference point (e.g., status quo). The reason for treating consequences in this way is that it allows gains and losses to be evaluated differently. The main postulate of PT states that losses loom larger than corresponding gains. PT also differs from EUT in the way it deals with the probabilities attached to particular outcomes. EUT assumes that a decision-maker evaluates, for example, a 50% chance of winning as exactly as 50% of winning, whereas PT treats human preferences as a function of decision (probability) *weights*, assuming that these weights do not always correspond to the real level of probability - people tend to overweight small probabilities and underweight moderate and high probabilities (Kahneman & Tversky).

Another claim made by PT is in agreement with the extant psychophysical approach in psychology (Helson, 1964; Garner, 1954; Medin & Schaffer, 1978; Nosofsky, 1986), in particular that humans' perceptual apparatus is attuned to the evaluation of changes or differences rather than to the evaluation of absolute magnitudes - people's judgments are affected by the set of stimuli in the task. The response to single psychophysical stimuli differing in for example brightness, loudness or temperature depends on immediate or previous experience with stimuli in the task. Kahneman and Tversky (1979) assumed that the same principle applies to non-sensory attributes such as health, prestige and wealth.

The basic principle of EUT is that decisions under risk are made to maximize people's EU for gains and minimize EU for losses (von Neumann and Morgenstern, 1947) and that equivalent formulations of a choice problem should give rise to the same preference order. Kahneman and Tversky's (1979, 1984) and Tversky and Kahneman's (1986, 1991, 1992) concept of loss aversion suggests an alternative functional form, different for the domains of gain and loss - a value function that is steeper (convex) for losses (the curve lying below the horizontal axis) than for gains (see Figure 2), implying that people feel losses more than they do gains of equivalent value. Contrary to the normative assumption, it was concluded that a framing effect (in terms of gains and losses) leads to systematically different preferences with people's decisions (Tversky & Kahneman, 1986).

However, recent research has shown that loss aversion is not a stable phenomenon and can vary depending on the decision task (abstract gamble or real-world protective task, the latter resulting in an absence of loss aversion) (Kusev et al., 2006, 2007), demographic differences (less loss aversion with a higher level of education) (Johnson, Gachter & Herrmann, 2006) or experience with the attributes probability and value (experts are less loss-averse) (Camerer, Babcock, Loewenstein & Thaler, 1997; List, 2003, 2004).



Subjective value

Figure 2. S-shaped value function for hypothetical decision-making tasks (Kahneman & Tversky, 1979).

#### 3.2 Probability Prospects - Cumulative Prospect Theory

Anticipated-Utility theory (AU) and its Rank-Dependent Expected-Utility model (RDEU) were originally proposed by Quiggin (1982) and later extended by Yaari (1987) to explain the Allais paradox. RDEU is typically applied to problems that have well-defined probabilities and are formulated in terms of hypothetical lotteries. AU is a generalization of EUT, where RDEU incorporates transformations of both a prospect's probabilities as well as its outcomes. One characteristic of RDEU is that it uses the transformation of the cumulative distribution of a prospect's probabilities (where the prospect's outcomes are ordered by attractiveness) rather than the direct transformation of probabilities that is hypothesized in PT. RDEU offered a new representation (called rank-dependent or cumulative functional) that transforms cumulative rather than individual subjective probabilities. This use of the cumulative distribution makes the weighting of an outcome dependent on its rank, hence the name RDEU. Quiggin (1982) proposed an S-shaped function for humans' perceptions of probabilities that was later generalized by Tversky and Kahneman's (1992) Cumulative Prospect Theory (CPT).

Tversky and Kahneman (1992) proposed CPT as an extension of PT. In this extended version of PT, the problem of how people estimate probabilities (subjective probabilities) is central. CPT uses the same basic principles as original PT, a value function, defined over gains and losses, and a probability-weighting function that captures people's risk preferences. The major technical innovation is the use of the cumulative functional rank-dependent form

to extend PT to uncertain and risky prospects with many hypothetical probabilities (ranks) and different numbers for outcomes.

The new theory incorporates the basic principles of PT and can explain descriptive violations of EUT: (a) *framing effects* in terms of gains and losses - EUT assumes description invariance, that is equivalent formulations of a choice problem should give rise to the same preference order, but human decision-makers violate this assumption; (b) *nonlinear preferences* - according to EUT's maximization principle, the utility of a prospect under conditions of risk is linear in outcome probabilities, but Allais (1953) successfully challenged this principle; (c) *source dependence* - Ellsberg (1961) observed that individuals systematically show a preference for well-defined probabilities, in contrast to EUT's rationality; and (d) *risk-seeking behavior* - people prefer a prospect with a small probability of winning a large prize over another prospect with larger probability and smaller prize. People prefer risk-seeking in choice between a sure loss and a substantial probability of a larger loss. However, risk-aversion is generally assumed in economic analyses of decision under uncertainty based on EUT (Tversky & Kahneman, 1992). Kusev et al. (2006, 2007) demonstrate that such a difference in risk preferences could be a result of the decision task given to the participants.

Tversky and Kahneman (1992) presented a comprehensive, empirical test of CPT consisting of two sessions, where participants' cash equivalents (certainty equivalents - CE) for a number of hypothetical gambles were calculated as a midpoint between the lowest accepted and highest rejected value in the second session. The prospects involved the domains of gain and loss, with different probabilities and outcomes. The vast majority of participants exhibited the four-fold pattern of risk attitudes, in particular risk-seeking preferences for low-probability ( $p \le 10\%$ ) gains and high-probability ( $p \ge 50\%$ ) losses and risk-averse preferences for high-probability ( $p \ge 50\%$ ) gains and low-probability ( $p \le 10\%$ ) losses (see Table 1).

A parametric regression analysis of the CEs produced a value function of the form of Figure 2 and a probability-weighting function of the form of Figure 3. As in the original prediction made by PT, Kahneman and Tversky's (1992) result demonstrates that the value function is concave for gains (above the reference point of 0) and convex for losses (below the reference point of 0) (see Figure 2). The function is also steeper for losses than for gains, implying the principle of loss aversion according to which losses loom larger then corresponding gains.

The one-parameter model proposed by Tversky and Kahneman (1992) to described the probability-weighting function has several specific features: (a) it has only one parameter irrespective of decision task (hypothetical or real-world), but this feature was not confirmed by experimental evidence from recent research by Gonzalez and Wu (1999); (b) it encompasses weighting functions with both convex and concave regions; and (c) it provides a reasonably good approximation to both aggregate and individual data for probabilities in the range between .05 and .95, demonstrating people's tendency to overweight small probabilities and underweight the large ones. Consequently, people are relatively insensitive to probability differences in the middle of the range. The pattern of probability estimates demonstrates that the probability weights for hypothetical gambles with gains and losses are relatively similar, although the probability functional form for gains is slightly more curved for gains (see Figure 3) (Abdellaoui, 2000; Camerer & Ho, 1994; Prelec, 1998; Stewart et al., 2006; Tversky & Kahneman, 1992; Tversky & Fox, 1995).

	Hypothetical, gain		Hypothetic	Hypothetical, loss	
	p ≤ .1	$p \ge .5$	$p \le .1$	$p \ge .5$	
Risk-seeking	78% <sup>a</sup>	10%	20%	87% <sup>a</sup>	
Risk-averse	10%	88% <sup>a</sup>	80% <sup>a</sup>	6%	

Table 1. Risk-Seeking and Risk-Averse Choices for Hypothetical Decision-Making
Tasks (Tversky & Kahneman, 1992)

Note. Percentages are mean values.

<sup>a</sup> Values corresponding to the fourfold pattern of risk-seeking (Tversky & Kahneman, 1992).

Many other studies, using a variety of methodologies, agree or find inverse S-shape functions as shown in Figure 3 (e.g., Abdellaoui, 2000; Camerer & Ho, 1994; Tversky & Fox, 1995). Abdellaoui (2000) agrees with the introduction of probability-weighting functions in CPT, but also notes the observed tendency that participants treat probabilities differently when passing from gains to losses and vice versa (see Figure 3). Abdellaoui suggests that CPT's experimental findings can be seen as a confirmation that the traditional utility elicitation methods could produce distorted utility functions. Camerer and Ho (1994) argue for a model that is nonlinear in probability in order to capture people's risk preferences. Their results are consistent with both CPT and RDEU models, assuming an inverse S-shaped weighting function. Other studies with hypothetical prospects (gambles) have produced similar parameter estimates as those found by Tversky and Kahneman (1992), even though these exercises varied considerably in terms of the data used and estimation techniques (see Abdellaoui, 2000; Camerer & Ho, 1994; Tversky & Fox, 1995; Wu & Gonzalez, 1996).

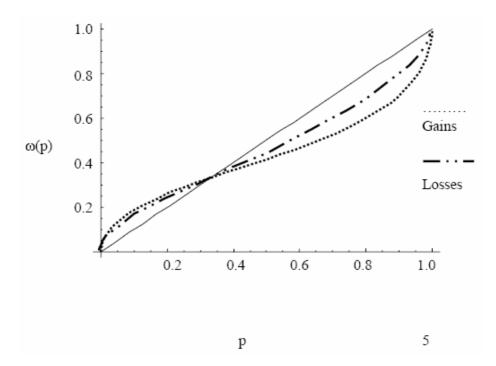


Figure 3. Inverse S-shaped probability-weighting function for hypothetical decision-making tasks (Tversky & Kahneman, 1992). *Note*.  $\omega(p)$ : humans' estimate of probability.

In accordance with humans' adeptness at evaluating changes or differences, another important feature of CPT is that the value function and the probability-weighting function exhibit *diminishing sensitivity* with distance from a reference point (Tversky & Fox, 1995). The characteristic of diminishing sensitivity gives rise to an S-shaped value function that is concave for gains and convex for losses. For probability, there are two natural reference points - 0% (impossible) and 100% (certain), which correspond to the endpoints of the scale. Diminishing sensitivity implies that increasing the probability of winning a prize by 10% has more impact when it changes the probability of winning from 90%-100% or from 0%-10% than when it changes the probability from 40%-50% or from 50%-60% (see Figure 3). However, the stability of the principle of diminishing sensitivity is questionable (Kusev et al., 2006, 2007); in particular, this phenomenon could vary depending on the given decision task.

Despite CPT's general applicability to abstract gambles, the model is unlikely to be accurate in specific contexts, in particular the probability-weighting might be sensitive to the formulation of the prospects and types of prospect. In this chapter, we suggest that such sensitivity is due to the format of the decision prospects, for example hypothetical or given protective real-world task. Although existing studies of decision-making, using hypothetical (abstract) gambles, corroborate the need for a distinction between normative and descriptive accounts, evidence presented by recent research (e.g., Gonzalez & Wu, 1999) suggests they cannot provide all the answers that we need to explain people's behavior in the face of risk. As a consequence, various recent theoretical developments have specifically addressed decision-making with real-world prospects under risk.

#### 4. THEORETICAL DIRECTIONS IN PROTECTIVE DECISION-MAKING

Why are people attracted to *protective* options such as insurance contracts, savings accounts, life- and health-assurance? In this section, we investigate what might affect people's probability-weighting with respect to described protective decision prospects. People's reactions in the face of risk imply that they may judge and weight the probability of risky events in characteristic ways that deviate from existing theory. We may expect a discrepant tendency concerning the probability-weighting process with precautionary behavior (Baron, Gurmankin & Kunreuther, 2007; Kunreuther, 2001; Slovic, Fischhoff & Lichtenstein, 1987) compared to decisions with abstract gambles (still the main context for investigating economic behavior).

We specifically address the applicability of PT and CPT, proposing a protective-decisionmaking approach. People can reduce risks by engaging in protective behaviors such as washing their hands before eating, wearing a seat belt, taking health-screening tests, avoiding second-hand smoke, limiting alcohol consumption and installing a smoke detector. Governments and other organizations often set up programs to try to encourage people to undertake such protective behaviors (Baron et al., 2007). The research on precautionary behavior is also focused on what factors lead people to adopt a particular precautionary behavior and what interventions are likely to increase its adoption. For example low-risk, highly devastating events (e.g., natural disasters) could stimulate people's precautionary behavior (Kunreuther, 2001). In this section, we review two approaches that contribute towards the explanation and prediction of human decision-making in both hypothetical and precautionary decision scenarios.

4.1 The Role of Experience in People's Judgments under Risk: Implications for Protective Decision-Making

In their everyday decisions (e.g., regarding insurance decisions, pension plans, savings) people may exaggerate the risk of described risky options, when protecting themselves against the risk of real-world negative outcomes. Instances of some protectable risks are encountered in everyday life (via TV, newspapers, advertisements and individual experience) disproportionately frequently and might affect the probability-weighting process with given protective (real-world) scenarios. We could suppose that there are differences in the shape (modality) of the probability-weighting function in different domains of judgment - hypothetical (abstract) gambles and described real-world protective prospects. This notion is supported by evidence from research in subjective probability - in particular, that different descriptions of the same event can prompt different subjective probabilities (e.g., Tversky & Koehler, 1994) - and insurance-decision studies (e.g., Viscusi, 1995; Wakker, 2003; Wakker, Thaler & Tversky, 1997).

Even today, many economists interpret people's insurance preferences in terms of and within the framework of EUT. The descriptive psychological alternatives (PT and CPT) revealed the reasons why we should question its descriptive adequacy. For example, Murray (1972) and Neter, Williams and Whitmore (1968) found that utility functions scaled individually for each participant failed to predict their lottery and insurance preferences. It has therefore been suggested that the application of a utility and a disutility function as proposed by EUT for decision-making is not successful.

Behavior inconsistent with EUT has been confirmed in many problems, including insurance choices. For example, EUT cannot account for insurance-purchasing behavior if rates are actuarially unfair - a very expensive product with extremely low risk-probability (Hogarth & Kunreuther, 1992; Hershey, Kunreuther & Schoemaker, 1982). Decisions about insurance protection (e.g., natural disasters) are usually based on a lack of knowledge about relevant probabilities and in these situations people are often ambiguity-averse - they prefer known probability distributions over uncertain ones. Actuaries suggest higher warranty prices for ambiguous probabilities than for well-specified probabilities and underwriters set higher insurance premiums for ambiguous probabilities and losses than for well-specified probabilities and losses (Hogarth & Kunreuther, 1992; Kunreuther, 2001). However, Kusev et al. (2006, 2007) present experimental evidence showing that even well-defined (fairly presented) high and low real-world probabilities can be exaggerated.

Investigating insurance prospects, Camerer and Kunreuther (1989) assumed that most people have difficulties in evaluating small probabilities, arguing that decision-makers often ignore the information regarding small probabilities (underweighting). In contrast, Viscusi and Chesson (1999) argue that people are *averse* to ambiguous probabilities, demonstrating that preference for a known probability is most prevalent for low-probability losses and high-probability gains. In addition, Johnson, Hershey, Meszaros and Kunreuther (1992) found that manipulation of descriptions in an insurance-decision task leads decision-makers to violate EUT's basic principles regarding probabilities and values. Participants exhibit distortions of

risk; in particular, the way that an insurance premium is framed can determine attractiveness of the insurance, in agreement with PT's *framing* effect - framing the prospects as gains or losses affects people's choices.

The theories of judgment and decision-making, ranging from normative frameworks that assume rational behavior to descriptive psychological concepts, differ in their approach to major issues. The descriptive psychological alternatives of EUT demonstrate that, breaching rational agents' rules, people usually treat probabilities non-linearly, overweighting low and underweighting high probability (Abdellaoui, 2000; Gonzalez & Wu, 1999; Prelec, 1998; Tversky & Kahneman, 1992; Tversky & Wakker, 1995). In EUT, every rational decision-maker should be able to trade off the value of all the possible outcomes by the likelihood of obtaining them (Ayton, 2005), highlighting that we all well deal with probabilities and monetary outcomes. However, there appears to be no evidence that normative EUT and descriptive theories of decision-making (tested with abstract gambles) - such as PT and CPT - can account for people's protective decision-making. The meaning of probability for decisions under risk and uncertainty has been the subject of numerous studies and the issues are still far from settled.

Research on decision-making (e.g., Simon, 1956; Ellsberg, 1961; Allais, 1953; PT and CPT) has suggested that people violate the principles of rational choice (EUT) and these violations are different in the domains of gain and loss. However, Baron et al.'s (2007) comparative approach demonstrates that real-world behavior is more complex and in some cases could be rational. The authors suggest that people's precautionary behavior might be rational (i.e., consistent with EUT) when considered in isolation, but not when considered in the context of other precautionary behaviors - people are less willing to undertake a precautionary behavior when they are considering several precautionary behaviors at the same time. People might have 'precautionary behavior budgets' - the extent to which people engage in a particular precautionary behavior decreases as the number of precautionary behaviors that they consider increases.

Further evidence of the complexity of people's real-world judgments has been presented by experience-based decision-making models, which recently gained popularity as psychological models of decision-making (Dougherty, Gettys & Ogden, 1999; Frisch & Jones, 1993; Fiedler, 2000; Gilboa & Schmeidler, 2001; Hertwig et al., 2004; Stewart et al., 2006; Tversky & Koehler, 1994). Kunreuther (2001) offered the notion of *fear*, based on individual experience with protective decision-making, as a powerful emotional factor; for example, people's fear of crime is what makes people more willing to spend on various precautionary measures. However, EUT does not predict a role of *affect* and *emotions* in people's precautionary behavior.

Case-Based Decision Theory - CBDT (Gilboa & Schmeidler, 2001) makes very similar predictions as Kunreuther, that is decisions under uncertainty are made by analogies to previously-encountered problems. The theory postulates a similarity function over decision problems and a utility function on outcomes, such that acts are evaluated by a similarity-weighted sum of the utility they yielded in past cases in which they were chosen. According to CBDT, decision-makers are not rational as EUT predicts - they count on their experience rather than attempt to figure out what the outcomes of available choices will be.

In agreement with the *experienced* decision-making orientation, Viscusi and Chesson (1999) assumed that there are in fact, few risks where probabilities are known with precision and decision-makers very often make risky decisions, in some instances after receiving highly

divergent risk information. The authors proposed that people exhibit fear in response to the ambiguity associated with small probabilities and loss and *hope* in response to large probabilities. Ignoring the *accessibility* paradigm (see Section 4.2 below), Kunreuther (2001) has suggested the notion of *fear-and-prudence* decisions, where only past individual experience with anxiety or fear have an important influence on decisions of whether or not to invest in protective measures.

The experience-based view suggests that estimates of probabilities are made by retrieving exemplars from memory and that frequently-occurring probabilities will be more strongly represented in memory. People's estimation of the distribution of probabilities is dependent on their experience with small or large probabilities (Stewart et al., 2006). According to Stewart et al.'s theory - Decision by Sampling (DbS), the attributes that constitute the decision sample come either from memory or from the immediate context of the decision and the distribution of values in memory is assumed to reflect the distribution of attribute values in the world. DbS accounts for a convex utility function (in accordance with existing psychological models) - losses loom larger than gains and decision weights are distorted (the overweighting of small probabilities, and the underweighting of large probabilities).

Fiedler's (2000) theoretical framework, suggests that the sampling process may draw on the external world or on internal memories. Because people lack the meta-cognitive ability to understand and control for sampling constraints, sampling biases carry over to subsequent judgments. Human behavior appears to be affected by a reliance on relatively small samples of information and overweighting of recently sampled information and there is evidence that people distinguish between decisions based on experience and decisions based on description (Hertwig et al., 2004). Recently, Hertwig, Pachur and Kurzenhäuser (2005) suggested two decision mechanisms - availability by recall and regressed frequency, where the former mechanism assumes that participants' risky judgment is a function of the number of cases recalled from participants' social circles and the latter offers the assumption that people keep track of the frequency of occurrences of risk. The two mechanisms were formally stated as follows:

Choice proportion<sub>*Risk a*</sub> = 
$$\frac{\sum recalled ins \tan ces_{Risk a}}{\sum recalled ins \tan ces_{Risk a}} + \sum recalled ins \tan ces_{Risk b}}$$

(recall)

and

Choice proportion<sub>Risk a</sub> = 
$$\frac{\sum occurrences_{Risk a}}{\sum occurrences_{Risk a}} + \sum occurrences_{Risk b}}$$

(regressed frequency).

In the framework of illness scripts, proposed as a pertinent type of knowledge structure operating in medical contexts (e.g., Charlin, Tardiff, & Boshuizen, 2000) - including decision-making (van Schaik, Flynn, van Wersch, Douglass & Cann, 2005), three types of

knowledge component are distinguished. One of these components ('enabling conditions' both medical and non-medical contextual patient-related factors that influence the probability of disease) develops as a function of experience and affects decision-making over and above and in interaction with 'objective' knowledge components ('consequences') (van Schaik et al., 2005), again highlighting the role of experience-based knowledge in decision-making.

#### 4.2 The Effect of Accessibility on Decision-Making under Risk

In this chapter, we suggest that people's past experience with the frequency of risks has an important influence on decisions whether or not to invest in protective measures. According to our assumption, in judging prospects in protective-decision scenarios, people's risk preferences may be affected by their subjective experience and the accessibility of frequencies of hazardous events (Higgins, 1996; Kahneman, 2003; Koriat, 1993, 1995; Koriat & Levy-Sadot, 2001; Schwarz, 1998; Tulving & Pearlstone, 1966), that is not all available observations of risks are equally accessible in memory.

It seems that the probability associated with highly accessible features or events tends to be exaggerated (see e.g., Kusev et al., 2006, 2007), whereas events of low accessibility (e.g., gambles and inaccessible real-world events) in memory are largely ignored. People judge highly accessible hazards in memory as more likely than less accessible events. Indeed, we do not have memory for any particular experienced frequencies with sampling of abstract gambles (or the capability of retrieving examples of abstract gambles from our memory, even though abstract gambles have typically been used in decision-making research), unrealistic gain insurance scenarios or low-accessibility risks.

The concept of *accessibility* originates from research on memory (Koriat, 1993; Tulving & Pearlstone, 1966) and social cognition (Higgins, 1996; Schwarz, 1998). According to Tulving and Pearlstone's theoretical framework, it is important to draw a distinction between what information or what traces are available in memory storage and what are accessible. Based on one experiment (N = 948), they proposed that availability refers to whether information is stored in memory, whereas *accessibility* is the ease of retrieving information that is available. This distinction parallels that between retention and recall or the distinction between trace storage and trace utilization.

More recent developments in *accessibility* (Koriat, 1993, 1995; Koriat & Levy-Sadot, 2001) report that when participants fail to recall an answer, their judgments of feeling-of-knowing are based on the amount and intensity of the partial information accessed in the course of the search for the target. The assumption is that even if the retrieval attempt is unsuccessful, it may generate a variety of partial clues and activations, such as fragments of the target (e.g., semantic and episodic attributes). These partial clues may induce the subjective feeling that the target is stored in memory although there is no direct access to the accuracy of the partial clues that come to mind.

In social-cognition research, *accessibility* is defined as the activation potential of available knowledge (Higgins, 1996; Schwarz, 1998). This definition is useful because it separates the concepts of *accessibility* and availability, which are often confused. However, in distinguishing the availability heuristic (Tversky & Kahneman, 1973) from accessibility, Kahneman (2003) described *accessibility* as a continuum that has characteristics of perception and of the intuitive system; specifically, *accessibility* might be viewed as an intuitive

judgment or preference where particular solutions come to mind. There appears to be a lack of evidence that the most accessible features are also the most relevant to making a good decision.

It is difficult for humans to make judgments about given real-world options and their possible outcomes. The complexity of the real-world environment, our risk preferences and individual experience make the decision process an arduous task. In contemplating most real-world risks, people suffer from a lack of knowledge about the probabilities of hazardous events (e.g., natural disaster, health and safety risk). As a consequence, most of our decisions will be based on accessibility of the frequencies of real-world observations - often via news reports on TV, newspapers and advertisements. Recent results (Kusev et al., 2007) show that risk exaggeration is caused by experienced frequencies and their accessibility in memory. We could explain the exaggeration of the risk of hazardous events by assuming that instances of some protectable risks are encountered in everyday life disproportionately frequently. The findings indicate that people estimate probability differently and are willing to drastically exaggerate given real-world high-frequency risks even when the probabilities are known and fairly presented.

The evidence presented by Kusev et al. (2006, 2007) demonstrates that individuals magnify the probability of given protective options compared to probabilities in hypothetical gamble scenarios in the domain of loss. In contrast, the well-known fourfold pattern of risk attitudes (Tversky & Kahneman, 1992; Tversky & Fox, 1995) suggests predominantly risk-seeking preferences for losses with medium-sized and large probabilities (see Table 1). People's risk preferences in hypothetical decision tasks (gambles), when using a repeated measures design - including a mix of trials with losses and trials with gains, are characterized by risk-seeking behavior for small ( $p \le .1$ ) probabilities of gains and risk-aversion for small ( $p \le .1$ ) probabilities of loss (see Table 1). (Tversky & Kahneman, 1992; Tversky & Fox, 1995). However, people's protective choice behavior (Kusev et al., 2006, 2007), as represented in Figure 4 and Table 2, highly overweights small and moderate probabilities and underweights or almost behaves neutrally in response to high probabilities; the pattern of risk-seeking preferences in protective decision-making (see Table 2) is different from that in abstract gambles (see Table 1).

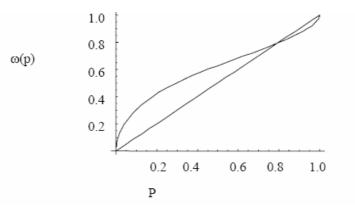


Figure 4. Probability-weighting function for protective decision-making tasks (Kusev et al., 2006, 2007). *Note*.  $\omega(p)$ : humans' estimate of probability.

	p =.01	$.01$	$p \ge .5$	p =.01	$.01$	p≥.5
	Hypothetical, gain			Hypothetical, loss		
Risk-seeking	78% <sup>a</sup>	57%	11%	13%	30%	88% <sup>a</sup>
Risk-averse	22%	43%	89% <sup>a</sup>	87% <sup>a</sup>	70%	12%
	Protective, gat	e, loss				
Risk-seeking	76%	56%	7%	0%	12%	45%
Risk-averse	24%	44%	93%	100%	88%	55%

# Table 2. Risk-Seeking and Risk-Averse Choices for Hypothetical and Protective Decision-Making Tasks (Kusev et al., 2006, 2007)

*Note.* Percentages are mean values. The pattern of results shows similarity among tasks in risk-seeking with gain, but inconsistency among tasks in risk-seeking preferences with loss.

<sup>a</sup> Values corresponding to the fourfold pattern of risk-seeking (Tversky & Kahneman, 1992).

In the past few decades, the question how people make decisions and deal with realworld choice options has become of fundamental significance to cognitive psychologists. Although it seems that the existing theories, models and concepts are adequate and well established, recent debate in decision-making demonstrates that we are far away from any general behavioral conclusions regarding core decision-making dimensions such as value or utility, probability, loss and gain.

People's incapacity to simultaneously deal with complex multidimensional tasks and the failure of existing models to explain the complexity of real-world decisions in terms of value, probability, loss and gamble motivated the notions presented in this chapter. Kusev et al. (2006, 2007) identified two independent probability-related psychological constructs: *probability discriminability* and *protective willingness*, which together affect probability-weighting for protective decision-making, resulting in significant overweighting of probability. Descriptive and normative models fail to explain the probability-weighting function for this type decision-making (see Figure 4), in which small and moderate probabilities are highly overweighted and higher probabilities are almost accurately estimated.

#### **5.** CONCLUSION

The normative approach to the study of decision-making (EUT, SEUT) does not distinguish between rational and human decision-making. The descriptive psychological approach (PT, CPT) does account for various violations of normative theory by human decision-making, but does not distinguish between decision-making in hypothetical and given real-world tasks.

Various theoretical approaches contribute to an explanation of decision-making in protective decision-making. In contrast to hypothetical decision-making, protective decision-making appears to be influenced by emotions, conceptualized for example as fear and prudence (Kunreuther, 2001), and hope and fear (Viscusi & Chesson, 1999). Furthermore, more than one decision-making parameter affects humans' risk preferences, in particular protective willingness is involved in protective decision-making (Kusev et al., 2006, 2007). The exaggeration of the risk of hazardous events can be explained by assuming that instances

of some protectable risks are encountered in everyday life disproportionately frequently. In deciding about real-world risks, people suffer from a lack of knowledge about the probabilities of hazardous events, indicating that risk exaggeration is caused by experienced frequencies and their accessibility in memory.

Future research should investigate the range of situations under which risk exaggeration occurs and the influence of experience and accessibility in different tasks. It is important therefore that any future research expands the scope of enquiry to include data from populations with more varied demographic characteristics, using a range of decision environments (e.g., real-world applications and computer-controlled individually administered experiments).

#### REFERENCES

- Abdellaoui, M. (2000). Parameter free elicitation of utilities and probability weighting functions. *Management Science*, 46, 1497-1512.
- Allais, M. (1953). Le comportment de l'homme rational devant le risqué: Critique des postulats et axioms de l'ecole americaine. *Econometrica*, *21*, 503-546.
- Ayton, P. (2005). Judgment and decision making. In N. Braisby & A. Gellatly (Eds.), *Cognitive psychology* (pp. 382-413). New York: Oxford University Press.
- Baron, J., Gurmankin, A. & Kunreuther, H. (2007). The comparative approach to protective behavior. *Journal of Behavioral Decision Making* (under review).
- Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22, 23-36. (Original work published 1738)
- Camerer, C., Babcock, L., Loewenstein, G. & R. Thaler. R. (1997). Labor supply of New York City cabdrivers: One day at a time. *Quarterly Journal of Economics, 112*, 407-41.
- Camerer, C. F. & Ho, T. H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, *8*, 167-196.
- Camerer, C.F. & Kunreuther, H.C. (1989). Experimental markets for insurance. *Journal of Risk and Uncertainty*, *2*, 265-300.
- Charlin, B., Tardiff, J., & Boshuizen, H. (2000). Scripts and medical diagnostic knowledge: Theory and applications for clinical reasoning instruction and research. *Academic Medicine*, 75, 182–190.
- Dougherty, M. R. P., Gettys, C. F. & Ogden, E. E. (1999). MINERVA-DM: A memory process model for judgments of likelihood. *Psychological Review*, 106, 180-209.
- Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75, 643-79.
- Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling theory of judgment biases. *Psychological Review, 107*, 659-676.
- Frisch, D. (1993). Reasons for framing effects. Organizational Behavior and Human Decision Processes, 54, 399-429.
- Frisch, D. & Jones, S.K. (1993). Assessing the accuracy of decisions. *Theory and Psychology*, *3*, 115-135.
- Garner, W. R. (1954). Context effects and the validity of loudness scales. *Journal of Experimental Psychology*, 48, 218-224.

- Gilboa, I. & Schmeidler, D. (2001). *A theory of case-based decisions*. New York, Cambridge University Press.
- Gonzalez, R. & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38*, 129-166.
- Grether, D.M. & Plott, C.R. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, 69, 623-638.
- Helson, H. (1964). Adaptation-level theory. New York: Harper & Row.
- Hershey, J.C., Kunreuther, H.C. & Schoemaker, P. (1982). Sources of bias in assessment procedures for utility functions. *Management science*, 28, 936-954.
- Hertwig, R., Barron, G., Weber, E. & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*, 534-539.
- Hertwig, R., Pachur, T., & Kurzenhäuser, S. (2005). Judgments of risk frequencies: Test of possible cognitive mechanisms. *Journal of Experimental Psychology: Learning, Memory,* and Cognition, 4, 621-642.
- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133-168). New York: Guilford.
- Hogarth, R. M. & H. Kunreuther. (1989). "Risk, Ambiguity and Insurance," Journal of Risk and Uncertainty 2, 5-35.
- Johnson, E., Gachter, S. & Herrmann, A. (2006). Exploring the nature of loss aversion. Manuscript in preparation.
- Johnson, E. J., Hershey, J., Meszaros, J. & Kunreuther, H. (1992). Framing, probability distortions, and insurance decision. *Journal of Risk and Uncertainty*, 7, 35-51.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*, 697-720.
- Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.
- Kahneman, D. & Tversky, A. (1984). Choices, values, and frames. American Psychologist, 39, 341-350.
- Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, 100, 609-639.
- Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124*, 311-333.
- Koriat, A. & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feeling of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 1*, 34-53.
- Kunreuther, H. (2001). Protective decisions: Fear or prudence. In Hoch, S., & Kunreuther, H. (Eds.), *Wharton on Making Decisions*. New York: Wiley.
- Kusev, P., Ayton, P., van Schaik, P. & Chater, N. (2006). Taking precautions is not the same as choosing gambles: Prospect theory and use of probability in risky choices. *The 37<sup>th</sup> Annual Conference of the Society for Judgment and Decision Making*. Houston, Texas. November 17-20, 2006.
- Kusev, P., Ayton, P., van Schaik, P. & Chater, N. (2007). Exaggerated risk. The 21<sup>st</sup> Biannual Conference on Subjective Probability, Utility and Decision Making. Warsaw. August 19-23, 2007.

- Lichtenstein, S. & Slovic, P. (1971). Reversal of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, 89, 46-55.
- List, J. A. (2003). Does market experience eliminate market anomalies? *Quarterly Journal of Economics*, *118*, 41-71.
- List, J. A. (2004). Neoclassical theory versus prospect theory. Evidence from the marketplace. *Econometrica*, 72, 615-625.
- Medin, D. L. & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85, 207-238.
- Murray, M. (1972). Empirical utility functions and insurance consumption decisions. *Journal* of Risk and Uncertainty, 1, 31-41.
- Neter, J., Williams, C. A. & Whitmore, G. A. (1968). Comparison of independent and joint decision-making for two insurance decisions. *Journal of Risk and Uncertainty*, 1, 87-105.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115,* 39-57.
- Plous, S. (1993). The psychology of judgment and decision making. New York: McGraw-Hill.
- Prelec, D. (1998). The probability weighting function. Econometrica, 66, 497-527.
- Quiggin, J. (1982). A Theory of Anticipated Utility. *Journal of Economic Behavior and Organization 3*, 323-343.
- Savage, L. (1954). The foundations of statistics, New York, Wiley.
- van Schaik P., Flynn D., van Wersch A., Douglass, A. & Cann, P. (2005). The influence of illness script components and medical practice on medical decision-making. *Journal of Experimental Psychology: Applied*, 11, 187-199.
- Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review*, 2, 87-99.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*, 99-118.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129-138.
- Slovic, P. (1987). Perception of risk. Science, 236, 280-285.
- Slovic, P. (2000). The perception of risk. London: Earthscan publication.
- Slovic, P., Fischhoff, B. & Lichtenstein, S. (1987). Behavioral decision theory perspectives on protective behavior. In N.D. Weinstein (Ed.), *Taking care: Understanding and encouraging self-protective behavior* (pp. 14-41). New York: Cambridge University Press.
- Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53, 1-26.
- Tulving, E. & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381-391.
- Tversky, A. (1969). Intransitivity of preferences. Psychological Review, 76, 31-48.
- Tversky, A. & Fox, C. R. (1995). Weighting risk and uncertainty. *Psychological Review*, 102, 269-283.
- Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207-232.
- Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453-458.

- Tversky, A. & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal* of Business, 59, 251-278.
- Tversky, A. & Kahneman, D. (1991). Loss aversion in riskless choice: A reference dependent model. *Quarterly Journal of Economics*, 106, 1039-1061.
- Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 204-217.
- Tversky, A. & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, 101, 547-567.
- Tversky, A. & Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica*, 63, 1255-1280.
- Viscusi, K. & Chesson, H. (1999). Hopes and fears: The conflicting effects of risk ambiguity. *Theory and Decision*, *47*, 153-178.
- Viscusi, W. K. (1995). Government action, biases in risk perception, and insurance decisions. The *Geneva Papers on Risk and Insurance Theory*, 20, 93-110.
- von Neumann, J. & Morgenstern O. (1947). *Theory of games and economic behavior*. Princeton, N.J., Princeton University Press, 2<sup>nd</sup> Edition.
- Wakker, P. (2003). The data of Levy and Levy (2003) "Prospect theory: Much ado about nothing?" actually support prospect theory. *Management Science*, 49, 979-981.
- Wakker P.P., Thaler, R.H. & Tversky A. (1997). Probabilistic insurance. Journal of risk and uncertainty, 15, 7-28.
- Wu, G. & Gonzalez, R. (1996). Curvature of the Probability Weighting Function. Management Science, 42, 1676-1690.
- Yaari, M. (1987). The dual theory of choice under risk. Econometrica 55, 95-117.

Chapter 41

# INSTABILITY AND RELATIVITY OF PREFERENCES: HOW CONTEXT DETERMINES UTILITIES AND DECISIONS\*

# Ivo Vlaev<sup>†</sup>

Department of Psychology, University College London, 26 Bedford Way, London, WC1H 0AP United Kingdom

# ABSTRACT

The aim of this paper is to draw attention to what is arguably a very general and pervasive feature of human cognition that may have important implications for our understanding of human decision making and also for some aspects of economics. The major claim, defended here, is that when people judge the attributes of choice options (like utilities, payoffs, and probabilities), they are not able to represent the absolute magnitudes of these attributes; instead, they represent magnitudes ordinally---in relation to other magnitudes that they can sample from memory or from the current environment. Also, when people represent a magnitude, they can only do so on the basis of whether it is larger or smaller than other sampled magnitudes. Such sampling of knowledge from memory and transferring it to the current situation produces certain biases in judgment because stimuli are judged only relative to each other and therefore utility of an option is dependent on the other options that can be retrieved from memory. As a consequence, there may be no ability to represent cardinal scales, for any magnitude and judgments involving such magnitudes are determined by the context. The core evidence for this claim comes from recent research in psychophysics on the perception of the intensity of basic psychophysical magnitudes such as the brightness of a light or the loudness of a sound, and also from research on the effects of context on decision making under risk and uncertainty.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Economics, Business and Finance*, edited by Klaus P. Hofmann published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> E- mail: i.vlaev@ucl.ac.uk; Tel: +44 207 679 7585

### INTRODUCTION

Economic theory is often conveniently formulated to assume that economic actors have perfect information, a perfect grasp of their objectives, and the perfect ability to use that information to further their objectives. Markets comprising such agents, and games played by such agents, can be understood using particularly elegant techniques, and producing general and valuable theoretical results. It is common for social scientists outside economics to criticise the use of such strong assumptions as hopelessly indefensible. But such criticism is not, on its own, any more persuasive than any other criticism of scientific idealization. The analysis of the world routinely involves wild simplifying assumptions. Most prosaically, Newtonian celestial mechanics treats planets as point masses; the analysis of the path of a projectile may ignore friction. These idealizations seem rather harmless in comparison to those of the economist; they seem to involve ignoring factors that may seem, in some particular context, to be of rather limited significance; and in any case there is the hope that we may be able to add them in later, albeit at the cost of complicating our calculations.

In reality, though, science is replete with much more extreme simplifications---for example, the physics of spin-glasses is described in terms of the Ising model, which completely and knowingly falsifies the spatial structure of charges inside the spin-glass (Hertz, Krogh, and Palmer, 1991); models of avalanches in "sand-piles" use local rules for "collapse" which are known to be qualitatively different from the rules that govern actual sand-piles (Bak, 1997); "neural network" models of brain function knowingly and fundamentally falsify a range of key neuroscientific facts (Sejnowski, 1986; although see O'Reilly and Munakata, 2000). This explanatory phenomenon is so universal, even in physics, that it has led some to doubt the global coherence of physical science (Cartwright, 1983). The moral here, though, is that classical economists need not be defensive purely because they use wildly implausible assumptions; or that, indeed, this simplifying assumptions seem to be essential rather than a mere matter of convenience (because once they are abandoned the entire analytical apparatus becomes unworkable). Idealizations may be radical simplifications, and to a degree, falsifications of, reality---but if their resulting predictions provide elegant explanations and good predictions, then this is merely business as usual for scientific enquiry (Friedman, 1953).

The burden, then, rests with those who believe that classical economic theory's idealizations are, in some crucial way, over-simple. The burden is, specifically, to show that by using a (presumably, slightly) more realistic model of economic agents, better economic explanations and predictions will be obtained. One theme in this special issue is a discussion of the ways in which economic idealizations may be over-simple, with an eye to providing the basis for a richer economic theory.

Note that economic and decision analysis can be made more complex along many different dimensions. One dimension is to allow that economic agents may have imperfect information; although analysis rapidly becomes difficult if we allow that each agent's imperfect information may be idiosyncratic. Nonetheless, to understand the competitive behaviour of firms, the geographical spread of technology and innovation, or the role of "inside knowledge" in markets, a richer representation of the information available to the decision making process of each type of economic agents may be necessary (e.g., Tirole, 1988). Another dimension is to allow that economic agents may have imperfect cognitive

abilities to process the available information, and then to investigate whether these limitations are universal across all people, which would give us the possibility to derive general decisionmaking theories able to predict choice behaviour. This very important psychological dimension, which is the focus of numerous publications in psychology, economics, and all other social sciences, is "bounded rationality"---the fact that, even given certain information, people are not able to use it optimally (Simon, 1959, 1992). At a broad level, bounded rationality is a mathematical necessity---in general, the sophisticated calculations involving probability and decision theory that are invoked in economic analysis are known to be, in general, computationally intractable, and therefore, presumably, beyond the computational powers of the brain (Oaksford and Chater, 1998). But, in the light of my previous discussion, this general point is methodologically uninteresting. The key question is how far specific cognitive limitations, or systematic departures from the dictates of economically rational thought and behaviour, can be identified, and used to explain economic and social phenomena (see Camerer, 1998 for a recent review of applications in individual decision making). This general programme also motivates much of the sub-fields of behavioural decision theory (Slovic, 1977), experimental economics (Kagel and Roth, 1995) and behavioural finance (e.g., Sheffrin, 1999).

The aim of this article is to show that adding a new dimension (of complexity) related to certain fundamental aspects of human cognition may have important implications for some aspects of economics and decision sciences.<sup>1</sup> The cognitive claim, I argue for, is that people are not able to represent absolute magnitudes of stimuli of any kind (including a choice option's attributes like utilities, payoffs, and probabilities). Instead, they represent magnitudes ordinally---in relation to other magnitudes that they can sample from memory or from the current environment. This framework was first described in the *decision by sampling* theory proposed by Stewart, Chater, and Brown (2006), who also argued that, when people represent a magnitude, they can only do so on the basis of whether it is *larger* or *smaller* than other magnitudes sampled from memory or from the immediate context. Here, I argue that such sampling of knowledge produces certain biases in judgment, because stimuli are judged only relative to each other and therefore utility of an option is dependent on the other options that can be retrieved from memory. As a consequence, there is no ability to represent on any cardinal scale, the absolute value of a magnitude of any kind.

The core evidence for this claim comes from the study of the perception of the intensity of basic psychophysical magnitudes such as the brightness of a light or the loudness of a sound. Much traditional research in psychophysics has assumed the existence of some cardinal internal scale of intensities, onto which physical stimulation must somehow be mapped; and there has been consequent debate concerning the nature of this mapping (e.g., whether it is logarithmic, as argued by Fechner, 1966; or a power law, as argued by Stevens 1957). But more recent theory (reviewed and analysed in Laming, 1997) suggests a different point of view---that the very idea of an internal scale is incoherent. In the next section I digress briefly into psychophysics, before applying the resulting conclusion to an economic context in the following sections. Note, though, that the parallel between the two cases is relatively close. After all, just as perceptual theorists traditionally assumed that people had internal scales for the representation of loudness and brightness, so a traditional economic

<sup>&</sup>lt;sup>1</sup> This is also a part of a more general program aiming to ground decision-making research more directly on the underlying cognitive mechanisms that produce choice behaviour (e.g., see Oaksford & Chater, 1998).

picture of an agent assumes that the agent must have internal scales for the representation of the utility of various outcomes; for representing the probability that they will occur; and so on. Without a scale for utility or probability, the model of the economic actor would look very different (and, as I will briefly consider later, perhaps different in the crucial way identified above---in providing the possible basis for a richer analysis of economic phenomena).

## **MOTIVATION FROM PSYCHOPHYSICS**

So let me begin with a brief digression into psychophysics. A well-known and puzzling paradox of psychophysical magnitude perception is that people are rather good at *discriminating* the intensity of different magnitudes, but remarkably poor at categorizing them absolutely. That is, people are typically able to tell which is the louder of two sounds, or the brighter of two lights, with an accuracy that would lead to the naïve impression that they can tell the difference between of the order of a hundred different physical intensity levels. Yet when people are asked to explicitly associate intensity levels with category labels (e.g., 1 for the least intense stimulus, 2 for the next most intense, to, say, 7 for the most intense), they find this astonishingly difficult. Rather than having of the order of a hundred or so different cognitive "bins" into which items can be reliably categorized, performance tails off when about *five* different intensity levels must be categorized. Critically, it matters very little either what the absolute intensity level of the items to be categorized actually is; or, more shockingly, what the range of those items is. That is, so long as the items are readily discriminable from each other, to a first approximation, all that matters is the number of items that must be classified.

This is hard to reconcile with the idea that items are represented on an internal scale, according to which it would seem almost inevitable that performance would (a) be far higher than is observed; (b) would be degraded, if at all, in proportion to the crowdedness of items along the internal scale. The alternative viewpoint is that people do not construct any kind of internal scale---instead they are only able to make ordinal judgements, concerning which stimulus items are more intense than others (Laming, 1997). That is, I assume no more than the ability to make binary discriminations, rather presupposing the existence of internal psychological scales. The motivation for the restriction to discrimination comes from previous work on the direct judgment and the absolute identification of psychophysical magnitudes, such as luminance and sound pressure. In particular, Laming (1997) has shown that empirical data in line with Stevens' power law relating psychophysical variables and free numerical judgments can arise without assuming any representation of absolute information. Stewart, Brown, and Chater (2005) developed a theory of this psychological task (absolute magnitude identification), which embodies these assumptions and successfully predict the approximate limit of five items, as well as makes detailed predictions about the correlations across trials, and the nature of confusion errors.

Another example supporting this viewpoint is an elegant experiment conducted by Garner (1954), who asked participants to judge whether tones were more or less than half as loud as a 90 dB reference loudness. Participants' judgments were entirely determined by the range of tones played to them. Participants played tones in the range 55-65 dB had a half-loudness point, where their judgments were "more than half as loud" 50% of the time and

"less than half as loud" 50% of the time, of about 60 dB. Another group, who received tones in the range 65-75 dB had a half-loudness point of about 70 dB. A final group, who heard tones in the range 75-85 dB, had a half-loudness point of about 80 dB. Garner's experiment indicates, therefore, that people have no idea of the absolute intensity of the sound or what it means for one sound to be half as intense as another. Instead, it seems that people adjust their responses depending on the presented sound intensities from which they are asked to choose.

Other examples of similar context effects abound in psychophysics. Thus, empirical investigations in absolute identification (e.g., Garner, 1953; Holland and Lockhead, 1968; Lockhead, 1984; Luce, Nosofsky, Green, and Smith, 1982; Ward and Lockhead, 1970), magnitude estimation (e.g., Jesteadt, Luce, and Green, 1977), relative intensity judgment (Lockhead and King, 1983), and matching tasks (Stevens, 1975), have shown that perceptual judgments of stimuli varying along a single psychological continuum are strongly influenced by the preceding material. A robust finding is that current responses (judgments) tend to be contrasted (i.e. negatively correlated) with immediately preceding stimuli and assimilated (positively correlated) toward previous responses. Laming (1997) provides an extensive discussion of other similar findings and summarises many decades of psychophysical research, the results of which are consistent with the idea that participants are unable to make reliable decontextualised judgements of absolute sensory magnitudes. He claims that only relative judgements can be made - so whenever isolated stimuli are presented, and a judgement about the magnitude of the resulting sensation must be made, there is always a implicit comparison baseline of some kind. This might be stimulus presented on a previous trial, or may be some undifferentiated amalgam of remembered experience.

In summary, context effects, like those found by Garner (1954), are consistent with participants making perceptual judgments on the basis of relative magnitude information, rather than absolute magnitude information (see also Laming, 1984, 1997; Stewart, Brown, and Chater, 2002, 2005).

If the representation of utility is analogous to the representation of any other magnitude information, and in particular like the simple perceptual dimensions discussed so far, then the evidence presented above suggest that there is no fixed zero point on the utility scale and the experienced utility will shift depending on the context provided by the other experienced (consumption or risky) options. How worrying should this result be for economists?

Theories of decision making under risk typically and historically start from a normative standpoint, which is an economic theory of how decisions should be made (standardly, expected utility theory, von Neumann and Morgenstern, 1947); then assess the degree to which people do make decisions as they should (e.g., Kahneman, Slovic and Tversky, 1982; Kahneman and Tversky, 2000); and finally attempt to modify the normative theory to bring its predictions into line with people's actual behavior (e.g., prospect theory [Kahneman and Tversky, 1979]; regret theory [Loomes and Sugden, 1982]; rank-dependent utility theory [Quiggin, 1982]). The result involves building a bridge from economics (the normative domain) to psychology (the descriptive domain). The core elements of the normative economic approach are maintained. Typically there is a numerical representation of value or utility; a kind of representation of probability, or some related notion; the value of an uncertain outcome is computed by multiplying its 'value' and 'probability' attributes (or something similar). Within these constraints, however, theorists have devised a range of elegant and important models that capture a great deal of empirical data.

The alternative approach that I presented in this article, however, takes a different stance. Rather than starting from a normative economic theory, and attempting to make modifications that render it descriptively acceptable, I start from assumptions about elementary cognitive processes, and attempt to construct an account that can address the economic problem of choice under uncertainty. That is, my attempt is to bridge from psychology to economics, rather than the other way round, with the ultimate aim of testing whether the resulting model predicts economic behaviour. So the context dependence of the utility scale should only be of relevance for economics as far as there is a model that explains how the context affects preferences and hence economic choices. Stewart, Chater, and Brown (2006) propose a theory based on such links between basic cognitive principles concerning the representation of magnitudes, and economic behaviour.

### **TRANSFER OF UTILITIES ACROSS SCALES AND CONTEXTS**

Sampling from the past and from the current environment would make it almost impossible for people to create absolute representation of the utility of various choice options, which makes very difficult for people to make consistent choices matching the standards of the normative decision theories. However, these theories will be even less predictive if we realise that people are probably unable to build single utility scale along which to judge and compare options that are not very similar to each other (like for example various consumption goods).

There is recent evidence that if objects differ on more than one attribute then such a transfer could be problematic. In the psychology of perceptual judgment there is ongoing debate about the integrality and separability of psychological dimensions, arguing whether or not, for some pairs of psychological dimensions, judgment of the level of a stimulus on one dimension is interfered with by irrelevant variation on another dimension. There is evidence that for many pairs of dimensions, orthogonal variation on one dimension interferes with judgments of the level of a stimulus on the other dimension (e.g., Garner and Felfoldy, 1970; Ashby and Townsend, 1986; Lockhead, 1992). Such dimensions are said to be integral. It seems that, for integral stimuli, stimulus attributes are not represented independently of one another.

Stewart and Chater (2003) did loudness judgment experiment, in which on each trial, participants were presented simultaneously with a tone and a hiss and asked to judge which was louder. Their results showed that the assimilation of the current stimulus towards the previous stimulus was stronger when the current and previous stimuli were of the same type, and was attenuated when the current and previous stimuli were of different types, as was the accuracy of responding on the current trial. These data suggest that information about the loudness of a stimulus is not represented separately from the information about other stimulus attributes. If the information was represented separately then, for example, the effect of a tone or a hiss (of equal loudness) on the previous trial should have been the same. Stewart and Chater suggest that there may not be a single underlying scale representing loudness independently of other stimulus attributes. In particular, these data appear to suggest that people cannot consistently compare the loudness of two different types of sound: their binary

discrimination between categories of sound can be manipulated depending on the sounds they have heard on the previous trial.

Another interpretation is that the successful abstraction of a common scale depends on how similar are the two stimuli whose attributes are being judged. Note that common finding in research on sequential effects on judgment is usually an interaction between the previous stimulus and the previous response (two time-lagged variables). The assimilation towards the previous response seems to be modulated by the difference between the two consecutive (previous and current) stimuli (Jesteadt et al., 1977; Petzold, 1981). The closer the stimuli, the stronger the assimilation. Therefore, the conclusion is that if stimuli are very dissimilar then people contrast them so much so it becomes impossible to compare them realistically on a single scale. Probably the findings of Stewart and Chater can be explained in terms of the similarity between the stimuli, i.e., if the two stimuli, or their attributes, are very similar, then an ordinal loudness scale could be extracted and generalised independently of other stimulus attributes; while if the stimuli are very dissimilar (e.g., hiss and tone might be perceived as categorically quite dissimilar stimuli), then each type of stimulus is ordinally represented only on its own loudness scale. In support of this idea, Garner and Felfoldy (1970) showed that for certain dimension pairs that are rather dissimilar, like for example circle size and diameter angle, there was little or no facilitation or interference in direct stimulus sorting (scaling) task.

If these findings are transferred into the decision making domain and assume that judgments of value on any dimension are similar in nature and context dependent, as I argue here, then the conclusion is that the utility scale cannot be generalised over domains, situations, and product types, if products and domains are very dissimilar. For example, when people choose between various consumer goods it might be impossible for them to compare the utility from a holiday with the utility from CD player. Therefore consumer theory cannot be based on the standard indifference curves and needs some revision in light of the presented evidence about the locality of utility scales.

In summary, the findings presented here seem to be problematic for those accounts that suggest such sequential effects should be considered as a biasing of absolute judgment. Instead, I argue for a more radical alternative: that all there is, cognitively, is relative judgment. When there is a change in the stimulus type, removing at least part of the context used to make a relative judgment, then there will be a reduction in accuracy. The evidence presented in this section suggests that most of the representation of perceptual magnitudes is context dependent. In the next section, I move from psychophysics to behavioural decision making and present some relevant psychological evidence on the context dependence of judgments and choices under risk.

#### THE ROLE OF CONTEXT IN DECISION MAKING

A small number of experiments have investigated the effect of the context, i.e., the set of available options, on decision making under risk in a way analogous to the effects I have described in a psychophysical context, above. For example, the set of options available as potential certainty equivalents has been shown to affect the choice of certainty equivalent for risky prospects (gambles). In making a certainty equivalent judgment, participants suggest, or select from a set of options, the amount of money for certain that is worth the same to them as

a single chance to play the prospect. Birnbaum (1992) demonstrated that skewing the distribution of options offered as certainty equivalents for simple prospects, whilst holding the maximum and minimum constant, influenced the selection of a certainty equivalent. When the options were positively skewed (i.e., most values were small) prospects were under-valued compared to when the options were negatively skewed (i.e., most values were large).

Benartzi and Thaler (1998, 2001) have found evidence of another effect of the choice set by studying how people allocate their retirement funds across various investment vehicles. In particular, they find evidence for a diversification bias, which they call the 1/n heuristic. The idea is that when an employee is offered n funds to choose from in her retirement plan, she divides the money approximately evenly among the funds offered. Use of this heuristic, or others only slightly more sophisticated, implies that the asset allocation an investor chooses will depend strongly on the array of funds offered in the retirement plan. Thus, in a plan that offered one stock fund and one bond fund, the average allocation would be 50% stocks, but if another stock fund were added, the allocation to stocks would jump to two thirds. Read and Loewenstein (1995) also reported that people tend to diversify equally between the set of available options.

Simonson and Tversky (1992) also reported strong context effects but their evidence was that there is a general preference for the central options in each choice set, which they explained with what they called the compromise effect. For example, when participants had to choose between \$6 or famous brand pen, the introduction of a pen from a lesser known brand name increased the proportion of participants selecting the famous brand pen, and reduced the proportion selecting the \$6. Plausible account for this type of data was the notion of trade-off contrast, where participants, who are assumed to have little knowledge about the trade-off between two properties, i.e., they do not have a clear idea what is the exact utility of each option, deduce what the average trade-off is from the current or earlier choice sets. These data may reflect a more general tendency to prefer central options when choosing amongst set of options (also called extremeness aversion), which might be due to the relativistic way people derive the utilities of the choice options. Similar trend was observed when people choose between products on a supermarket shelf as shown by Christenfeld (1995). These results back up my earlier suggestion that preferences between different types of good may not be stable.

Simonson and Tversky (1992) also provide several cases where preceding material significantly influences current judgments in decision making. For example, when choosing between pairs of computers that vary in price and amount of memory, the trade-off between the two attributes in the previous choice affects the current choice. This result shows that by varying the preceding products the preference can be reversed. Such an effect of the preceding material is similar to the sequential context effects found in the psychophysical studies of perceptual judgment reported in the previous section.

In summary, the results reported above seem to indicate a variability of decision behaviour, which cannot be explained in terms of the existing normative economic theories. In the next section, I present a review of the research that contributed to the development of the idea about the relativistic nature of human judgments and decisions, and also present some recent models that could account for these data.

There have been quite a few recent publications on the role of cognition in individual decision making (Kahneman and Tversky, 2002, for a review) as well as interactive game

playing (e.g., Colman, 2003). These persuasive accounts describe psychological phenomena in economic behaviour by introducing various non-standard principles of judgment and reasoning processes (like for example, loss aversion, non-linear and weighting of probabilities, team reasoning, stackelberg reasoning, etc.). The goal has been to explain psychological phenomena in decision making that orthodox economic theory, and its conventional extensions, cannot explain. I argue that in order to better understand human preferences, in addition, a model is needed of how the economic agent perceives and mentally represents the decision problem initially before any judgment and reasoning, consequently, takes place.

As an illustration of such an account, here I offer some results from a general research program that aims to ground accounts of rationality in general, and decision theory in particular, on the underlying cognitive mechanisms that produce the seemingly paradoxical behaviour. Existing models of rational choice, like expected utility theory for individual decision making, and also game theory for interactive decision making, are typically based on the underlying assumption that only the attributes of the risky prospect or the game need be considered when reaching a decision. In other words, these theories assume that the utility of a risky prospect or strategy is determined by the utility of the outcomes of the prospect or game, and transforms of the probabilities of each outcome. Then the assumption is that the decisions are based on these utilities.

There is recent evidence, however, that the attributes of the previously or currently seen risky prospects and games influence the decisions in the current prospect and game, which suggests that prospects and games are not considered independently of the previously played ones (Stewart et al., 2003; Vlaev and Chater, 2003, 2006; Vlaev, Chater, and Stewart, in press). In particular, Stewart et al. (2003) have argued for the existence of what they call "prospect relativity": That the perceived value of a risky prospect (e.g., "p chance of x") is relative to other prospects with which it is presented. In particular, Stewart et al. studied peoples' perception of utilities in individual decision making tasks in gambling situations. The initial expectation based on the psychophysical studies described above, is that the option set (i.e., the context) will affect peoples' choices because there is no fixed internal scale according to which people make their judgements of the values of certain options. The results demonstrated a powerful context effect in judging the value of different risky prospects - the set of options offered as potential certainty equivalents for simple prospects was shown to have a large effect on the certainty equivalents selected. For example, when during judging the value a 50% chance of winning £200 people have options of 40, 50, 60, and 70 pounds, the most popular choice is 60 and then second choice is 50. When people have options of 90, 100, 110, 120 pounds, the most popular choice is 100, and then second choice is 110. So the set of alternatives affects valuation by a factor of nearly 2! This effect was replicated despite monetary incentives designed to encourage participants to deliver accurate and truthful certainty equivalents. In another experiment, the set from which a simple prospect was selected was also shown to have a large effect on the prospect that was chosen.

Vlaev, Chater, and Stewart (in press) further verified the prospect relativity principle, originally discovered with abstract gambles, by demonstrating relativity of human preferences in financial decision making under risk. This study investigated how the range and the rank of the options offered as saving amounts and levels of investment risk influence people's decisions about these variables. In the range manipulation, participants were presented with either a full range of choice options or a limited subset, while in the rank manipulation they

were presented with a (positively or negatively) skewed set of feasible options. The results showed that choices of saving rates and investment risk are affected by the position of each option in the range and the rank of presented options, which suggests that such judgments and choices are relative.

Similar context effects were also found in a sequential setting during interactive decisionmaking when people play many one-shot Prisoner's Dilemma games with appropriate anonymity (Vlaev and Chater, 2003; Vlaev and Chater, 2006), thus providing a new type of anomaly for orthodox game theory. In particular, I found that the degree to which people cooperate in these games is well predicted by a function of the pay-offs in the game, the cooperation index (Rapoport and Chammah, 1965).<sup>2</sup> In particular, the participants were asked on each round of the game to predict the likelihood that their co-player will cooperate, and then to make a decision as to whether to cooperate or defect. The results demonstrated that the average cooperation rate and the mean predicted cooperation of the co-player in each game strongly depended on the cooperativeness of the preceding games, and specifically on how far the current game was from the end-points of the range of values of the cooperation index in each session. In particular, the actual and predicted cooperation rate for a particular game was higher if this game was closer to the highest cooperation index value in the given sequence, compared to a condition in which the same game was closer to the lowest value in the sequence. Thus the perceived cooperativeness of a game did not depend only on the absolute value of its cooperation index, but also on the position of this index value in comparison with the minimum and the maximum index values in each experimental session (condition) as would be expected if the "cooperativeness" of a game could not be represented absolutely. Another results was that in games with identical cooperation indices, people cooperated more and expected more cooperation in a game with higher rank position (in terms of its cooperation index) relative to the other games in the sequence. These results present a challenge to game theoretic models that assume that the attributes of each game in a sequence are independently considered from the other games that are played.

In summary, these findings on the role of context in decision making, which are reviewed here, present another challenge to the standard rational choice theory and game theory. But they also challenge descriptive theories of decision-making under uncertainty, including rank dependent utility theory (Quiggin, 1982, 1993), configural weight models (Birnbaum, Patton, and Lott, 1999), and prospect and cumulative prospect theories (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), which all assign a risky prospect with a value or utility that depends only on the attributes of that prospect. In addition, the results by Vlaev and Chater (2006) present another challenge to the standard game theory, which assumes that games in a sequence are considered independently (e.g., Fudenberg and Tirole, 1991).

However, there are some theories, in which the utility or value of a prospect is not independent of the other prospects in the choice set, and hence they are potential candidates that could account for the prospect and game relativity phenomena described here. Thus these theories embody the relativistic idea (defended in this article) that choice alternatives are judged only relative to each other, instead of being based on some absolute internal cardinal

<sup>&</sup>lt;sup>2</sup> The cooperation index was proposed by Rapoport and Chammah (1965) to give a measure of the probability with which human players tended to cooperate when playing the game. Specifically, the cooperation index is a simple function of the values in the payoff matrix in Prisoner's Dilemma. Roughly, it depends on the degree of incentive the players have in playing "defect," in terms of the size of the gain they may achieve, and the disincentive from an altruistic standpoint in terms of potentially harming the payoff of the other player.

scale (for value, utility, etc.), These theories are *regret theory* (Loomes and Sugden, 1982), the *stochastic difference model* (González-Vallejo, 2002), the *multi-alternative decision field theory* (Roe, Busemeyer, and Townsend, 2001), the *componential-context model* (Tversky and Simonson, 1993) and the *range frequency theory* (Parducci, 1965, 1974). I briefly present the basic ideas in each of them.

According to *regret theory* (Loomes and Sugden, 1982) when choosing between outcomes people expect and estimate possible feelings of regret they may have on experiencing each outcome of a prospect. Such anticipated feelings of regret modify the utility of an outcome that results from a particular choice with respect to the outcomes that would have resulted from other choices (which where not taken). Therefore in regret theory, the utility of a prospect is not independent of the other prospects in the choice set.

In the *stochastic difference model* (González-Vallejo, 2002) prospects are also judged relative to one another and the function comparing prospect attributes gives the difference between them as a proportion of the larger attribute (the theory also assumes that subjective prospect attributes are the real prospect attributes). This proportional difference strategy is a special case of the stochastic difference model. The proportions are summed over all attributes to give the overall preference for one prospect over another. Hence in this theory the utility of a prospect is also not independent of the other prospects in the set.

In the *multi-alternative decision field theory* (Roe, Busemeyer, and Townsend, 2001) attribute values are compared across options, and these differences are summed across all dimensions to produce the momentary "valences" for each option. Preferences are constructed for each option by integrating valences over time (the relative weight for each dimension is assumed also to vary in time). This process contrasts with the accumulation of absolute attribute values. Instead, valences represent the comparative affective evaluations. Thus, the choice between options is made in relative rather than absolute terms, as in the stochastic difference model.

Tversky and Simonson (1993) proposed the *componential-context model* as a model of context dependent preference devised to provide an account of trade-off contrast and extremeness aversion (Simonson and Tversky, 1992). According to the model, each attribute has a subjective value depending on its magnitude and the value of an option is a weighted sum of its attribute values. The background context is assumed to be the previous choice set, which modifies the weighting of each attribute (dimension) according to the trade-off between the attributes in that set. Thus, after the weighting of each attribute has been modified, the value of an option in the current set is then modified by the relative value of the option averaged over pair-wise comparisons with the other options in the choice set (i.e., the choice between options is made again in relative terms).

Range frequency theory (Parducci, 1965, 1974) models how people value or rate items that vary along a single psychological dimension. According to the theory, the subjective value given to an attribute is a function of its position within the overall range of attribute values, and its rank among the other attribute values (here I mean the attribute values considered, or available, when the judgment is made). Thus attributes are judged purely in relation to one another, which is what the prospect relativity phenomenon is demonstrating when the attributes of the risky prospects are compared; and Stewart et al. (2003) indeed claim that the range frequency theory can account for the effects of the choice set. In particular, their results showed that people choose consistently either the more risky position

within the total range of risk, and low or high ranking prospect when prospects are ranked by risk. Birnbaum (1992) also found his data on context effects in decision under risk to be consistent with the theory.

In summary, in theories where prospects are judged in relation to one another, as in regret theory, the stochastic difference model, multi-alternative decision field theory, the componential-context model, and range frequency theory, the same effects of the choice set can, under some circumstances, be predicted. These relational theories all have in common the idea that preferences are constructed for a given choice set.

My account for the results presented in this section is summarised by the claim that people have poor notions of absolute cooperativeness, risk, and utility, and instead make their judgments and decisions in relative terms (analogously to the presented psychophysical and cognitive theories of perception and judgment of information about magnitudes representing intensities of stimulus attributes). Recall that the experiments by Stewart et al. (2003), Vlaev and Chater (2003; 2006), and Vlaev, Chater, and Stewart (in press), were all motivated by evidence from absolute identification and magnitude estimation paradigms, which demonstrates that participants typically have poor access to absolute magnitude information, and instead they rely upon comparisons with recent or concurrent stimuli, as evident from the strong effect of preceding material demonstrated in these paradigms. The key message here is that these experiments challenge the assumption that people have access to some stable internal cardinal scale representing the absolute magnitude/intensity of stimuli during perceiving and judging the differences between them.

Such account departs fundamentally from previous work in this field, by modelling the highly flexible and contextually variable way in which people represent magnitudes (like sums of money, probabilities, time intervals, cooperativeness, etc.), rather than assuming that these magnitudes can be represented on absolute internal (cardinal) psychological scales. My conjecture is that the results from the studies presented here suggest that people use the context in order to derive the utility of a risky prospect or a strategy. Thus, if absolute judgments are impossible and judging by how much one option is better/worse than another option changes depending by the other options that are available, then the only reliable judgment that can be made is that one option is *just* better/worse than the other (without being able to say by how much). Therefore, the best a decision maker can do is to rank order the available choice options on each dimension for judgment (i.e., constructing an ordinal scale). Note, however, that I do not even need to postulate the existence of a stable internal ordinal scale, because all the decision maker is doing is making ordinal binary comparisons between the choice alternatives available in the working memory (perceived from the environment or retrieved from memory), which was originally proposed by Stewart, Chater, and Brown (2006). Using this decision strategy, the decision maker can determine which option is the best one, for example, by simply counting how many times each option was better than another option (in the context). As a result, the option that was most often better off relative to the sample, is obviously the best option to select. Given people's inability to make absolute judgment, such simple judgment heuristic is probably the most efficient to use. And indeed, there is some recent evidence that similar fast and frugal mental heuristics can lead to nearoptimal results in various cognitive domains (Gigerenzer and Goldstein, 1996; Gigerenzer, Todd, and the ABC Group, 1999).

In summary, any descriptive account of decision behaviour, should incorporate a model of agents' basic cognitive perceptual processes. In summary, I believe that the standard decision theory needs to be supplemented by a more general "cognitive decision theory," which grounds decision-making in the underlying cognitive mechanisms that produce choice behaviour.

#### **IMPLICATIONS FOR MODELS OF ECONOMIC BEHAVIOUR**

One novel aspect of this approach is that, although its primary focus is decisions involving key economic variables, money, risk, and time, these variables are treated as undifferentiated 'attributes,' which must be traded off against each other (which was recently proposed by Stewart, Chater, and Brown, 2006). This is in stark contrast to a normative economic account, where the nature of the trade-offs between these dimensions is specific to the attributes involved. Thus, trade-offs between pay-off and probability should be governed by a multiplicative combination of (transformed) pay-off and (transformed) utility (von Neumann and Morgenstern, 1947; Starmer, 2000); and trading off the time delay after which a pay-off is received against the size of the pay-off is governed by an exponential, or perhaps hyperbolic, time-discounting function (Loewenstein and Prelec, 1992). However, although the different structure of these trade-off is normatively appropriate, it might not reflected in people's psychological processes. Thus, the principles governing such trade-offs might be the same, whatever attributes are being combined. This hypothesis should be tested in future research; but if it is correct, then the theory of decision under risk, and decisions concerning time, appear to be special cases of the multi-attribute decisions, where these attributes might just as well concern different qualities of an article of clothing (price, look, material), as key economic variables (Roe, Busemeyer, and Townsend, 2001). From a normative economic perspective, such an approach would be scandalous---because it blithely ignores vitally important normative criteria. But from a psychological perspective, this approach is not unreasonable, if the mental processes underlying decision making draw on common psychological mechanisms.

Stewart, Chater, and Brown (2005) argue that the assumption that people do not have internal scales for value rejects Bentham's (1789/1970) notion that utility is calibrated on an internal psychological scale, which is also in sharp contrast from the psychological theories derived from economics, which make a similar assumption. Note that economists have also shifted away from assuming the existence of internal utility scales. For example, the standard "revealed preference" interpretation of utility in economics (Samuelson, 1937) takes utilities to be revealed by observable choices without further specification about the psychological nature of these utilities. Recall also that Savage (1954) generalized this assumption to utility and probability information simultaneously. Thus, from the revealed preference perspective, the utility and probability scales are derived from choice preferences, rather than from assumptions about psychological scales.

The conceptual framework presented here has interesting similarities with respect to this traditional view in economics. In both, the reveal preference perspective and the decision by sampling approach proposed by Stewart, Chater, and Brown (2005), people are assumed to have access only to their own binary preferences (or more generally, to binary comparisons between perceptual magnitudes). Therefore, to the extent that people have broader grasp of their own, more global, values (probabilities, etc), this must be inferred from sampling their

own past choices and other memories revealing their preferences. So, for a given person to gain any "global" insight into how much pleasure is gained from consuming a specific product, this person has to sample from her memory some related, comparable events where she consumes that product. If the consumption episode in question is preferred to these events sampled from memory, this "reveals" to the person that this was a good experience; if it is preferred to some past episodes, but dispreferred to as many, this reveals to the same person that the experience was moderate, and so on. Thus, to the extent that people have any global grasp of their views concerning their perspective on how valuable or probable some event is, they must "reveal" this, by sampling from their own binary preferences, just as the economist attempts to reconstruct utility and probability values from the entire set of a person's' binary preferences.

Yet, the account presented here also has very different properties from those of economics. The psychological approach outlined here assumes that sampling from memory is extremely limited, and also stochastic, people's judgments concerning magnitudes will be strongly influenced by the particular items that they happen to sample (see Stewart, Chater, and Brown, 2006, for a precise specification of this sampling model and some simulation results). In this account, these may be drawn from long-term memory of recent events, but also from magnitudes that have been presented in the decision problem that the person faces. Hence, people's assessments of pay-offs, probabilities and intervals of time will vary capriciously, and may be highly malleable, rather than corresponding to a stable ordering, as in normative economic theory. Indeed, the effects of sequential and simultaneous context discussed above confirm this prediction.

It is evident from the discussion so far that crucial to the framework presented here is the process of sampling comparison magnitudes from long-term memory. A critical question will be how such magnitudes are distributed in memory. If, for example, there are many memories of small sums of money, and few of large sums, then a medium-sized sum will be likely to be judged as relatively large---because it is larger than most comparison items. Stewart, Chater, and Brown (2006) assume that the distribution of magnitudes in memory reflects the distribution of magnitudes in the natural environment. This work connects with a range of recent research which views cognitive processes as adaptive reflections of environmental structure (Anderson, 1990; Anderson and Schooler, 1991; Chater and Brown, 1999; Oaksford and Chater, 1998).

Thus, it seems that the cognitive system is naturally wired to be history and context dependent and to search for comparison points in the memory of recent events that relate to the current problem. I could even argue that such constant sampling from our memories, which are derived from our interactions with the social and natural environment, and the inbuilt tendency to rely on the information reflecting the statistical structure of the environment in order to make adaptive decisions, has vital consequences for the knowledge transfer in the economy and society at large. Thus, we sample from people around us by observing what they do, and they also observe our preferences, and then in the future we all sample from our memories that are reflection of this sampling. Such view suggests that preferences are unstable within the individual even at most basic level and are socially (environmentally) determined and transmitted, and my major claim is that the transmission of preferences can be viewed as a form of knowledge transfer. Note that the traditional view in economics is that knowledge is socially generated and transmitted while preferences are endogenous and stable. Of course, there are fashions and tastes but these appear to be marginal to the economic theory of rational choice. Is there some implication of this view for the functioning of markets and the geographical propagation of preferences and prices? Well, my answer is that because preferences are so socially malleable, they may sustain any fixed equilibrium, but perhaps would still tend to stabilize as markets are connected (e.g.,

globalization may not just help integrate markets from the point of view of allocation of resources and prices, but also in integrating preferences).

In summary, inconsistency in judgment and decision making can come from imperfect biased sampling, which happens because people sample in reality mostly from one side of the scale (e.g., only bad quality products of certain type) and they adapt to that level and might become unable to make adequate judgments of value and utility when they encounter the same product in a context where its value is out of the range of values experienced before. Therefore people might find it difficult to integrate or transfer absolute knowledge about utility between sequences of experiences and the transfer is biased by the past cases. I also conclude that because of the inseparability and locality of some judgment scales, people usually end up with at best separate ordinal utility scales for different classes of good. But if there is no a single utility scale along each different products are compared, then these products cannot be compared systematically in order to derive the indifference curves ubiquitous in consumer theory. Therefore consumer theory and marketing analysis have to take into account the particular scale that is characteristic for each consumer group and product type.

### THE ROLE OF CONTEXT IN COGNITION

The approach and the evidence presented so far suggest that decision making is fundamentally context-dependent. Here I take a broader perspective by arguing that there is enough evidence that most of human cognition is context dependent and I believe that this is a consequence of the goal of the cognitive system to adapt flexibly to the dynamic environment. Thus, my account can be considered as an attempt to provide an adaptionist approach to decision making. This approach would require adaptive, efficient, robust, context-specific, domain-specific, species-specific behaviour. This approach is contrasted to the traditional rational approach that demands consistency, transitivity and contentindependence and context-independence for the resulting decisions. Therefore, current and future research should investigate decision making, preferences, and utility functions, from an adaptively normative viewpoint. The aim is to explore the possible adaptive value of people's deviation from the prescriptions of the traditional rational choice theories by taking adaptiveness and context-dependence in natural decision environments rather than consistency in arbitrary problems as the ultimate standard for good judgement and decision making. Such framework promises also to propose an evolutionary-normative alternative to the standard rational choice theory. Here I also argue that context-dependence is nor restricted to decision making, but is a general feature of our cognitive system, which affects all cognitive processes, and which appears to serve adaptive purposes.

Advocates of such ecological views of rationality (Evans and Over, 1996, 1997; Gigerenzer and Goldstein, 1996; Gigerenzer and Todd, 1999; though see Chater, Oaksford, Nakisa, and Redington, 2003, for a critical analysis of some aspects of these views) emphasise the contrast between everyday human behavior, the success of which must be judged in the context of a specific and complex environment, and abstract classical principles of rationality, which appear to be justified a priori, but which may crucially ignore constraints

imposed by the environment about which learning occurs. In short, the concern is that classical principles of rationality, on their own, are inappropriate as standards of real-world reasoning. For example, in Anderson's rational analysis framework (Anderson, 1990; Oaksford and Chater, 1994), agent's goals and environmental constraints are all used to modify one's understanding of what is optimal behavior in a particular context, although rational principles still play a key role in determining what behaviour is optimal (given the agent's goals and the environmental structure). This emphasis on the environment does count against the decontextualized study of human inference, which ignore content and context.

Note that this emphasis on the environment is a reaction to the attempt of some researchers to devise empirical tests of descriptive rational theories that are independent of specific contexts, beliefs, or utilities; which has led to a focus on internal consistency of behavior in highly artificial conditions, rather than on how behavior meshes with the environment. For example, there are psychological models of inference where normative theories are interpreted as models of mental calculation, not merely behavioral description. The paradigm example of such models are "mental logic" theories in the psychology of reasoning, which regard the syntactic proof theory for logic as the basis of the algorithms that implement logical inference in the mind (e.g., Braine, 1978; Fodor and Pylyshyn, 1988; Rips, 1994). However, these algorithms appear to be intractable and therefore cannot apply to complexities of real-world contextualised inference (Chater and Oaksford, 1990; Cherniak, 1986; McDermott, 1987; Oaksford and Chater, 1991). On the other side, reasoning that may appear poor in an ecologically invalid laboratory context may be highly adaptive in the natural environment, as has been extensively argued (Gigerenzer, Hell and Blank, 1988; Gigerenzer and Hoffrage, 1995; Gigerenzer and Murray, 1987; Oaksford and Chater, 1991, 1993, 1998). Thus, it is important to stress essential role of the environmental context in which reasoning takes place in order to understand everyday human inference and rationality (Oaksford and Chater, 1995).

Note that people do not need to calculate their optimal behaviour functions in order to behave adaptively, because deriving the optimal behaviour function is frequently very complex. They simply have to use successful algorithms; they do not have to be able to make the calculations that would show that these algorithms are successful (Chater et al., 2003). This viewpoint is standard in rational explanations of human and animal behaviour across a broad range of disciplines. As I pointed out, economists do not assume that people actually make complex game-theoretic or macroeconomic calculations (Harsanyi and Selten, 1988); zoologists do not assume that animals calculate how to forage optimally (e.g., McFarland and Houston, 1981); and, in psychology, rational analyses of, for example, memory, do not assume that the cognitive system calculates the optimal forgetting function with respect to the costs of retrieval and storage (Anderson and Milson, 1989; Anderson and Schooler, 1991). Chater et al. (2003) claim that such behavior may be built in by evolution or be acquired via a long process of learning--but it need not require real time computation of the optimal solution.

According to this framework, one way to account for at least the prospect relativity effects described in this paper, is to assume that the internal scale used to represent the items in question (e.g., the scale representing the utility of a prospect, or the cooperativeness of a game) not fixed but is stretchable and adaptive to the environmental stimuli; as if people have fixed quantity of dimensional capacity, which can be stretched or contracted to accommodate the task demand most efficiently (although even the assumption of such a flexible scale may

not be necessary as demonstrated by Stewart, Brown, and Chater, 2005). Therefore, the resolution of the scale will depend upon the task demands. This flexibility has the advantage that sensitivity can be task dependent, and that adaptation can occur. As a consequence, people are unable to make reliable judgements of absolute magnitudes because they do not have direct access to information about absolute magnitudes.

There are at least two existing models of context effects in perceptual identification conforming to these principles. The first is the *adaptation level theory* (Helson, 1964), which states that the judgement of a particular event is proportional to its deviation from the mean value of all other events – the adaptation level – which itself is assigned a neutral value. This implies that the sum of the judgements of all experiences (e.g., pain and pleasure) will not depend on the shape of the distribution of events, because the sum of the deviations from the mean is zero. Note, though, that adaptation level theory does not explain how people's judgments of absolute magnitudes seem relatively sensitive to the range of the items involved, and would require some extension in order to deal with this. The second possible account is the range frequency theory proposed by Parducci (1968, 1974), which I have already discussed. Parducci found that the neutral point of the scale did not correspond to the mean of the contextual events (contrary to the adaptation level theory), but rather to a compromise between the midpoint (defined by the range) and median (depending on the skew of the distribution) of the distribution of contextual events. For example, satisfaction judgements depend on the skew of the distributions even when the means of the distributions are the same. Thus, the range principle reflects tendency to judge an event relative to the proportion of the range of stimuli lying below that event on the specified dimension of judgment, while the frequency principle reflects a tendency to judge an event relative to the proportion of contextual stimuli lying below that event on the specified dimension of judgment (which therefore depends on the rank order of the stimuli). In summary, the subjective value given to an attribute is a function of its position within the overall range of attributes, and its rank. Thus attributes are judged purely in relation to one another.

In summary, by claiming that judgments and decisions are based on the context rather than on some absolute judgment of value or utility, I also attempt to provide, a speculative general theoretical framework that will present the *context* as a explanatory concept which could account for the effects of different factors employed in the models of choice attempting to account for people's deviation from the predictions of the traditional rational choice models (such explanatory constructs for example are 'frames', 'categories', and 'discourse'). In the framework presented here, the major assumption is that every form of decision making happens in some cognitive context and this context plays the main causal role in people's behaviour.

#### CONCLUSION

In summary, the key difference between the approach I developed here and those approaches derived from normative economic accounts is that I do not assume that people internally represent values, probabilities, temporal durations, or indeed any other magnitudes. Instead, I assume only that people can sample items from memory, and can judge whether those items are associated with a higher or lower value (or probability, or duration) than the

present item. That is, I assume no more than the ability to make binary discriminations, rather presupposing the existence of internal psychological scales. The motivation for the restriction to discrimination, comes from previous work on judgment and absolute identification of psychophysical magnitudes, and also from recent work on context effects on judgments and decisions under risk and uncertainty (see Stewart, Chater, and Brown, 2006, for a model based on these assumptions). The assumption that people do not have internal scales for value, probability etc., constitutes a break from Bentham's (1789/1970) notion that utility is calibrated on an internal psychological scale; and thus a break too from the psychological theories derived from economics, that make a similar assumption. Finally, I argue that solving the heterogeneous problems arising from the standard rational theories of choice requires looking at how people represent the decision problems and the quantities that define them depending on the decision context.

Here one could raise the question whether this added level of detail is really going to matter to economists. After all, as I discussed at the beginning of this paper, although economists use radical simplifications, if their resulting predictions provide elegant explanations and good predictions, then this appears to be characteristic for every scientific enquiry. I argue that it may matter to incorporate the level of detail that I discuss here because by taking into account the factors governing perception and judgment of magnitudes, I can obtain better economic explanations and predictions.

For example, if the assumption is that judgments of value on any dimension are similar in nature and context dependent, as I argue here, then the conclusion is that the utility scale probably cannot be generalised over domains, situations, and product types, if products and domains are not very similar. The important implication of such a view for consumer choice theory is that it cannot be based on the standard indifference curves analysis, and needs some revision in light of the presented evidence about the locality and relativity of utility scales (for example, by taking into account the particular scale that is characteristic for each consumer group and product type). Note that context dependent models of magnitudes perception can be incorporated without substantially complicating the calculations (for example, the range frequency theory is formally much simpler than most of the standard normative and descriptive decision theories).

Another consequence of the view presented in this article is that preferences are unstable within the individual even at most basic level and are socially determined and transmitted, which also implies that they may sustain any fixed equilibrium. Therefore, the transmission of preferences can be viewed also as a form of knowledge transfer, and because markets are interconnected, preferences would still tend to stabilize at certain equilibrium points.

I believe that further research is required to integrate the findings presented in this paper with the standard economic theory in a way that might justify economists in replacing their current rather severe idealizations.

#### References

Anderson, J. R. (1990). *The adaptive character of thought*, Hillsdale, N.J.: Lawrence Erlbaum Associates.

- Anderson, J. R., and Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96, 703-719.
- Anderson, J. R., and Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 1, 396-408.
- Ashby, F. G., and Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, 93, 154-179.
- Bak, P. (1997). How Nature Works: the science of self-organized criticality. Oxford: Oxford University Press.
- Benartzi, S., and Thaler, R. (1998). Illusory diversification and retirement savings, working paper, University of Chicago and UCLA.
- Benartzi, S., and Thaler, R. (2001), Naive diversification strategies in defined contribution saving plans, *American Economic Review*, 91, 79-98.
- Bentham, J. (1970). *An introduction to the principles of morals and legislation* (J. H. Burns and H. L. A. Hart, Eds.). London: The Athlone Press. (Original work published in 1789),
- Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choice-based certainty equivalents. *Psychological Science*, 3, 310-314.
- Braine, M. D. S. (1978). On the relation betwen the natural logic of reasoning and standard logic. *Psychological Review*, 85, 1-21.
- Cartwright, N. 1983. How the Laws of Physics Lie. Oxford: Oxford University Press.
- Chater, N., and Brown, G. D. A. (1999). Scale-invariance as a unifying psychological principle. *Cognition*, 69, 17-24.
- Chater, N., and Oaksford, M. (1990). Autonomy, implementation and cognitive architecture: A reply to Fodor and Pylyshyn. *Cognition*, 34, 93-107.
- Chater, N., Oaksford, M., Nakisa, R., and Redington, M. (2003). Fast, frugal and rational: How rational norms explain behaviour. *Organisational Behavior and Human Decision Processes*, 90, 63-86.
- Cherniak, C. (1986) Minimal rationality. Cambridge, MA: MIT Press.
- Christenfeld, N. (1995). Choices from identical options. Psychological Science, 6, 50-55.
- Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences*, *26*, 139-153.
- Evans, J. St. B. T., and Over, D. E. (1996). *Rationality and reasoning*. Hove, Sussex: Psychology Press.
- Evans, J. St. B. T., and Over, D. (1997). Rationality in reasoning: The problem of deductive competence. *Cahiers de Psychologie Cognitive*, 16, 1-35.
- Fechner, G. T. (1966). *Elements of psychophysics* (H. E. Adler, Trans.). New York: Holt, Rinehart, and Winston. (Original work published 1860).
- Fodor, J. A., and Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3-71.
- Friedman, M. (1953). *Methodology of positive economics*. Chicago: Chicago University Press.
- Fudenberg, D., and Tirole J. (1991). Game Theory. Cambridge, MA: MIT Press.
- Garner, W. R. (1953). An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology*, 46, 373-380.
- Garner, W. R. (1954). Context effects and the validity of loudness scales. *Journal of Experimental Psychology*, 48, 218-224.

- Garner, W. R., and Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, *1*, 225-241.
- Gigerenzer, G., and Goldstein, D. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650-669.
- Gigerenzer, G., Todd, P., and the ABC Group (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.
- Gigerenzer, G., Hell, W., and Blank, H. (1988). Presentation and content: The use of baserates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 513-525.
- Gigerenzer, G., and Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684-704.
- Gigerenzer, G., and Murray, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum.
- González-Vallejo, C. (2002). Making trade-offs: A probabilistic and context-sensitive model of choice behavior. *Psychological Review*, 109, 137-154.
- Harsanyi, J. C., and Selten, R. (1988). *A general theory of equilibrium selection in games. Cambridge*, MA: MIT Press.
- Helson, H. (1964). Adaptation-level theory. New York: Harper and Row.
- Hertz, J., Krogh, A., and Palmer, R. G. (1991). *Introduction to the Theory of Neural Computation*. Addison-Wesley, Redwood City CA,
- Holland, M. K., and Lockhead, G. R. (1968). Sequential effects in absolute judgments of loudness. *Perception and Psychophysics*, 3, 409-414.
- Holyoak K., and Koh K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition*, 15, 332-340.
- Holyoak, K., and Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.
- Jesteadt, W., Luce, R. D., and Green, D. M. (1977). Sequential effects of the judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 92-104.
- Kagel, J. H., and Roth, A. E. (1995). The handbook of experimental economics. Princeton University Press.
- Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263-291.
- Kahneman, D, and Tversky, A. (Eds.). (2000). *Choices, values and frames*. New York: Cambridge University Press and the Russell Sage Foundation.
- Kahneman, D., Slovic, P., and Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
- Laming, D. R. J. (1984). The relativity of 'absolute' judgements. *British Journal of Mathematical and Statistical Psychology*, 37, 152-183.
- Laming, D. R. J. (1997). The measurement of sensation. London: Oxford University Press.
- Lockhead, G. R. (1984). Sequential predictors of choice in psychophysical tasks. In S. Kornblum and J. Requin (Eds.), *Preparatory states and processes* (pp. 27-47). Hillsdale, NJ: Erlbaum.
- Lockhead, G. (1992). Psychophysical scaling: Judgments of attributes or objects? *Behavioral and Brain Sciences*, *15*, 543-601.

- Lockhead, G. R., and King, M. C. (1983). A memory model of sequential effects in scaling tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 461-473.
- Loomes, G., and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal*, *92*, 805-824.
- Loewenstein, G., and D. Prelec. (1992). Anomalies in Intertemporal Choice: Evidence and an Interpretation. *Quarterly Journal of Economics*, 107, 573–597.
- Luce, R. D., Nosofsky, R. M., Green, D. M., and Smith, A. F. (1982). The bow and sequential effects in absolute identification. *Perception and Psychophysics*, *32*, 397-408.
- McDermott, D. (1987). A critique of pure reason. Computational Intelligence, 3, 151-160.
- McFarland, D., and Houston, A. (1981). *Quantitative ethology: The state space approach*. London: Pitman.
- O'Reilly, R. C., and Munakata, Y (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain.* MIT Press.
- Oaksford, M., and Chater, N. (1991). Against logicist cognitive science. *Mind and Language*, *6*, 1-38.
- Oaksford, M., and Chater, N. (1993). Reasoning theories and bounded rationality. In K. I. Manktelow, and D. E. Over (eds.), *Rationality*, (pp. 31-60). London: Routledge.
- Oaksford, M., and Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*, 608-631.
- Oaksford, M., and Chater, N. (1995). Theories of reasoning and the computational explanation of everyday inference. *Thinking and Reasoning*, *1*, 121-152.
- Oaksford, M., and Chater, N. (1998). *Rationality in an uncertain world: Essays in the cognitive science of human reasoning*. Hove, Sussex: Psychology Press.
- Parducci, A. (1965). Category judgment: A range-frequency theory. *Psychological Review*, 72, 407-418.
- Parducci, A. (1974). Contextual effects: A range-frequency analysis. In L. Carterette and M. P. Friedman (Eds.), *Handbook of Perception*, Vol. II, 127-141. New York: Academic Press.
- Petzold, P. (1981). Distance effects on sequential dependencies in categorical judgments. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1371-1385.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organisation*, *3*, 323-343.
- Quiggin, J. (1993). Generalized expected utility theory: The rank-dependent model. Boston: Kluwer Academic.
- Rapoport, A., and Chammah, A. (1965). *Prisoner's dilemma: a study in conflict and cooperation*. Ann Arbor: University of Michigan Press.
- Read, D., and Loewenstein, G. (1995). Diversification bias: explaining the discrepancy in variety seeking between combined and separated choices, *Journal of Experimental Psychology: Applied, 1,* 34-49.
- Rips, L. J. (1994). The psychology of proof. Cambridge, MA: MIT Press.
- Roe, R. M., Busemeyer, J. R., and Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, 108, 370-392.

- Samuelson. P. (1937). A note on the measurement of utility. *Review of Economic Studies*, *4*, 155-161.
- Savage, L. J. (1954). The foundations of statistics. New York: Wiley.
- Sejnowski, T.J. (1986) Open questions about computation in cerebral cortex. In: Parallel Distributed Processing: Explorations in the Microstructure of Cognition Vol. 2: Psychological and Biological Models, ed. McClelland, J.L. and Rumelhart, D.E., eds. MIT Press.
- Shefrin, H. (1999). Beyond Greed and Fear: Understanding Bahvioral Finance and the *Psychology of Investing*. Harvard Business School Press.
- Simon, H. A. (1959). Theories of decision-making in economics and behavioral science. *American Economic Review*, 49, 253-283.
- Simon, H. A. (1992). *Economics, bounded rationality and the cognitive revolution*. Aldershot, England: Elgar.
- Simonson, I., and Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29*, 281-295.
- Slovic, P. (1977). Behavioural decision theory. Annual Review of Psychology, 28, 1-39.
- Starmer, C. (2000). Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk, *Journal of Economic Literature*, *38*, 332-382.
- Stevens, S. S. (1957). On the psychophysical law. Psychological Review, 64, 153-181.
- Stevens, S. S. (1975). Psychophysics. New York: Wiley.
- Stewart, N., and Chater, N. (2003). No unified scales for perceptual magnitudes: Evidence from loudness. In: *Proceedings of the 25rd Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.
- Stewart, N., Brown, G. D. A., and Chater, N. (2002). Sequence effects in categorization of simple perceptual stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 3-11.
- Stewart, N., Brown, G. D. A., and Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review*, 112, 881-911.
- Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53, 1-26.
- Stewart, N., Chater, N., Stott, H. P., and Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General*, 132, 23-46.
- Tirole, J., (1988). The theory of industrial organization, Cambridge, Mass.: MIT Press.
- Tversky, A., and Simonson, I. (1993). Context-dependent preferences. *Management Science*, *39*, 1179-1189.
- Vlaev, I., and Chater, N (2006). Game relativity: How context influences strategic decision making. Journal of Experimental Psychology: Learning, Memory, and Cognition, 32, 131-149.
- Vlaev, I., and Chater, N (2003). Effects of sequential context on judgments and decisions in prisoner's dilemma game. In: *Proceedings of the 25rd Annual Conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ.
- Vlaev, I., Chater, N., and Stewart, N. (in press). Financial prospect relativity: Context effects in financial decision-making under risk. *Journal of Behavioral Decision Making*.
- von Neumann, M., and Morgenstern, O. (1947). Theory of games and economic behavior. (2<sup>nd</sup> ed.). Princeton, NJ: Princeton University Press.

Ward, L. M., and Lockhead, G. R. (1970). Sequential effect and memory in category judgment. *Journal of Experimental Psychology*, 84, 27-34.

Chapter 42

# WHY DO PEOPLE CHANGE THEIR MINDS? EVIDENCE FROM THE PURCHASE OF LONG-TERM CARE INSURANCE<sup>\*</sup>

Anne Theisen Cramer<sup> $\dagger 1$ </sup> and Gail A. Jensen<sup>2</sup>

<sup>1</sup>Davenport University, 27650 Dequindre Road, Warren, MI, 48092, USA <sup>2</sup>Institute of Gerontology, Wayne State University, Detroit, MI, 48202, USA

### ABSTRACT

Today more than ever, people try to anticipate financial needs and to plan wisely for a lifetime of financial security. Information about financial options is plentiful, and financing for health and long-term care (LTC) is no exception. With all of the information and advice that is available, under what circumstances would a person decide that his/her decision was no longer the best option?

We address this question by looking at the market for LTC insurance, and estimate logistic regressions to model consumer decisions to drop or renew an existing LTC insurance policy. We explore events that occurred after the policy was last purchased and before the current policy was dropped or renewed. The price and benefit design of each policy is not directly observable so several proxy measures of the price of a policy are explored.

Data is obtained from the publicly available Health and Retirement Survey (HRS). Data from 2002 is used to identify those who have a LTC policy and to establish baseline financial circumstances. Data from 2004 is used to determine whether the policy was renewed, and to identify potentially influential events that occurred since 2002.

The study sample includes 1,375 individuals who reported an existing, private LTC insurance policy in 2002, and were therefore eligible to renew the existing policy before 2004. Proxy prices were calculated and assigned using publicly available price schedules.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Economics, Business and Finance*, edited by Klaus P. Hofmann, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Email: anne.cramer@davenport.edu. Or Email: atcramer@wayne.edu.

Preliminary findings suggest that price was an influential factor in the decision to drop an existing policy, even though the price of the policy did not increase as a result of age. Those with newer policies were less likely to allow a policy to lapse. Those with low levels of assets (less than \$200,000) were more likely to allow a policy to lapse, as were those with more than \$1.5 million in assets.

Our results suggest that financial considerations are important, and a thorough review of an individual's financial circumstances may be effective in enabling people to make a lasting choice when they decide how to plan for LTC.

#### INTRODUCTION

One of the biggest challenges facing Americans today is that of planning their retirement. Life expectancy is longer, retirement age is changing, and medical advances are providing opportunities that did not exist a few years ago. In 2004, life expectancy at birth was 77.8 years (National Center for Health Statistics, 2006), a full seven years higher than in 1970, and the average duration of retirement was 17.4 years, four years longer than in 1970 (Gendel, 1998). In order to adequately prepare for these retirement years, financial advisors are recommending earlier planning, and are recommending that the planning specifically address potential health care needs (Vanderhei, 2006; Sahadi, 2006), including long-term care (LTC).

In 1960, the majority of LTC spending was for nursing home care, with paid long-term care (LTC) accounting for only 3.1% of national health expenditures (National Center for Health Statistics, 2006). Today, LTC can be provided in a variety of settings including the home, and LTC spending accounts for 8.4% of national health expenditures (National Center for Health Statistics, 2006). LTC insurance policies have been marketed since the 1980s, but relatively few adults purchase them, and of those who purchase policies, many allow their policies to lapse before services are used. The industry association AHIP (America's Health Insurance Plans, 2004) reports that only about seven out of every ten purchased policies remains in force. In addition, industry underwriters cite lapse rates that have been as high as 7% per year, and approximate 2% per year in recent years (Morisato, 2004; O'Brien, 2004).

The potential for lapsed policies is one of great concern. Advisors typically warn seniors against purchasing LTC policies that they cannot afford. CNN and Money Magazine editors advise readers "Unless you're confident that you can afford the premiums for the long haul, don't sign on. You could waste thousands on a policy that lapses before you need it." (Feldman, et. al., 2000). If a policy is allowed to lapse, not only will the individual be without insurance coverage for LTC, but he or she will also be without the money paid as premiums. So why might an individual purchase a LTC policy, just to let it lapse at some future date? What factors might cause an individual to reverse an initial decision to purchase and decide that LTC insurance is no longer the best option?

#### BACKGROUND

LTC is assistance with basic care and can include nursing care, skilled care such as physical therapy and occupational therapy, personal care, custodial care, and household services. LTC insurance typically covers care when the insured becomes cognitively impaired or can no longer perform two or more standard "activities of daily living" (ADL's) such as dressing, eating, or bathing. The price of most LTC insurance is based on the age of the purchaser at the time of purchase and does not increase as a result of aging, providing a strong incentive to continue to renew the policy once it is in force.

Studies regarding long-term care insurance have typically focused on the initial decision to purchase LTC insurance, identifying influential factors such as information and knowledge, family circumstances, and availability of informal care. Several of these prior works are used as bases for empirical tests in this paper. The basic health insurance purchase decision is modeled by McKenna (1986). The purchase of long-term care insurance is described theoretically in works by Pauly (1990) and by Zweifel and Strüwe (1998). Gupta and Li (2004) model the LTC insurance purchase decision as an optimization problem across two periods (pre- and post-retirement).

In addition to studies that focus on the initial purchase decision, two recent studies have focused specifically on lapsed LTC policies using early HRS data. Finkelstein and McGarry (2005) acknowledge the high lapse rate – over 25% over five years in their sample – and suggest that information that becomes available after the initial purchase might induce subscribers to drop their policies. McNamara and Lee (2004) look at HRS data, and find high lapse rates, with only 23.2 percent of subscribers in their sample keeping their coverage over the five-year period of the study. They suggest that one explanation is a lack of accurate information about LTC insurance. Another is inaccurate information about the risk of needing LTC.

In the study presented here, we again consider long-term care insurance lapses. We focus on whether or not the consumer allows his or her LTC insurance policy to lapse, given that a policy is already in force. We include price as a determinant of the decision to allow an existing LTC policy to lapse, along with other potential determinants of this decision. We also use the newly introduced validation question in the Health and Retirement Survey to verify that respondents are not mistakenly describing Medicare as a long-term care policy.

### **METHODS**

Data

Data for the study are drawn from the Health and Retirement Study (HRS), an on-going, nationally representative survey of older adults in the U.S. (Health and Retirement Study, 2004). In 2004, the survey included 20,147 respondents. Each respondent was born in 1953 or earlier (age 51 or older in 2004), or was married to a respondent who was born in 1953 or earlier.

In this study, we focus on the decision to allow an existing LTC insurance policy to lapse. We include in our sample those who initially purchased a policy prior to 2002, and report that the policy is still in force in 2002. By including these individuals, we focus on policies that are comparable to those offered today, and we focus on potential repeat purchasers. We eliminate from the sample persons for whom essential data was missing, yielding a final analytic sample of 1,375 individuals who were eligible to renew an existing LTC insurance policy in 2004.

#### **Theoretical Framework**

We use a simple two period model of the decision to purchase long-term care insurance that takes into account the intertemporal nature of such policies. Consider an individual who is deciding whether to renew an existing insurance policy to cover the risk of a loss that could occur in the current period, the future, or both. The individual is risk averse with preferences defined by a von Neumannn-Morgenstern utility function u for consumption in each period. Her risk of a loss, L, increases over time. A loss occurs in the current period with probability  $p_1$ , and it occurs in the future with probability  $p_2$ , where  $p_1$  is less than  $p_2$ . Assume this individual is endowed with  $W_1$  units of wealth today, and  $W_2$  units in the future. Absent insurance, her ex-ante expected utility is given by:

$$EU_0 = [p_1 u(W_1 - L) + (1 - p_1)u(W_1)] + b[p_2 u(W_2 - L) + (1 - p_2)u(W_2)]$$
(1)

where *b* reflects her valuation of tomorrow's consumption compared to today's.

Suppose that a market for LTC insurance contracts exists in which there are willing suppliers of policies that pay an indemnity Y should a loss occur. Policies are renewable and there is a fixed premium-per-period, h, at which such contracts are traded. As long as h is paid to the seller each period, full coverage remains in force. One distinguishing feature of LTC insurance policies is that h remains constant over time. This model reflects this feature.

If the individual purchases a contract, her ex-ante expected utility will be

$$EU_{I} = [p_{1}u(W_{I}-h-L+Y)+(1-p_{1})u(W_{I}-h)] + b[p_{2}u(W_{2}-h-L+Y)+(1-p_{2})u(W_{2}-h)]$$
(2)

Clearly, she will purchase a policy today if and only if  $EU_I > EU_0$ . From this simple model we see that a number of factors are likely to influence the decision to renew the policy. The price, h, and the characteristics of coverage, Y, are clearly relevant. The decision also depends on an individual's resources ( $W_1$  and  $W_2$ ), her rate of time preference (b), the likelihood of incurring a loss and the magnitude of that loss  $(p_1, p_2, \text{ and } L)$ , as well as the individual's attitudes towards risk, which are reflected in u(). These considerations guide the selection of variables in our empirical work. Factors that influence the individual's attitudes toward risk and perceived likelihood of incurring a loss are relevant to the model, as discussed in prior studies (Pauly, 1990; Zweifel and Strüwe, 1996, 1998; Sloan and Norton, 1997; McCall et al., 1998; Finkelstein and McGarry, 2004; McNamara and Lee, 2004). We also recognize that information may have become available and events may have occurred after the initial decision to purchase and may therefore prompt the individual to change her decision. In our model, the consumer will allow LTC insurance to lapse if her expected utility in the insured state  $(EU_l)$  falls below expected utility in the uninsured state  $(EU_l)$ . More formally, we model the decision to allow LTC insurance to lapse by estimating variations of the equation:

$$D_i^* = \beta_0 + \beta_1 P_i + \beta_2 X_i + \beta_3 W_i + \beta_4 E_i + \varepsilon_i$$
(3)

where i indexes the individual. If  $D_i^* > 0$ , the individual chooses to let their LTC policy lapse, whereas if  $D_i^* < 0$  the individual renews their policy.  $P_i$  is the annual premium to be incurred,

 $X_i$  is a vector of individual characteristics,  $W_i$  is a measure of financial variables including wealth and income,  $E_i$  is a vector of events that have occurred since the policy was last reported in force, the  $\beta$ 's are vectors of coefficients, and  $\varepsilon_i$  is a random disturbance. We do not observe  $D_i^*$  directly. Rather, we observe whether each LTC policy holder has actually let their coverage lapse. Define  $D_i = 1$  if the policy has lapsed, and  $D_i = 0$  if it has not.  $D_i$  is the variable we observe. We assume  $\varepsilon_i$  follows a logistic distribution and estimate the parameters of (3) by maximum likelihood.

In 2002 and 2004, HRS participants were asked about basic health insurance (both private and government). They were then asked, "Not including government programs, do you now have any long term care insurance which specifically covers nursing home care for a year or more or any part of personal or medical care in your home?" If the individual answered, "Yes," then they were asked, "Is that one of the plans you have already described, or a different plan?" For our study, we include the individual in our sample if he replied "Yes" to the first question in 2002, indicating a LTC policy, and then replied "a different plan" to the second question, indicating that the LTC policy under discussion was different from his Medicare, Medicaid, or basic health insurance policy. We then use responses to the same questions in 2004 to determine whether the policy was renewed or allowed to lapse in 2004.

Overall, about 15% of people in our sample allowed a private LTC insurance policy to lapse in 2004. Of the 1375 observations in our sample, 1171 individuals decided to renew LTC insurance, and 204 decided to allow the policy to lapse. Based on data reported by many individual insurers, AHIP (America's Health Insurance Plans) estimates that 30% of all policies purchased are no longer in force. In our sample, only 15% of policies have lapsed, probably for two reasons. First, the AHIP rate includes policies that have lapsed due to death, while our sample includes only those who are living and choose not to renew. Second, the AHIP rate is cumulative, while we examine lapses over only the two-year period from 2002 to 2004.

#### **Explanatory Variables**

Explanatory variables fall into four general categories: price and characteristics of the policy, individual characteristics, financial variables, and recent events. A list of explanatory variables and descriptive statistics is presented in Table 1.

For this analysis the price of insurance, P, is measured by the annual premium the individual would have faced when the LTC policy was first purchased. For each respondent in the sample, the premium was calculated and assigned on the basis of the individual's age and health status at the time of initial purchase, using publicly available price schedules for LTC coverage (described below). The year of initial purchase was determined using self-reported duration of the contract. Consistent with standard insurance company practices, price discounts were given for non-smokers who did not report an identified pre-existing condition. Using this as our measure allows us to assign the price of a standard policy to all potential purchasers. This methodology has been used in past studies regarding health insurance (Leibowitz and Chernew, 1992) and term life insurance (Pauly et al. 2003, Brown and Goolsbee 2003), where prices were assigned based on industry practices.

## Table 1. Descriptive Characteristics of Potential Lapses (N=1375)

Variable	Mean	Standard Deviation	Expected Sign
Dependent Variable			·
Lapse: Equal to "1" if LTC insurance is Lapsed (not renewed) in 2004	0.15	0.36	
LTC Insurance			
G E Capital Price	2155.98	1357.28	+
AHIP Price	2404.00	1439.05	+
TIAA Price	3244.53	1512.11	+
Years Contract Was In Force	7.60	5.10	+
Demographic Variables			
Male	0.41	0.49	+
White	0.96	0.19	?
Black	0.03	0.17	?
Age	71.56	8.58	-
Married or Partnered	0.73	0.45	?
Number of Children	2.89	1.85	?
Education in Years (Max = 17, graduate level)	13.97	2.51	-
Geographic Variables			
Urban/Suburban = 1; $Rural = 0$	0.71	0.45	_
New England (ME, NH, VT, MA, RI, CT)	0.04	0.20	
Mid Atlantic (NY, NJ, PA)	0.09	0.29	
East North Central (OH, IN, IL, MI, WI)	0.16	0.37	
West North Central (MN, IA, MO, ND, SD, NE, KS)	0.18	0.38	
South Atlantic (DE, MD, DC, VA, WV, NC, SC, GA, FL)	0.23	0.42	
East South Central (KY, TN, AL, MS)	0.05	0.22	
West South Central (AR, LA, OK, TX)	0.06	0.24	
Mountain (MT, ID, WY, CO, NM, AZ, UT, NV)	0.05	0.21	
Pacific (WA, OR, CA, AK, HI)	0.13	0.34	

Variable	Mean	Standard	Expected
		Deviation	Sign
Health			
Health Status: Equal to "1" if self-reported status is at least "good."	0.84	0.37	?
ADL Count: Number of ADL's reported as needing help (maximum = 10).	1.95	2.28	_
Finances			
High Asset Level: Equals "1" if household assets > \$1.5 million.	0.06	0.25	+
Low Asset Level: Equals "1" if household assets < \$200,000.	0.38	0.49	+
Real Estate Equity (\$000's)	90	81	?
Income (\$000's). Household income divided by 2 if married; divided by 1 if not married.	51	114	_
Recent Events			
Recently Moved Residence	0.14	0.34	+
Spouse Recently Died	0.03	0.18	+
Parent Recently Died	0.03	0.18	_
Recently Married	0.01	0.08	-
Health Improved	0.73	0.44	+
Other			
Uses email and internet	0.45	0.50	_
Satisfied: Satisfaction with health insurance. (Equals "1" if very satisfied, "0" if not satisfied.)	0.97	0.16	_
Has a Will	0.85	0.36	

#### Table 1. Descriptive Characteristics of Potential Lapses (N=1375) (Continued)

In this study, we consider price (premium) variables based on three different industry models. The first variable, "GE Capital Price" is based on rate schedules filed by GE Capital, a major insurer in the LTC insurance market (State of Michigan, 1996). The second, called "AHIP Price," is based on a table of average rates nationwide for LTC insurance published by America's Health Insurance Plans (AHIP, 2004). The third, "TIAA Price," uses a mathematical formula developed by Gupta and Li using premium data from TIAA- CREF (Gupta, 2004) and assuming a standard set of benefits. It is expected that price will be positively correlated with the probability of allowing a LTC insurance policy to lapse. We also control for the age of the policy.

In addition to price and the policy's age, explanatory variables also include individual characteristics that are unlikely to have changed since the original purchase decision was

made, individual characteristics that may change, and events that may result in a decision to stop renewing a LTC policy.

Characteristics that are unlikely to have changed since a LTC policy was first purchased include gender, race, and level of education. Gender is included as a dummy variable. It is expected that men will be more likely to allow a LTC policy to lapse because of a lower life expectancy, implying that they may need LTC for less time. Race is included as a set of dummy variables in order to capture cultural differences that might impact the decision to renew LTC insurance. According to a national survey conducted for AARP (Belden, 2001), white individuals are less likely to provide informal care if LTC is needed, relying instead on paid formal care. In addition, prior studies (Pandya, 2005; Wallace, 1998) have shown that non-white cultures have traditionally relied more on family members to provide informal LTC. If this is the case, white individuals might be less likely to allow a LTC policy to lapse.

Education is expected to be correlated with a lower probability of lapsing LTC insurance for a number of reasons. First, education can potentially improve efficiency of health production (Kenkel, 1991), improving the individual's ability to make a lasting decision about LTC insurance when the policy is first purchased. Second, education can improve the individual's knowledge about health (Kenkel, 1991) and the need for LTC in the first place, making it less likely that new information will cause the individual to change his LTC decision.

Age is also expected to influence the decision to allow a LTC policy to lapse. First, because most policies are rated according to age at the time of purchase, an older policyholder who purchased the policy at an early age is paying less than the market price. Second, as a policyholder ages, there are fewer payments remaining. For these reasons, policies become relatively more valuable at older ages, and a policyholder may be less likely to allow the policy to lapse as he or she ages.

Demographic variables that may have changed include marital status, number of children, geographic variables, health status, and financial circumstances. Marital status and the number of children are expected to influence the LTC insurance decision in two ways. First, married individuals and those with children might expect informal care from their spouse or children to serve as a substitute for formal paid care. If this is the case, the presence of children or a spouse might diminish the need for LTC insurance (Pauly, 1990; Zweifel and Strüwe, 1998), making it more likely that the policy will be allowed to lapse. Alternatively, the desire to provide for a spouse or children is a possible incentive to retain a policy, since assets that would be used to pay for LTC in absence of insurance could be preserved for family members if the LTC policy is renewed. In our model, we include marital status, the number of children, and a dummy variable that indicates whether or not the individual has a legal will.

Geographic location of the respondent is measured by a set of regional dummy variables reflective of the nine Census regions in the US. In addition, a dummy variable reflecting an urban or suburban setting is included to approximate the likelihood of living near a nursing home or intermediate care facility. It is expected that proximity to a facility may decrease the probability of allowing a LTC policy to lapse. In our sample, 71% of respondents live in an urban or suburban setting.

It is expected that people who require help with Activities of Daily Living (ADL's) such as walking and rising from a chair will be less likely to allow a policy to lapse, since they may have a more immediate need for long-term care. Therefore, we include the variable "ADL Count" which indicates how many ADL's each person requires help with. "ADL Count" is expected to enter negatively, since those with a high ADL need are not expected to allow the policy to lapse.

Financial circumstances are expected to influence the probability of renewing LTC insurance in several ways. With regard to assets, we include two dummy variables that indicate if assets fall in the "high" range (over \$1.5 million) or "low" range (less than \$200,000). The reference category is mid-range assets (\$200,000 to \$1.5 million), and is based on a recommendation by Consumer Reports (2003) that those with assets in this mid-range are most likely to benefit from LTC insurance. For consumers with assets in the "low" range, Medicaid would likely cover LTC costs if care were needed. For those with assets in the "high" range, self-insurance (relying on personal funds) is a viable option if LTC is needed. Therefore, we expect LTC insurance lapse to be positively correlated with high and low asset levels.

In addition to asset level indicators, we include real estate equity, including the main home. Because LTC insurance is perceived by many as a way to protect one's home, equity may be negatively correlated with LTC insurance lapse. Alternatively, individuals can draw on the equity in their homes to pay for LTC, so equity may serve as a substitute for LTC insurance and be positively correlated with a lapse.

Income is expected to influence LTC insurance renewal because individuals with higher incomes are better able to continue to afford the LTC insurance premiums. In the model presented here, we include household income divided by 2 if married and divided by 1 if single. We expect that income will be negatively correlated with a lapse.

Finally, we include a number of dummy variables that reflect changes occurring over the last two years, since the policy was last reported "in force." These variables could influence the individual's financial circumstances, perception of risk, perception of the need for long-term care, and family circumstances. Any of these changes could result in a re-evaluation of the long-term care policy, and could possibly cause the policy to be dropped. The self-reported health improvement variable is of particular interest. A perception of improving health can reflect a lower probability of needing care, thereby increasing the probability of allowing a policy to lapse. Alternatively, it can reflect a higher life expectancy and therefore a higher probability of living long enough to require care, and a lower likelihood of allowing the policy to lapse.

Finally, we include two additional items of interest. We attempt to capture awareness of current information by including a dummy variable for whether the person uses email. We also include a dummy variable indicating whether the individual is satisfied with his or her overall health insurance package. If a policy is dropped, this satisfaction variable may indicate whether it was due to dissatisfaction with administration of the policy as opposed to inability to afford the premium or perception that long-term care insurance will not be needed.

#### RESULTS

Regression results are shown in Table 2. In this model, which includes only potential renewals (individuals who already had LTC coverage in the prior period), the coefficient on

price is positive and statistically significant. This suggests that price is a critical determining factor in the decision to renew, consistent with anecdotal evidence that individuals sometimes allow their policies to lapse because the policies become unaffordable. Marginal effects of price are small but significant, with a \$1000 increase in price corresponding to an increase of approximately 4 percentage points in the probability of lapse (Table 3).

	G E Capital Pricing		AHIP Pricing		TIAA Pricing	
	β	S. E.	β	S. E.	β	S. E.
Price	0.0004**	0.0001	0.0004**	0.0001	0.0004**	0.0001
Years Contract		0.000	0.0004.000		0.000044	
Was In Force	0.0902**	0.0206	0.0991**	0.0229	0.0909**	0.0205
Age	-0.0887**	0.0203	-0.0973**	0.0234	-0.0902**	0.0202
Male	0.3869*	0.1721	0.3935*	0.1716	0.3869*	0.1724
White	0.2107	0.6694	0.2086	0.6716	0.2412	0.6704
Black	1.3899	0.7642	1.4104	0.7647	1.4355*	0.7648
Education	-0.0478	0.0374	-0.0468	0.0371	-0.0458	0.0375
Married	-0.3193	0.2109	-0.3319	0.2094	-0.3330	0.2111
Children	0.0218	0.0513	0.0219	0.0511	0.0226	0.0517
Urban	-0.0403	0.2220	-0.0415	0.2215	-0.0432	0.2221
Mid Atlantic	-0.7236	0.4846	-0.7215	0.4854	-0.7380	0.4846
EN Central	-0.9631*	0.4864	-0.9581*	0.4881	-0.9731*	0.4867
WN Central	-2.0051**	0.5490	-2.0011**	0.5490	-2.0105**	0.5484
South Atlantic	-0.9176	0.4752	-0.9202	0.4766	-0.9326*	0.4752
ES Central	-0.3364	0.5347	-0.3483	0.5334	-0.3419	0.5342
WS Central	-0.7330	0.5845	-0.7408	0.5865	-0.7475	0.5865
Mountain	-1.0721	0.5905	-1.0437	0.5896	-1.0842	0.5892
Pacific	-1.1563*	0.5173	-1.1566*	0.5183	-1.1692*	0.5174
High Asset Level	0.9448*	0.4394	0.9540*	0.4407	0.9556*	0.4381
Low Asset Level	0.8174**	0.2022	0.8059**	0.2021	0.8159**	0.2022
Real Estate Equity (\$000's)	-0.0001	0.0002	-0.0001	0.0002	-0.0001	0.0002
Income (\$000's)	-0.0003	0.0006	-0.0003	0.0006	-0.0003	0.0006
ADL Count	0.1234**	0.0363	0.1242**	0.0362	0.1233**	0.0364
Health Status Improved	0.2628	0.1953	0.2774	0.1945	0.2625	0.1952
Changed Residence	0.2751	0.2410	0.2710	0.2426	0.2673	0.2415
Spouse Recently Died	-0.6467	0.6106	-0.6819	0.6115	-0.6627	0.6143
Parent Recently Died	-0.4314	0.4915	-0.4138	0.4925	-0.4329	0.4935

#### Table 2. Logit Model of Decision to Allow LTC Insurance Policy to Lapse in 2004

	G E Capital Pricing		AHIP Pricing		TIAA Pricing	
	β	S. E.	β	S. E.	β	S. E.
Recently Married or						
Partnered	0.6735	0.7772	0.6973	0.7730	0.6632	0.7842
Has Will	-0.9617**	0.2160	-0.9712**	0.2173	-0.9698**	0.2167
Uses Email	-0.3797	0.2245	-0.3788	0.2234	-0.3803	0.2250
Satisfied With Health						
Insurance	-0.6545	0.4337	-0.6854	0.4301	-0.6547	0.4350
Constant	5.0474**	1.6253	5.4606**	1.7220	5.1508**	1.6167
Number of						
Observations	1375		1375		1375	
McFadden's adjusted						
$R^2$	0.126		0.124		0.126	
Log Likelihood	-472.73		-473.60		-472.38	
LR (31)	209.16		207.41		209.86	

Table 2. (Continued)

\* Statistically significant at the  $\alpha = 0.05$  level.

\*\*Statistically significant at the  $\alpha = 0.01$  level.

#### **Table 3. Marginal Effects of Price Variables**

		Marginal Effect
	GE Capital Price	0.0372**
Price Proxy (\$000's)	AHIP Price	0.0400**
	TIAA Price	0.0344**

\*\*Statistically significant at the  $\alpha = 0.01$  level.

Having an older policy is positively and significantly correlated with LTC insurance lapse. This is expected, because current insurance policies are typically more comprehensive than older policies, and information about LTC financing is more prevalent today than it was 25 years ago. In addition, public coverage is more comprehensive than it was prior to 1989 (Norton and Kumar, 2000). Therefore, those with older policies may have based their original purchase decisions on the paucity of substitutes available at the time, and may be less in need of a private policy today.

While the coefficient on race suggests that black respondents are more likely to allow a policy to lapse, the finding was not significant, with only a 94% confidence level. It should be noted, however, that non-white individuals represented only 4% of the sample.

Age is negatively correlated with LTC insurance lapse, as expected. This may reflect the increased probability of needing care as one ages, and also the increasing value of a renewed policy that was priced at a younger age, resulting in a tendency to continue renewing the policy as one ages. Other demographic factors, including education, marriage, and children are not significant. It is possible that these factors influence the initial decision to purchase LTC insurance, but are not factors in the decision to renew or drop a policy once it is in force.

Whether an individual lived in an urban setting does not significantly influence the decision to keep a policy. At the same time, the results show that lapses in coverage are more common in some areas than others, the most notable example being the "West North Central" region (MN, IA, MO, ND, SD, NE, KS), which is negatively correlated with policy lapses. It is interesting to note that in 2002, only five states in the U.S. had a LTC market penetration of over 15% (AHIP, 2004). Four of these states are in the "West North Central" census region.

As expected, an individual's current financial circumstances are correlated with the probability of lapse. Those with household assets in the lowest range of "less than \$200,000" are more likely to allow a policy to lapse. It is possible that individuals with assets below this level expect to eventually qualify for Medicaid, and therefore do not perceive a need to renew a private LTC policy. Those with household assets above \$1.5 million are also more likely to allow a policy to lapse, possibly choosing to rely on personal funds if LTC becomes necessary. This finding is consistent with recommendations by financial advisors that those with assets in the middle range are most likely to benefit from LTC insurance. Real estate equity and income are not significantly correlated with a lapse.

Those who have a will are significantly less likely to allow a policy to lapse. This is as expected, because one motive for purchasing a LTC insurance policy is to preserve assets for children or a spouse.

Surprisingly, individuals who have difficulty with ADL's are more likely to allow a LTC insurance policy to lapse. LTC insurance is marketed as way to provide flexibility if care is needed, so one might expect those who are in a position to use LTC insurance benefits to continue to renew their policies. The finding here suggests the opposite.

Why might this occur? There are several possibilities. First, some individuals may have found that in the absence of a policy, they received some coverage of LTC services under Medicare or Medicaid, causing them to re-evaluate the necessity of continuing with a private policy. In early years of the HRS, respondents were asked if they had ever allowed a LTC insurance policy to lapse, and if so, why (HRS 1996-2002). Our finding here would be consistent with the second most frequently cited reason for allowing a policy to lapse: the private insurance was not needed.

Second, it is quite possible that those who have tried to obtain benefits under an existing policy had trouble receiving benefits they were entitled to (Duhigg, 2007). This would be consistent with another frequently cited reason for a policy lapse in the HRS survey: "general dissatisfaction" with the policy.

Finally, events that occurred over the past year, including deaths of family members and changes in residence and marital status do not significantly affect the decision to allow the policy to lapse.

#### CONCLUSION

In this empirical analysis of the decision to allow a LTC insurance policy to lapse, we find that financial considerations have a significant influence. Price was a significant factor in deciding whether to renew, even though most policies become a better deal as the person ages. Individuals with less than \$200,000 in household assets were more likely to allow a policy to lapse, possibly reflecting the potential to qualify for Medicaid. Those with assets of

over \$1.5 million were also likely to allow the policy to lapse, possibly relying on personal funds for LTC if care becomes necessary. Finally, those who had a will were less likely to drop the policy, probably with the intent of preserving assets for their heirs.

In this study, those who had trouble with ADL's were more likely to allow a policy to lapse. This may suggest a perception that the policy is not needed after all, or it may reflect dissatisfaction with administration of the policy. Older individuals were less likely to drop a policy.

Our findings suggest several approaches for the design and regulation of long-term care policies. First, the statistical significance of the price supports the claim that when an individual does allow an existing policy to lapse, it may be because of affordability. Second, the lower probability of renewal for people in the low-asset range suggests that Medicaid may, in fact, be crowding out some private purchase. These findings suggest that a thorough review of an individual's financial circumstances may be effective in enabling people to make a lasting choice when they decide how to plan for LTC. The importance of careful assessment has been a focus for senior advisors, since premiums paid into a policy are lost if the policy eventually lapses before care is needed.

There has been much discussion about people's understanding of LTC insurance and, in particular, concern that they believe they have LTC insurance coverage when in fact they have only a standard Medicare or basic health policy (AARP, 2001). Through the follow-up question "Is this a policy you've already described?" we are able to validate responses, identifying and recoding those who describe Medicare, supplemental, or basic health insurance as a LTC insurance policy. While this verification has greatly improved reporting accuracy, there is still a possibility that some people may report LTC insurance when in fact they have Medicare. Since people do not drop Medicare, this would tend to understate the lapse rate.

#### ACKNOWLEDGMENTS

This research was supported by a Ruth L. Kirschstein National Research Service Award (Grant No. 1-F32AG023438-01) from the National Institute of Health/National Institute on Aging. The conclusions presented here are the responsibility of the authors and do not necessarily reflect the views nor the endorsement of the research sponsor.

#### REFERENCES

- America's Health Insurance Plans (AHIP). 2004. Long-Term Care Insurance in 2002 [accessed on February 11, 2005]. Available at: http://www.ahipresearch.org/pdfs/ 18\_LTC2002.pdf
- American Association of Retired Persons (AARP). 2001. "The Costs of Long-Term Care: Public Perceptions versus Reality." [accessed on April 4, 2007.] Available at http://assets.aarp.org/rgcenter/health/ltc\_costs.pdf.
- Belden, Russonello and Stewart. 2001. "In the Middle: A Report on Multicultural Boomers Coping With Family and Aging Issues." National Survey conducted for AARP by

Research/Strategy/Management [accessed on August 26, 2005]. Available at: http://assets.aarp.org/rgcenter/il/in\_the\_middle.pdf.

- Brown, J. and A.Goolsbee. 2002. "Does the Internet Make Markets More Competitive?" *Journal of Political Economy* 110: 481-507.
- Consumer Reports. 2003. "Long-term Care Insurance: Should You Buy?" accessed on February 11, 2005]. Available at: http://www.consumerreports.org/main/content/ display\_report.jsp? FOLDER%3C%3Efolder\_id=344787&ASSORTMENT%3C%3East\_id=333147&bmUI
- D=1108757501241. Duhigg, Charles (2007, March 26). "Aged, Frail and Denied Care by Their Insurers." *New York Times*, p. A1.
- Finkelstein, Amy, Kathleen McGarry, and Amir Sufi. 2005. "Dynamic Inefficiencies in Insurance Markets: Evidence From Long-Term Care Insurance." *National Bureau of Economic Research, Inc.* NBER Working Paper 11039. [accessed on August 30, 2005]. Available at: http://www.nber.org/papers/w11039.
- Feldman, Amy, Leslie Haggen, Laura Lallos, Teresa Tritch, Walter Updegrave, and Penelope Wang. 2000. CNNmoney 29 (7) [accessed on August 30, 2005]. Available at: http://premium.money.cnn.com/pr/subs/magazine archive/2000/07/QUEB.html.
- Gendel, Murray. 1998. "Trends in Retirement Age in Four Countries, 1965 95." *Monthly Labor Review* 121(8): 20 30. [accessed on April 4, 2007]. Available at: http://www.bls.gov/opub/mlr/1998/08/art2full.pdf.
- Gupta, Aparna and Lepeng Li. 2004. "A Modeling Framework for Optimal Long-Term Care Insurance Purchase Decisions in Retirement Planning." *Health Care Management Science* 7: 105-17.
- Health and Retirement Study. 1992-2004. Public use dataset produced and distributed by the University of Michigan with funding from the National Institute on Aging. Ann Arbor, MI [accessed on August 30, 2005]. Available at: http://hrsonline.isr.umich.edu.
- Kenkel, Donald S. 1991. "Health Behavior, Health Knowledge, and Schooling." Journal of Political Economy 99: 287 – 305.
- Leibowitz, A. and M. Chernew. 1992. "The Firm's Demand for Health Insurance." *Health Benefits and the Workforce*. U.S. Department of Labor. Washington, D.C.: U.S. Government Printing Office.
- McCall, Nelda, Steven Mangle, Ellen Bauer, and James Knickman. 1998. "Factors Important in the Purchase of Partnership Long-Term Care Insurance." *Health Services Research* 33: 187-204.
- McKenna, C. J. 1986. The Economics of Uncertainty. Brighton, Sussex: Wheatsheaf.
- McNamara, Paul E. and Nayoung Lee. 2004. "Long-term Care Insurance Policy Dropping in the U.S. from 1996 to 2000: Evidence and Implications for Long-term Care Financing." *The Geneva Papers on Risk and Insurance* 29 (4): 640-651.
- Morisato, Susan. 2004. "Growing Pains: Industry at the Crossroads" panel discussion. 17<sup>th</sup> Private Long Term Care Insurance Conference, June 23-25, 2004.
- National Center for Health Statistics. 2006. *Health, United States, 2006 With Chartbook on Trends in the Health of Americans.* Hyattsville, MD: 176, 378. [accessed on 4/3/2007.] Available at http://www.cdc.gov/nchs/data/hus/hus06.pdf#027.

- Norton, Edward C. and Virendar Kumar. 2000. "The Long-Run Effect of the Medicare Catastrophic Coverage Act." *Inquiry Blue Cross and Blue Shield Association* 37 (2): 173-87.
- O'Brien, Dennis. 2004. "Pricing Issues in Individual LTC Insurance" panel discussion. 17<sup>th</sup> Private Long Term Care Insurance Conference, June 23-25, 2004.
- Pandya, Sheel. 2005. "Racial and Ethnic Differences Among Older Adults in Long-Term Care Service Use." AARP Public Policy Institute Research Report. [accessed on August 26, 2005]. Available at: http://www.aarp.org/research/longtermcare/trends/fs119\_ ltc.html.
- Pauly, Mark V. 1990. "The Rational Nonpurchase of Long-Term-Care Insurance." Journal of Political Economy 98: 153-168.
- Pauly, Mark V., Kate Withers, H. Krupa Subramanian-Viswanathan, Jean Lemaire, John C. Hershey, Katrina Armstrong, and David A. Asch. 2003. "Price Elasticity of Demand for Term Life Insurance and Adverse Selection." *National Bureau of Economic Research, Inc.* NBER Working Paper 9925. [accessed on February 11, 2005]. Available at: http://www.nber.org/papers/w9925.
- Sahadi, Jeanne. 2006. "Retirement: You'll Need Far More Than You Think." *CNNMoney.com*. [accessed April 4, 2007.] Available at http://money.cnn.com/2006/09/21/pf/retirement/new\_calcs\_savings/index.htm?postversion=2006092214.
- Sloan, Frank A. and Edward C. Norton. 1997. "Adverse Selection, Bequests, Crowding Out, and Private Demand for Insurance: Evidence from the Long-term Care Insurance Market," *Journal of Risk and Uncertainty* 15 (3): 201-19.
- State of Michigan, Office of Financial and Insurance Services. 1996 1998. Rate filings. Lansing, MI.
- VanDerhei, Jack. 2006. "Measuring Retirement Income Adequacy: Calculating Realistic Income Replacement Rates." *Employee Benefit Research Institute Issue Brief* 297. [accessed April 4, 2007.] Available at http://www.ebri.org/pdf/briefspdf/EBRI\_IB\_09-20061.pdf.
- Wallace, S.P., L. Levy-Storms, R.S. Kingston, and R.A. Anderson. 1998. "The Persistence of Race and Ethnicity in the Use of Long-Term Care. *Journal of Gerontology B* 53: S104-12.
- Zweifel, Peter and Wolfram Strüwe. 1996. "Long-Term Care Insurance and Bequests as Instruments for Shaping Intergenerational Relationships." *Journal of Risk and Uncertainty* 12 (1): 65-76.
- Zweifel, Peter and Wolfram Strüwe. 1998. "Long-Term Care Insurance in a Two-Generation Mode." *Journal of Risk and Insurance* 65 (1): 33-56.

Chapter 43

# TOWARD A COMPARISON-INDUCED DISTORTION THEORY OF JUDGMENT AND DECISION MAKING\*

# Jessica M. Choplin<sup> $\dagger$ </sup>

DePaul University, Chicago, Illinois, USA

### ABSTRACT

This chapter demonstrates how biases produced by verbal comparisons (Choplin & Hummel, 2002) might produce a variety of phenomena in the psychology of judgment and decision making. The biases produced by verbal comparisons cause people to overestimate the importance of small differences and underestimate the importance of large differences. Simulations will demonstrate that overestimating the importance of small differences and underestimating the importance of large differences from a reference point (default value, status quo, etc.) would produce s-shaped evaluation functions. Simulations will also demonstrate that overestimating the importance of small differences and underestimating the importance of large differences from other contextual stimuli would produce distribution-density effects (often called frequency effects). Because large differences are underestimated, when the variance in people's unbiased estimates is large as in anchoring effects, these biases will often look like assimilation effects. Biases produced by verbalized social comparisons would also overestimate the importance of small differences and underestimate the importance of large differences so that moderate downward comparisons will produce higher evaluations of the self than will extreme downward comparisons and moderate upward comparisons will produce lower evaluations of the self than will extreme upward comparisons. Comparison strategies might also help explain decoy effects such as asymmetric dominance effects and phantom decoy effects. Testable and falsifiable assumptions of this model are described thereby laying a foundation for future empirical research.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Research supported by a grant from the National Science Foundation (NSF-0621664) awarded to Jessica M. Choplin. Correspondence to: Jessica M. Choplin; Department of Psychology; DePaul University; 2219 North Kenmore Avenue; Chicago, IL 60614-3504; jchoplin@depaul.edu

#### INTRODUCTION

Hedonic and other types of attribute evaluations have wide ranging implications for decision-making and the consequences of those decisions for well-being. Evaluations of price and other consumer-product attributes affect the decisions people make as consumers and, in turn, how satisfied they are with the products they buy (e.g., Choplin & Hummel, 2002; Choplin & Hummel, 2005; Huber, Payne, & Puto, 1982; Wedell & Pettibone, 1996). Evaluations of people's own personal attributes—often informed by social comparisons (e.g., Festinger, 1954; Lavine, Sweeney, & Wagner, 1999; Morse & Gergen, 1970; Mussweiler, 2003; Wedell & Parducci, 2000)—affect how people see themselves on any personal attribute one might imagine. Evaluations of food intake (taste, calories, portion size, Riskey, Parducci, & Beauchamp, 1979; Schifferstein & Frijters, 1992; Wansink, 2004) affect people's dietary decisions and all of the consequences of those decisions. People's subjective evaluations of their own well-being (i.e., how happy they see themselves, Kahneman, 1999; Parducci, 1995; Schwarz & Strack, 1999) often affect actual well-being.

In this chapter, I present an introduction to a model of attribute evaluation, comprehension, memory, and estimation wherein verbal, language-expressible comparisons of attribute values systematically bias how people evaluate, comprehend, remember, and estimate the compared attribute values (comparison-induced distortion theory, CID theory, Choplin & Hummel, 2002, 2005). Judging one product to be more expensive than a second product, for example, might bias people's judgments of how affordable those products are. I will present simulation data that suggests that comparison-induced biases might create a variety of phenomena in the psychology of judgment and decision making and provide key insights into yet other phenomena. My goal in this chapter is merely to introduce readers to the basic tenants of CID theory. I will use simulation data to demonstrate that biases produced by verbal comparisons are sufficient to explain these phenomena and to describe some key novel predictions that can be tested in future research. It is not my focus in this basic introduction to argue that it is necessary to appeal to comparison-induced biases to explain these phenomena nor is it my focus to argue that CID theory provides a better account of these phenomena than other theories provide. Arguments and empirical results demonstrating that CID theory provides a better account than other accounts are presented elsewhere (Choplin & Hummel, 2005, and in papers currently in process or preparation, and ongoing empirical research). Furthermore, the fact that CID theory can account for such a large variety of phenomena (a larger variety of phenomena than any of its competitors) is itself strong evidence in its favor.

#### **COMPARISON-INDUCED DISTORTION THEORY**

Comparison-induced distortion theory (CID theory, Choplin & Hummel, 2002) is a theory of attribute evaluation in which language-expressible magnitude comparisons (e.g., "I am older than Susan is") systematically bias how people evaluate, comprehend, remember, and estimate attribute values. The basic idea behind CID theory is that language-expressible magnitude comparisons suggest quantitative values. For example, to investigate the meanings of English age comparisons Rusiecki (1985) gave his participants sentences such as "Mary is

older than Jane" and "Martin's wife is older than Ken's wife" and asked them to report the ages they imagined. Rusiecki found considerable agreement in the values imagined by his participants. In response to the comparison "Mary is older than Jane" participants imagined Mary to be 20.2 years old on average and Jane to be 17.9 years old on average. In response to the comparison "Martin's wife is older than Ken's wife" participants imagined Martin's wife to be 37.2 years old on average and Ken's wife to be 33.0 years old on average.

Of particular interest to the current discussion, the age differences imagined by Rusiecki's (1985) participants were remarkably similar. Regardless of the ages they imagined, participants imagined a difference between the ages of approximately 2 to 5 years (slightly larger for larger values; more on this topic below)—not 1 month or 30 years. Inspired by these results, Rusiecki argued that comparisons suggest quantitative differences between compared values. I will henceforth call these quantitative differences "comparison-suggested differences," because they are the differences suggested by comparisons. In the case of age comparisons, Rusiecki's results demonstrate that comparison-suggested differences are approximately 2 to 5 years. For ease of discussion I will operationally define the difference suggested by age comparisons to be 3.5 years in the discussion that follows. Please note, however, that the actual size of the difference likely depends upon many factors as I discuss below and that empirical measurements of the comparison-suggested difference will have to be cognizant of these factors.

Choplin and Hummel (2002) proposed that language-expressible magnitude comparisons (like those investigated by Rusiecki, 1985) might bias evaluations toward the quantitative values suggested by comparisons. For example, if the actual age difference between two people were 1.5 years (i.e., less than the comparison-suggested difference of 3.5 years), then a comparison might bias evaluations of their ages apart—toward a difference of 3.5 years. The younger person might be evaluated younger than she or he would have been evaluated otherwise and the older person might be evaluated older than she or he would have been evaluated otherwise. If the actual age difference between two people were 5.5 years (i.e., more than 3.5 years), then a comparison might bias evaluations of their ages together—again toward a difference of 3.5 years. The younger person might be evaluated older person might be evaluated older and the older person might be evaluated otherwise of the actual age difference between two people were 5.5 years (i.e., more than 3.5 years), then a comparison might bias evaluations of their ages together—again toward a difference of 3.5 years. The younger person might be evaluated older and the older person might be evaluated older and the older person might be evaluated younger.

Formally, the comparison-suggested value of the smaller of two compared values ( $E_S$ ; E for Expected) and the comparison-suggested value of the larger of two compared values ( $E_L$ ) can be calculated from Equations 1a and 1b respectively:

$$E_S = S_L - D \qquad (1a) \qquad \qquad E_L = S_S + D \qquad (1b)$$

where  $S_L$  and  $S_S$  (S for Stimulus values) are the values of the larger and smaller values unbiased by comparisons respectively and D is a parameter representing the comparisonsuggested difference (Choplin & Hummel, 2005). Participants probably do not calculate these comparison-suggested values on the fly. Rather, these values likely come from their memory of previous times in which they judged one value to be larger or smaller than another value (e.g., memory of the age of previous people they had judged to be older than 33.0 years old).

In the tradition of previous judgment models (Anderson, 1965; Huttenlocher, Hedges, & Vevea, 2000), CID theory is a weighted average model in that represented values are assumed to be a weighted mean of values unbiased by comparisons and comparison-suggested values:

$$R_{S} = wE_{S} + (1-w)S_{S}$$
 (2a)  $R_{L} = wE_{L} + (1-w)S_{L}$  (2b)

where w is the relative weights of the two values, is bound between 0 and 1, and is constrained so as to prevent impossible values (e.g., negative years or sizes of geometric figures) from being represented. For example, assuming a comparison-suggested difference, D, of 3.5 years (an oversimplification used here for demonstration purposes only; see discussion below on measuring comparison-suggested differences), a comparison between a 22-year old and a 28-year old would bias evaluations of their ages toward each other. If the weight given to comparison-suggested values were 0.2 (I use this value for demonstration purposes only; to model real data this value would be found by fitting the model to empirical data), then the represented age of the 22-year old would be 22.5 years and the represented age of the 28-year old would be 27.5 years. That is, the 22-year old would be evaluated, i.e., treated, half a year older and the 28-year old would be evaluated half a year younger.

Analogous to previous weighted-average judgment models that weight information from various sources (e.g., Huttenlocher et al., 2000), Parameter w represents the degree to which people rely upon comparisons to evaluate attribute values. If people have an accurate understanding of how large or small or good or bad attribute values are and they are completely certain of their own understanding, then they will have little need to use comparison information to evaluate attribute values and Parameter w would take a low value. By contrast, if they do not have an accurate understanding and need to evaluate attribute values relative to other contextual attribute values, then they will need to rely upon comparison information and Parameter w would take a high value. In cases where people need to recall attribute values from memory, if they remember exact values, they will have little need to rely upon comparison information to aid their recall and Parameter w would take a low value. If they do not remember exact values, they might use verbal comparison information to aid their recall and Parameter w would take a high value. Many cognitive processes could potentially produce this averaging and like other weighted-average models CID theory does not need to commit itself to any one specific cognitive process.

Sometimes people might hesitate to describe one value as larger (or smaller) than another value even if it is larger (or smaller). It might seem odd, for example, to describe a person who is 28 years and 4 months old as "older" than a person who is 28 years and 2 months old. People might prefer to describe these ages as "approximately the same" or "similar." These comparisons suggest that there is little to no difference between the compared values and so biases in evaluation created by these comparisons can be modeled by setting parameter D in Equations 1a and 1b to zero. If Parameter w in Equations 2a and 2b were set to .2, the comparison "a 28 year and 4 month old person is approximately the same age as a 28 year and 2 month old person" would bias the evaluation of 28 years and 4 months to be about 12 days younger and the evaluation of 28 years and 2 months to be about 12 days older.

The decision to describe a given difference as "approximately the same" or "similar" can be modeled stochastically using Shepard's (1987) law of generalization. That is, the probability of describing a difference between two stimuli (i.e.,  $S_L$  and  $S_S$ ) as approximately the same or similar can be modeled as:

$$p("same" | S_L - S_S) = e^{-c(S_L - S_S)}$$
(3)

where c is a sensitivity parameter.

Equations 1a through 3 employ three parameters (D, w, and c). Note that although the theory uses these three parameters, Parameters D and c can usually be measured empirically so these two parameters will generally not be free. The ability to measure Parameters D and c empirically leaves only Parameter w as a free parameter in CID theory-the same number of free parameters as some of its competitors such as adaptation-level theory (AL theory, Helson, 1964) and range-frequency theory (RF theory, Parducci, 1965, 1995) and fewer free parameters than other competitors (Frederick & Loewenstein, 1999; Stevens, 1961). Parameter D in Equations 1a and 1b can be measured by asking a control group of participants about the types of differences they imagine in response to a comparison (as Rusiecki, 1985, did) or by looking at real-world differences under the assumption that participants have been exposed to these differences and that their experiences with these differences have shaped their understanding of the differences implied by these comparisons. Parameter c in Equation 3 which reflects the probability that people will judge values to be "approximately the same" can be measured by asking a control group of participants whether or not they would describe a given difference as "approximately the same." I discuss the ways one might measure these values in the next sections.

## **MEASURING COMPARISON-SUGGESTED DIFFERENCES**

As described above, comparison-suggested differences (Parameter D in Equations 1a and 1b) and values ( $E_S$  and  $E_L$  in Equations 1a and 1b) can be measured empirically by asking control groups of participants to imagine that one value is more than, less than, larger than, smaller than, better than, worse than, etc. another value and then ask them about the differences they imagine. My students, colleagues, and I have measured comparison-suggested differences in this manner many times.

For example, in one study my students and I measured comparison-suggested differences and values in personal attributes (grade point average, height, weight, income, dates per month, and commute to campus) by asking a control group of participants about the differences they imagined. In particular, we described a fictional DePaul University undergraduate student (Jennifer for women; Brad for men; the personal attributes of these two people were collected from yet another control group of participants who imagined the average female or male undergraduate DePaul University student) and asked the control group to imagine the personal attributes of another DePaul University undergraduate student (Michelle for women; Michael for men). Michelle was better on every personal attribute than Jennifer (higher grade point average, taller, weighed less, earned more, went out on dates more often, and lived closer to campus); and Michael was better on every personal attribute than Brad (higher grade point average, taller, more muscular and so weighed more, earned more, went out on dates more often, and lived closer to campus). Jennifer and Brad's personal attributes as well as the median imagined personal attributes of Michelle and Michael are presented in Table 1. The median imagined personal attributes of Michelle and Michael represent comparison-suggested values. Comparison-suggested differences were calculated from these comparison-suggested values. Perhaps unsurprisingly the attributes of average female and male undergraduate DePaul University students imagined by these pretest participants fit gender stereotypes. The average female student imagined by women (Jennifer) earned a higher grade point average, was shorter, weighed less, earned much less money ( $65\phi$  on the dollar), went out on more dates, and lived closer to campus than did the average male student imagined by men (Brad). We will use these values later in the chapter when we discuss social comparisons.

	Jennifer's personal attributes (collected in another pretest)	Michelle's median imagined personal attributes (comparison-suggested values)	Comparison- suggested differences
g.p.a. Height Weight Income Dates	3.17 64 inches 134 pounds \$7,868.00 3.1	3.5 66 inches 125 pounds \$9,000.00 5	0.33 2 inches 9 pounds \$1,132.00 1.9
Commute	9.2 miles	5 miles	4.2 miles
	Brad's personal attributes (collected in another pretest)	Michael's median imagined personal attributes	
g.p.a.	3.1	3.38	0.28
Height	70 inches	72 inches	2 inches
Weight	173 pounds	190 pounds	17 pounds
Income	\$12,084.00	\$15,000	\$2,916.00
Dates	2.45	4	1.55
Commute	18.9 miles	10 miles	8.9 miles

The sizes of comparison-suggested differences depend upon many factors. For one, the size of the comparison-suggested difference often depends upon the size of the base of the comparison. For example, Choplin and Hummel (2005) studied comparison-induced distortions in judgments of line length. To measure comparison-suggested values and differences in judgments of line length, John Hummel and I showed a control group of participants lines of various lengths (each control participant only saw one line) and then asked them to imagine a line that was longer or shorter than that line and draw it. Looking first at participants who drew longer lines than the lines they saw: Of those who viewed a line that was 10.0 mm long, the median redrawn longer line was 36.3mm (a difference of 26.3mm). Of those who viewed a 14.0 mm-line, the median redrawn line was 42.5mm (a difference of 28.5mm). Of those who viewed a 22.0 mm-line, the median redrawn line was 53.0 mm (a difference of 31.0 mm); and of those who viewed a 30.0 mm-line, the median redrawn line was 60.0 mm (a difference of 30.0 mm). Next looking at participants who drew shorter lines than the lines they saw: Of those who viewed a line that was 34.0 mm long, the median redrawn shorter line was 18.0 mm (a difference of 16.0 mm). Of those who viewed a 30.0 mm-line, the median redrawn line was 15.0 mm (a difference of 15.0 mm). Of those who viewed a 22.0 mm-line, the median redrawn line was 9.0 mm (a difference of 13.0 mm); and of those who viewed a 14.0 mm-line, the median redrawn line was 6.0 mm (a difference of 8.0 mm). In this case, the sizes of the comparison-suggested differences depended upon the size of the base of the comparison such that as a general trend comparison-suggested differences were larger when the base of the comparison was larger.

Comparison-suggested differences might not always be larger when the base of the comparison is larger. Rather, the size of the comparison-suggested difference might depend upon the distribution of values out in the world. The comparison "a man taller than 5'8"," for example, might suggest a larger difference than would the comparison "a woman taller than 5'8"," because the distribution of men's heights is larger than the distribution of women's heights. Also if people have some idea of the amount of variance in a category, comparison-suggested differences might be larger when variance is larger. The variance in the high school g.p.a.s of all high school graduates, for example, might be larger than the variance in the high school g.p.a. of DePaul University undergraduate students are a self-selecting group and the self-selection process involves high school g.p.a. among other factors. Comparison-suggested differences would then likely be larger if one were comparing the high school g.p.a.s of two DePaul University undergraduate students.

Furthermore, not all ways of wording comparisons will suggest the same quantitative difference. Wording a comparison "slightly larger," for example, will suggest a smaller quantitative difference than will wording it "larger;" and wording a comparison "much larger" will suggest a larger quantitative difference than will wording it "larger." The fact that alternative ways of wording comparisons will suggest smaller or larger quantitative differences gives us a method of isolating the evaluation biases created by verbal comparisons from the evaluation biases produced by some other factors. Once a researcher has measured the quantitative differences suggested by alternative ways of wording verbal comparisons among control participants, an experimenter will be able to manipulate the quantitative differences suggested to experimental participants by manipulating how a comparison is worded while keeping all other factors constant. Once all other factors have been held constant, any observed effects would have to be due to the verbal comparisons that were used. Note that special care would be needed to hold emotional factors constant as emotional factors could also potentially be affected by different ways of wording comparisons. Emotional factors might be held constant by testing for the effects of alternative ways of wording comparisons on dimensions that have no hedonic consequences such as line length or size of geometric shapes.

#### MEASURING LIKELIHOOD OF JUDGING TWO VALUES THE SAME

The likelihood that people will describe two values as approximately the same can be modeled using Shepard's (1987) law of generalization (shown in Equation 3). The law of generalization has one parameter (Parameter c). A value for Parameter c can be measured empirically by asking control groups of participants how they would describe differences (e.g., as "same"—or perhaps "approximately the same" or "same ballpark"—versus "different," "larger than," or "smaller than"), fitting Equation 3 to these results using a root mean squared error (RMSE) criterion, and thereby finding the best-fit value for Parameter c. Empirically measuring the value of Parameter c for control participants allows CID theory to make a priori predictions about how experimental participants will evaluate attributes.

My students and I measured the likelihood that people will judge values to be the same and found the best-fitting value for Parameter c as described above in an experiment in which they judged whether prices for an all-you-can-eat lunch were the same as or different than \$6.53. In particular, participants imagined that they worked for a summer camp that had been charging \$6.53 for an all-you-can-eat lunch. For some of the participants, the camp was considering raising the price and they were asked whether they would consider each of the prices \$6.62, \$6.71, \$6.80, \$6.89, \$6.98, \$7.07, and \$7.16 to be the same as or different than \$6.53. Half of these participants simply judged whether the prices were "the same as" or "different than" \$6.53. The other half of these participants judged whether the prices were in "the same ballpark as" or "a completely different ballpark than" \$6.53. For other participants the camp was considering lowering the price and they were asked whether they would consider each of the prices \$6.44, \$6.35, \$6.26, \$6.17, \$6.08, \$5.99, and \$5.90 to be the same as or different than \$6.53. The proportions of participants describing each price as the same as \$6.53 are plotted in Figure 1. Shepard's (1987) law of generalization (Equation 3) was fit to the results using a root mean squared error criterion. Best-fits are also plotted in Figure 1. In the condition wherein participants simply judged whether the prices were "the same as" or "different than" \$6.53, the best-fitting value for Parameter c for raising prices was 3.33 (RMSE = 0.05) and the best-fitting value for lowering prices was 2.46 (RMSE = 0.12). In the condition wherein participants s judged whether the prices were in "the same ballpark as" or "a completely different ballpark than" \$6.53, the best-fitting value for Parameter c for raising prices was 1.77 (RMSE = 0.09) and the best-fitting value for lowering prices was 1.58(RMSE = 0.09).

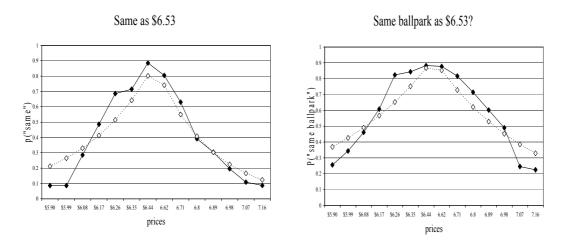


Figure 1. Probability of describing prices as the same as \$6.53. Black diamonds represent the proportion of participants who described each price as the same as \$6.53. White diamonds represent the best-fitting values when Equation 3 was fit to participants' responses.

Once comparison-suggested differences (Parameter D in Equations 1a and 1b) and likelihood of judging two values approximately the same (Parameter c in Equation 3) have been empirically measured, only the weighting of the information provided by comparisons (Parameter w in Equation 2a and 2b) will remain as a free parameter. Additionally, as described below CID theory will only be able to explain certain phenomena, if Parameters D

and c take certain values. If CID theory is to explain these phenomena, then, it will have to predict that these parameters will take these values and these predictions can then be empirically verified (or else the CID theory account of these phenomena would be empirically falsified).

Quantitative modeling (discussed shortly) demonstrates that biases in evaluation created by verbal comparisons might play a key role in a number of well-known phenomena in the psychology of judgment and decision making. In particular, biases in evaluation created by verbal comparisons might play a key role in producing s-shaped evaluation functions, distribution-density effects, anchoring effects, social comparisons, and decoy effects. Although I will not be pitting CID theory against alternative theories here in this chapter, note that to compete with CID theory any alternative theory would have to provide a more parsimonious account of all of these phenomena combined, not just a more parsimonious account of one phenomenon. The fact that CID theory can account for such a large variety of phenomena is itself strong evidence in favor of CID theory. The remainder of this chapter will describe the role that biases produced by verbal comparisons might play in these phenomena. To lay the ground work for future empirical research, empirically testable predictions of the CID theory account of these phenomena will also be described.

## **S-Shaped Evaluation Functions**

Previous research has found considerable evidence that evaluations often follow s-shaped functions (see Figure 2, Frederick & Loewenstein, 1999; Kahneman & Tversky, 1979). Following Helson's (1964) AL theory, these evaluation functions predict that evaluations are made relative to a single reference point (RP), i.e., the point at which people consider values "normal" or "average." This function differs from Helson's original formalization of AL theory, however, in that it has an s-shape. Evaluations are concave (downward) for positive changes from the RP and convex (concave upward) for negative changes from the RP. Helson assumed that evaluations are linearly transformed around the RP.

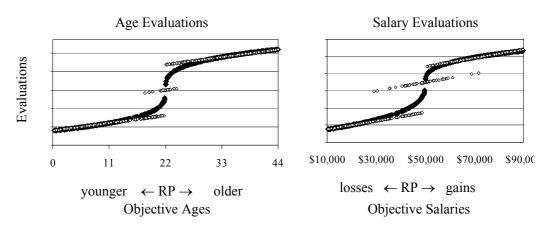


Figure 2. Evaluations made relative to reference points (RPs; 22 years and \$50,000). Evaluation functions are generally s-shaped. They are formalized here using Stevens' (1961) Power Law (black s-shaped functions) and CID theory (speckled lines).

To see that evaluations are often s-shaped, consider how evaluations of age change over one's lifespan. As a child (e.g., the RP might be 6 years), all adults—and even teenagers seem old. The portion of the curve representing adults and teenagers (i.e., the upper portion of the s-shape) is relatively flat. As a young adult of 22 years (i.e., the RP is 22 years; see Figure 2 Age Evaluations), children all seem young, mature adults all seem old, and the difference between 18-year olds and 25-year olds is huge. The portion of the curve representing children (i.e., the lower portion of the s-shape) and the portion of the curve representing mature adults (i.e., the upper portion of the s-shape) are both relatively flat, but the portion of the curve representing the difference between 18-year olds (just to the left of the RP) and 25-year olds (just to the right of the RP) is the steepest portion of the curve. As an adult of retirement age (e.g., the RP might be 65 years), the differences between children, teenagers, and young adults all seem insignificant, young parents appear to be kids having kids, and the lower portion of curve representing these ages is relatively flat.

Since this s-shape is generally thought to reflect the psychophysical law that sensitivity to differences decreases at greater magnitudes, it is often formalized as plotted in Figure 2 solid black lines using Stevens' (1961) Power Law (Kahneman & Tversky, 1979):

$$Evaluation(X) = \begin{cases} a_{more}(X - RP)^{b_{more}}, & \text{if } X > RP \\ -a_{less}(RP - X)^{b_{less}}, & \text{otherwise} \end{cases}$$
(4)

where X is the value being evaluated, RP is the reference point,  $a_{more}$  scales sensitivity to positive differences from adaptation,  $a_{less}$  scales sensitivity to negative differences from adaptation,  $b_{more}$  scales the curvature of the portion of the function representing positive differences, and  $b_{less}$  scales the curvature of the portion of the function representing negative differences. To produce s-shaped evaluation functions like those in Figure 2, the parameters  $b_{more}$  and  $b_{less}$  will take values greater than zero and less than one. They will be equal to one, if evaluations are linear (see Briesch, Krishnamurthi, Mazumdar, & Raj, 1997). They will be greater than 1 only in rare cases (e.g., perhaps evaluations of electric shocks, Stevens, 1962).

CID theory provides an alternative way of formalizing the s-shaped evaluation function than does Stevens' Law. The top-most lines in Figure 2 represent biases created by more-than comparisons simulated using Equations 1a through 2b. For example, the comparison "a 35year old is older than a 22-year old" would produce the evaluation of the 35-year old shown in Figure 2 Age Evaluations. The middle lines represent biases created by approximately-thesame comparisons. The likelihood that people will describe values as approximately the same to the RP is simulated using Equation 3 and the biases produced by these approximately-thesame comparisons are simulated using Equations 1a through 2b with Parameter D set at 0 (since the comparison-suggested difference of describing two values as approximately the same would be zero difference). Unlike Stevens' Law, CID theory can capture indifference around the RP (Kalyanaram & Little, 1994), if people judge a wider range of values as "approximately the same as" the RP. The bottom lines represent biases created by less-than comparisons.

Continuous s-shaped evaluation functions similar to the Stevens' Law evaluation function would be produced by averaging across the three types of comparisons. When values are very close to the RP almost all participants will describe differences as the same, but the tendency to do so usually tapers off quickly for values that are farther from the RP (see Figure 1). It tapers off at a rate that usually produces a function that, once biases from all three types of comparisons are averaged together, is very similar to (but not identical to) the function produced by Stevens' Law.

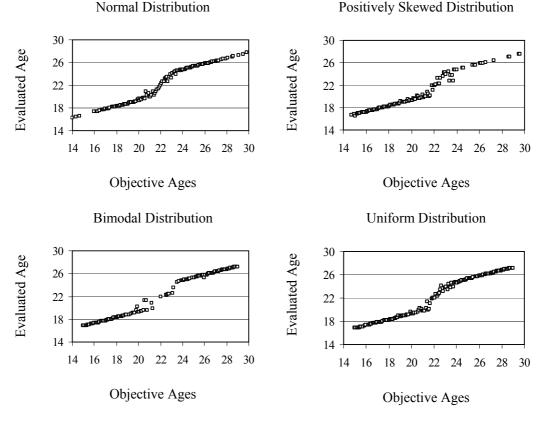


Figure 3. Simulated comparison-biased age evaluations—All ages are compared to 22 years.

To demonstrate what the evaluation function looks like when evaluations for less than, approximately the same, and more than comparisons are averaged together, I again simulated age evaluations from the perspective of a 22-year old who compares all ages to her or his own age (see Figure 3). Five hundred ages were randomly drawn from each of a normal, a positively skewed, a bimodal, and a uniform distribution. For demonstration purposes only, Parameter D in Equations 1a and 1b was set to 3.5 years, Parameter c in Equation 3 was set to 1.5 (before modeling empirical data, these two values would have to be empirically measured), and Parameter w in Equations 2a and 2b was set to .5 (to model empirical data, this value would be found by fitting this free parameter to data).

At c set to 1.5, the age 22 years and 3 months was judged to be the same age as 22 years 69% of the time. The comparison "22 years and 3 months is approximately the same age as 22 years" biased the evaluation of 22 years and 3 months to be approximately 22 years and 1.5 months. The other 31% of the time 22 years and 3 months was judged older than 22 years. The comparison "22 years and 3 months is older than 22 years" biased the evaluation of 22 years and 3 months was judged older than 22 years. The comparison "22 years and 3 months is older than 22 years" biased the evaluation of 22 years and 3 months to be approximately 23 years and 10.5 months. Averaging across these two types of comparisons, 22 years and 3 months was evaluated as approximately 22 years

and 4.3 months. An actual difference of 3 months was then evaluated as if it were (treated as if it were) 4.3 months. This extra sensitivity close to the RP can be seen in Figure 3 by the steeper slope around 22 years (the immediate region around the RP is not always steep, sometimes there is a region in which values are judged to be the same, see Kalyanaram & Little, 1994). Sensitivity quickly decreased for values that were farther away from the RP. An age of 24 years was evaluated as if it were 24 years and 10 months and an age of 26 years was evaluated as if it were 25 years and 10 months making a difference of 2 years seem like 1 year.

Notice also from Figure 3 that CID theory predicts that evaluations will be s-shaped and unaffected by the type of distribution from which values are drawn, if all values are verbally compared to a central RP and compared to no other contextual values. Simulation results presented in the next section demonstrate that evaluation functions will depend upon the type of distribution from which values are drawn, if values are compared to other contextual values.

A well-known phenomenon associated with s-shaped evaluation functions is that people's aversion to losses is often greater than their affinity for gains (Kahneman & Tversky, 1979). Consider evaluations of income (Figure 2 Salary Evaluations). For someone who earns \$50,000 a year, the subjective difference between \$40,000 and \$50,000 (a difference that represents a loss) is large, likely larger than the subjective difference between \$50,000 and \$60,000. Reflecting this fact, the evaluation function for values to the left of the RP in Figure 2 Salary Evaluations is steeper than the evaluation function for values to the right of the RP. This pattern would reverse, if large values represented loss (e.g., being charged more).

Previous research has shown that loss aversion is due to affective factors. These affective factors might affect how people use language to describe differences. For example, the comparison words they use to describe losses (e.g., loss, less, deficit, shortfall, or shortage) might suggest larger differences than do the comparison words they use to describe gains (e.g., gain, more, surplus, or excess). They also might be less likely to describe losses as approximately the same as RP values than to describe gains as approximately the same as RP values than to describe gains as approximately the same as RP values that the biases produced by verbal comparisons would not only be s-shaped, but would also reflect the fact that people are loss averse. This model of the s-shaped evaluation function assumes that verbal comparisons mediate the effects of affective factors on evaluations. That is, this model assumes that affective factors play a role in evaluations by changing comparison words; verbalized comparisons in turn bias evaluations.

The speckled lines in Figure 2 Salary Evaluations demonstrate the point that biases produced by verbal comparisons would reflect loss aversion by plotting what the comparisonbiased evaluation function would look like, if—because of loss aversion—comparisonsuggested differences (parameter D in Equation 1) were larger for losses than for gains. As seen in Figure 2 Salary Evaluations, biases produced by verbal comparisons nicely capture the asymmetry in the evaluation function caused by loss aversion. To mathematically demonstrate this point, comparisons to an RP were simulated with Parameter D set higher for losses than for gains. Sensitivity to losses and gains was assessed by fitting Equation 4 to the simulation results using a minimized root mean squared error criterion. Demonstrating increased sensitivity to losses, the best-fitting value for Parameter a in Equation 4 was greater for losses than for gains. Not plotted in Figure 2, a tendency to avoid describing losses as "approximately the same" (Parameter c) would also nicely capture the asymmetry in evaluations of losses and gains. To investigate the effects of asymmetries in Parameter c on sensitivity to losses and gains, comparisons to an RP were simulated with Parameter c set higher for losses than for gains. Again, sensitivity to losses and gains was assessed by fitting Equation 4 to the simulation results using a minimized root mean squared error criterion; and again the best-fitting value for Parameter a in Equation 4 was greater for losses than for gains.

Importantly, once Parameters D and c are measured empirically as described above, the CID theory account of the s-shaped evaluation function has fewer free parameters than does Steven's Law. The Steven's Law account has 4 free parameters ( $a_{more}$ ,  $a_{less}$ ,  $b_{more}$ , and  $b_{less}$ ), while the CID theory account only has one (w, or two, if w takes different values for losses than for gains). Not only can Parameters D and c be measured empirically, but also to explain some phenomena CID theory will have to predict that Parameters D and c will take certain values. These predictions can then be tested empirically. These parameters, therefore, represent testable predictions of the model rather than free parameters. For example, the CID theory account of the s-shaped evaluation function predicts that either comparison-suggested differences will be larger for losses than for gains or people will be less likely to describe losses as "approximately the same" than to describe gains that way, or both. If people do neither of these things, then CID theory would be unable to capture loss aversion or explain the s-shaped evaluation function.

Of course, finding the predicted correlations between how people use language and biases in their judgments and choice behavior would not in itself establish that these biases are caused by verbal comparisons (as CID theory maintains) as both could be caused by third variables (i.e., affective factors, psychophysical factors, or both) or the causal direction could be reversed. To establish, that verbal comparisons cause biases in judgment, one can manipulate these parameters. Parameter D might be manipulated, for example, by tagging the adjectives "slightly" or "much" on to comparisons; Parameter c might be manipulated by describing sameness as "same," "approximately the same," or "in the same ballpark." To control for affective factors, one might study evaluations of non-hedonic dimensions such as line lengths or sizes of geometric shapes.

Previous research has found that evaluation functions are not always s-shaped. Rather, evaluation functions will often depend upon the distribution of contextual attribute values. Typically, for example, (if the range of values is held constant) an attribute value that is drawn from a positively skewed distribution will be judged larger than the same attribute value that is drawn from a negatively skewed distribution (Birnbaum, 1974; Hagerty, 2000; Haubensak, 1992; Niedrich, Sharma, & Wedell, 2001; Parducci, 1965, 1995; Riskey et al., 1979; Stewart, Chater, & Brown, 2006). I will discuss these distribution-density effects and how CID theory might explain them in the next section.

#### **DISTRIBUTION-DENSITY EFFECTS**

Niedrich, Sharma, and Wedell (2001) described models of evaluation—like Helson's (1964) adaptation-level theory and the model formalized in Equation 4 as prototype models. An alternative to these models is suggested by the view that values might be compared to

other exemplars that are drawn from the distribution. That is, instead of being compared to a central prototypical example, a to-be-evaluated attribute value might instead be compared to large values sometimes, small values sometimes, and frequent values most commonly of all (i.e., most commonly compared to central values in a normally skewed distribution, large values in a negatively skewed distribution, and small values in a positively skewed distribution). The most successful exemplar model of evaluation in the literature is Parducci's (1965; 1995) range-frequency theory (RF theory) in which evaluations are affected by two types of information: the range score of x in distribution k ( $R_{xk}$ ) and the frequency score (percentile rank) of x in distribution k ( $F_{xk}$ ). The range score is calculated as shown in Equation 5:

$$\boldsymbol{R}_{xk} = \frac{\boldsymbol{X} - \boldsymbol{Min}_k}{\boldsymbol{Max}_k - \boldsymbol{Min}_k} \tag{5}$$

where  $Max_k$  is the largest value in distribution k and  $Min_k$  is the smallest value in distribution k (see also Janiszewski & Lichtenstein, 1999; Volkmann, 1951). The frequency score represents the percentile rank of value x among all values in distribution k as shown in Equation 6:

$$\boldsymbol{F}_{xk} = \frac{\boldsymbol{Rank}_{xk} - 1}{\boldsymbol{N}_k - 1} \tag{6}$$

where  $Rank_{xk}$  is the rank of value x in distribution k and  $N_k$  is the number of exemplars in distribution k. It is these frequency scores that create distribution-density effects. Category rating evaluations of x are assumed to be a linear function of the range-frequency compromise score of x (RFscore<sub>x</sub>) calculated as a weighted average of  $R_{xk}$  and  $F_{xk}$  as shown in Equation 7:

$$\boldsymbol{RFscore}_{x} = \boldsymbol{wR}_{xk} + (1 - \boldsymbol{w})\boldsymbol{F}_{xk} \tag{7}$$

where w is a weighting parameter.

RF theory predicts s-shaped evaluation functions for any distribution that is dense at the center of the distribution and becomes sparse toward the tails on each side (e.g., normal distributions; see Figure 4, top left). It predicts s-shaped evaluation functions in these cases, however, not because people are less sensitive to differences at greater distances from the RP (i.e., not because of Stevens', 1961, Law), but rather because the density or frequency of exemplars in the distribution produces frequency scores (i.e.,  $F_{xk}$ ; Equation 6) that are s-shaped. The middle portion of the curve is steep, because values are dense at the center of the distribution. The slopes of the lower and upper portions of the s-shaped curve are shallow, because values are sparse at the tails of the distribution.

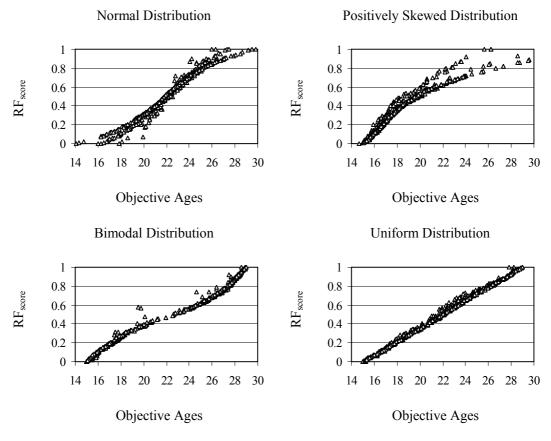


Figure 4: Age evaluation functions predicted by range-frequency (RF) theory. Evaluations follow range-frequency compromise scores (R-F Scores) and produce s-shaped evaluation functions for normal distributions, but not for skewed, bimodal or uniform distributions.

Unlike prototype models of evaluation, RF theory does not predict s-shaped evaluation functions for skewed (Figure 4; upper right), bimodal (Figure 4; lower left; the bimodal distributions in the simulations and proposed experiments below are normal distributions with the tails flipped in giving the end points the greatest density), or uniform (Figure 4; lower right) distributions. Rather, RF theory predicts that evaluation functions will be concave (downward) for positively skewed distributions, because the dense region at the lower end of positively skewed distributions produces larger frequency scores than would have been produced otherwise. It predicts that evaluation functions will be convex (concave upward) for negatively skewed distributions, because the sparse region at the lower end of negatively skewed distributions produces smaller frequency scores than would have been produced otherwise. Evaluation functions ought to flatten somewhat at the center of bimodal distributions because frequency at the center of bimodal distributions is sparse. Evaluations ought to be linear for uniform distributions because frequency is uniformly distributed.

Consistent with both prototype models and RF theory, Niedrich et al. (2001) found that evaluations of prices were s-shaped for normal distributions. But contrary to the predictions

of prototype models<sup>1</sup>, evaluations of prices in skewed and bimodal distributions were exactly as RF theory predicted they ought to have been. That is, evaluation functions were concave (concave downward) for positively skewed distributions and convex (concave upward) for negatively skewed distributions. Unlike normal distributions, the evaluation functions for bimodal distributions flattened in the sparse middle region. These results are consistent with previous research that has found distribution-density effects on evaluations (Birnbaum, 1974; Mellers & Birnbaum, 1982; Parducci, 1965, 1995; Riskey et al., 1979; Sokolov, Pavlova, & Ehrenstein, 2000; Wedell, Parducci, & Roman, 1989).

Although these results support RF theory, this theory is not able to account for several important findings. Most notably, consistent with CID theory and inconsistent with RF theory, in some cases people seem to have explicit RPs in mind (Holyoak & Mah, 1982). When evaluating body size, for example, people sometimes have a very particular body-size ideal in mind (Irving, 1990; Phelps et al., 1993; Richins, 1991). Unlike CID theory, RF theory also has no explanation for why evaluation functions are often steeper for changes that represent losses than for changes that represent gains (Kahneman & Tversky, 1979; but see Stewart et al., 2006).

Like prototype models of evaluation, CID theory predicts that whenever people verbally compare values to prototypical RPs, comparison-induced biases will generally produce sshaped evaluation functions regardless of the type of distribution from which values are drawn (see Figure 3). Like RF theory, CID theory predicts that whenever people compare values to other exemplars drawn from the distribution evaluation functions will depend upon distribution density. Like RF theory, CID theory predicts s-shaped evaluation functions for normal distributions, but not skewed, bimodal, or uniform distributions when values are compared to other exemplars drawn from the distribution. Unlike RF theory, CID theory provides a natural account of how greater aversion to losses than affinity for gains could produce asymmetries in sensitivity to losses and gains.

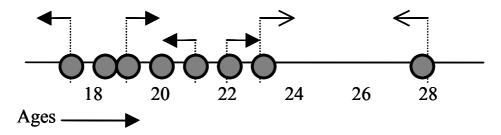


Figure 5: Comparison-induced biases in a positively skewed distribution of ages.

To understand how comparisons to other exemplars in the distribution could create distribution-density effects, consider the positively skewed distribution of ages presented in Figure 5. Filled-in arrows represent biases created by comparisons between values that are closer together than the comparison-suggested difference and that are, therefore, biased apart. Outlined arrows represent biases created by comparisons between values that are farther apart than the comparison-suggested difference and that are, therefore, biased together. Values in

<sup>&</sup>lt;sup>1</sup> A model utilizing Equation 4 in which the RP is always updated to be the value presented on the most recent trial or trials (Frederick & Loewenstein, 1999) would also produce distribution-density effects as described here.

dense regions (i.e., 18 - 22 years in Figure 5) are more likely to be closer than the comparison-suggested difference biasing evaluations apart. Values in sparse regions (i.e., 23 - 28 years) are more likely to be farther apart than the comparison-suggested difference biasing evaluations together. This difference in the effects of comparisons within dense versus sparse regions allows comparisons to create density effects.

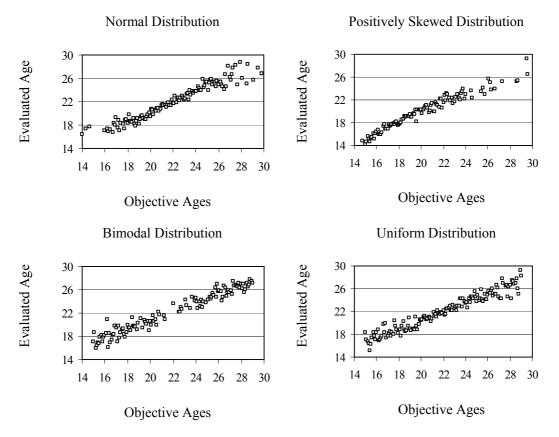


Figure 6. Simulated comparison-biased age evaluations. All values compared to contextual values.

To demonstrate how biases produced by verbal comparisons could produce distribution density-effects, five hundred ages were randomly drawn from each of a normal, a positively skewed, a bimodal, and a uniform distribution. As in the simulation presented in Figure 3, Parameter D in Equations 1a and 1b was set to 3.5 years, Parameter c in Equation 3 was set to 1.5 (these values were chosen for demonstration purposes only; before modeling empirical data, these two values would have to be empirically measured), and Parameter w in Equations 2a and 2b was set to .5 (this value was also chosen for demonstration purposes only; to model empirical data, this value would be found by fitting this free parameter to data). Each of the 500 values was compared to one other randomly drawn value. The results are presented in Figure 6 and they demonstrate that, like RF theory, CID theory predicts s-shaped evaluation functions for normal distributions (upper left), but not skewed (upper right), bimodal (lower left), or uniform (lower right) distributions when values are compared to other exemplar values in the distribution.

Before I conclude the discussion of distribution-density effects, please note that CID theory does not predict that evaluations will always be affected by distribution density. As demonstrated above, CID theory predicts that evaluation functions will be s-shaped whenever people compare all values to one RP and compared them to no other values. In addition, CID theory predicts no biases, if people do not compare values or Parameter w in Equation 2 is very small.

### **ANCHORING EFFECTS**

In the anchoring effect paradigm, participants compare an unknown value to an arbitrary value called the anchor and then estimate the unknown value (but see Wilson, Houston, Brekke, & Etling, 1996, for a slightly different anchoring effect paradigm). Participants might, for example, be asked whether the length of the Mississippi River is longer or shorter than 3,000 miles and then estimate the length of the Mississippi River. Previous research has consistently found that estimates are biased toward the anchor. In our example, estimates of the length of the Mississippi River and the setimates of the length of the Mississippi River.

CID theory offers one key insight into these anchoring effects, namely that estimates will usually be biased toward comparison-suggested values rather than toward anchor values per se (see Wilson et al., 1996, for a possible exception). This insight should be commonsensical. It bears to reason that if people think that the Mississippi River is shorter than 3,000 miles, then their estimates will not be biased toward 3,000 miles per se, but rather their estimates would be biased toward whatever value they think of as "less than 3,000 miles." That is, their estimates would be biased toward a comparison-suggested difference less than 3,000 miles. Before I describe simulations demonstrating the role that comparison-induced biases might play in creating anchoring effects, it is important to note that CID theory's possible contribution to our understanding of anchoring effects is orthogonal to the contributions offered by previous models of anchoring effects. Following previous models (Anderson, 1965; Huttenlocher et al., 2000), CID theory is a weighted average model of attribute evaluation. Like those previous models, it does not specify the mechanism underlying this weighted averaging. Comparisons could create anchoring effects by changing search strategies (Tversky & Kahneman, 1974), bringing to mind different diagnostic attributes (Strack & Mussweiler, 1997), priming values (Wilson et al., 1996), or changing conversational norms (Schwarz, 1990). Regardless of the particular mechanism that creates this weighted-average biasing, CID theory predicts a slightly different pattern of estimation than the pattern predicted by previous models. CID theory might thereby contribute key insights to our understanding of anchoring effect phenomena, but would not replace these previous models.

As described above, CID theory predicts that comparisons will bias estimates toward comparison-suggested values. If comparison-suggested values are closer to anchor values than are unbiased estimates, then comparisons will bias estimates toward anchor values. In particular, CID theory predicts biases toward anchor values whenever unbiased estimates are farther than a comparison-suggested difference away from the anchor or people judge the unknown value to be approximately the same as the anchor in which case the comparisonsuggested value would be the anchor value. Previous research on anchoring effects has typically found such biases toward anchor values. By contrast, however, if comparisonsuggested values are farther away from anchor values than are unbiased estimates, then comparisons will bias estimates away from anchor values. In typical anchoring effect scenarios such biases away from anchor values are rare for two reasons. First, because comparison-suggested differences are small relative to the range of people's estimates, most unbiased estimates will be more than a comparison-suggested difference away from the anchor. Most estimates will then be biased toward the anchor value leaving only a few estimates that would not be. Second, many people whose unbiased estimates are less than a comparison-suggested difference away from the anchor will judge the unknown value to be approximately the same as the anchor. These estimates are less than a comparison-suggested difference away from the anchor and those unbiased estimates are judged to be different from the anchor (i.e., more or less than the anchor) will estimates be biased away from the anchor.

Even though the vast majority of values are biased toward the anchor value and this result is consistent with both CID theory and alternative theories, CID theory nevertheless makes novel predictions regarding the magnitude of the predicted biases toward the anchor value. In particular, CID theory predicts large biases toward the anchor among unbiased estimates at the extremes, smaller biases toward the anchor for unbiased estimates closer to the comparison-suggested difference away from the anchor, and large biases toward the anchor again as more and more participants start to describe the unknown value as "approximately the same" as the anchor.

To demonstrate how comparison-induced distortions could create anchoring effects, 500 values between 500 and 3100 were randomly drawn from a normal distribution to represent simulated unbiased estimates of the length of the Mississippi River. A histogram of the results is presented in the top panel of Figure 7. I then simulated comparison-biased estimates by comparing all 500 values to an anchor of 2,900 miles. For demonstration purposes only, the comparison-suggested difference (Parameter D in Equation 1) was set at 300 miles. To model real data, one would have to measure this value by asking a control group of participants about the length they imagine when they hear about a river that is shorter than 2.900 miles as described in the measuring comparison-suggested differences section above. Also for demonstration purposes only Parameter c in Equation 3 was set at 0.0002. Again to model real data, one would have to measure this value by asking control groups of participants whether they would describe a variety of values as the same as or less than the anchor value, fit Equation 3 to their responses, and find the best-fitting value for Parameter c. With Parameter c set at 0.0002, 18.6% of values were judged to be approximately the same as the anchor value of 2,900 miles. Parameter w in Equation 2 was set at 0.5. To model real data, Equations 1a through 3 would be fit to the data using the empirically measured values for Parameters D and c to find the best-fitting value for Parameter w. The results are presented in the middle panel of Figure 7. Notice that values were biased toward the anchor value of 2,900 miles. Only 2.2% of values were biased away from the anchor value. I also simulated comparison-biased estimates by comparing all 500 values to an anchor of 600 miles. All parameter were given the same values. With Parameter c set at 0.0002, 20.8% of values were judged to be approximately the same as the anchor value. The results are presented in the bottom panel of Figure 7. Notice that values are biased toward the anchor value of 600 miles.

#### Simulated unbiased estimates

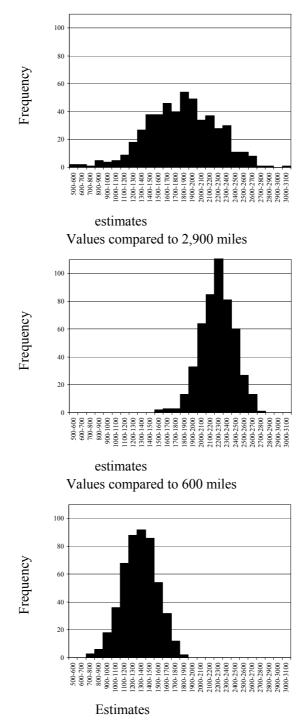


Figure 7. Simulated comparison-induced anchoring effects. The top panel presents a simulated unbiased distribution of estimates of the length of the Mississippi River. The middle panel presents what the distribution would look like, if all values were compared to 2,900 miles. The bottom panel presents what the distribution would look like, if all values were compared to 600 miles

Only 2.0% of values were biased away from the anchor value. Note that these results do not depend upon the particular parameter values I used here for demonstration purposes only. The same basic pattern of results is observed no matter what values these parameters take.

#### **SOCIAL COMPARISONS**

Festinger (1954) suggested that people evaluate their own personal-attribute values by comparing their attribute values to the attribute values of their peers. For example, people might evaluate their income, their intelligence, their appearance, their body size, and so forth by comparing themselves to their friends and others who are similar to them. Verbal comparisons of these personal attribute values are likely to produce the same biases as other verbal comparisons. That is, if actual differences are the same size as comparison-suggested differences, then there will be no biases in evaluation. If actual differences are smaller than comparison-suggested differences, then people will overreact to the difference. That is, evaluations will be biased toward the larger comparison-suggested difference. The smaller value will be evaluated smaller than it would have been evaluated otherwise and the larger value will be evaluated larger than it would have been evaluated otherwise. If actual differences are larger than comparison-suggested differences, then people will under react to the difference. That is, evaluations will again be biased toward the comparison-suggested difference, but this time the comparison-suggested difference will be smaller. The smaller value will then be evaluated larger than it would have been otherwise and the larger value will be evaluated smaller than it would have been otherwise. The overall effect of these biases would be such that moderate downward comparisons would produce higher evaluations of the self than would extreme downward comparisons and moderate upward comparisons would produce lower evaluations of the self than would extreme upward comparisons.

Consider, for example, Jennifer and Michelle, the two fictional DePaul University undergraduate students described in Table 1. If Jennifer whose income was \$7.868.00 were to compare her income to Michelle's income of \$9,000.00, then there would be no biases. The \$1,132.00 difference in their salaries would be exactly the same size as the \$1,132.00 comparison-suggested difference so reliance on comparison information would create no biases (recall that these were the incomes imagined by actual DePaul University undergraduate women for full-time female students working part time jobs to help pay for school expenses, notice how much larger the incomes imagined by DePaul University undergraduate men were for full-time male students). However, if Michelle had an income of \$8,000.00, instead of \$9,000.00, and Jennifer were to describe her \$7,868.00 income as less than Michelle's \$8,000.00 income or Michelle's \$8,000.00 as more than her \$7,868.00 income (a difference of \$132.00), then Jennifer would likely overreact to the difference (unless she were to describe these salaries as "approximately the same" in which case she would under react to the difference), because her evaluations would be biased toward the larger \$1,132.00 comparison-suggested difference. Her income would be evaluated smaller than it would have been otherwise and Michelle's income would be evaluated larger than it would have been otherwise. By contrast, if Michelle had an income of \$10,000, instead of \$9,000.00, and Jennifer were to describe her \$7,868.00 income as less than Michelle's \$10,000 income (a difference of \$2,132.00), then Jennifer would likely under react to the difference. Once again her evaluations would be biased toward the \$1,132.00 comparisonsuggested difference, but this time the \$1,132.00 comparison-suggested difference will be smaller than the \$2,132.00 actual difference. Her \$7,868.00 income would then be evaluated larger than it would have been evaluated otherwise and Michelle's \$10,000.00 income would be evaluated smaller than it would have been evaluated otherwise.

Simulated biases created when Jennifer compares her own and other people's salaries are presented in Figure 8. Jennifer's comparison-biased evaluations of other people's salaries (Figure 8 left panel) would follow an s-shaped evaluation function. This function is almost identical to the evaluation function shown in Figure 2 Salary Evaluations. The only difference would be that her reference point (RP) would be her own \$7,868.00 income, rather than the RP of \$50,000.00 shown in Figure 2 Salary Evaluations. Like Figure 2 Salary Evaluations, all other incomes would be evaluated relative to her RP. The top line represents evaluations of other people's salaries when she describes their salaries as more than her own \$7,868 salary. The middle line represents evaluations of other people's salaries when she describes their salaries as less than her own salary. The Stevens' (1961) Law formalization of the evaluation function function (Kahneman & Tversky, 1979) is also plotted. CID theory predicts a continuous function similar to Stevens' Law when evaluations from less than, approximately the same, and more than comparisons are averaged together (see Figure 3).

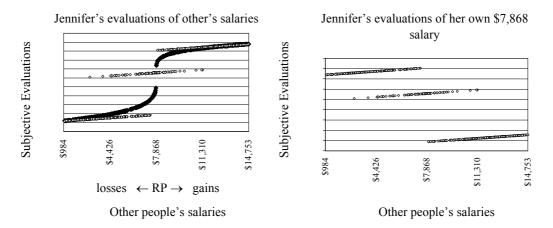


Figure 8. Biases created by social comparisons. The left panel presents comparison-biased evaluations of other people's salaries. The right panel presents comparison-biased evaluations of Jennifer's own salary. The middle lines in both panels represent biases created when Jennifer describes her salary as approximately the same as other people's salaries. The right most lines (top line on the left panel; bottom line on the right panel) represent biases created when Jennifer describes other people's salaries as more than her salary or her salary as less than other people's salaries. The left most lines (bottom line on the left panel; bottom line on the right panel) represent biases created when Jennifer describes other people's salaries as more than her salary or her salary as less than other people's salaries. The left most lines (bottom line on the left panel; top line on the right panel) represent biases created when Jennifer describes other people's salaries.

Jennifer's comparison-biased evaluations of her own salary are plotted in Figure 8 right panel. The top line represents evaluations of her own \$7,868 salary when she describes it as more than other people's salaries. The middle line represents evaluations of her salary when she describes it as approximately the same as other people's salaries. The bottom line represents evaluations of her salary when she describes it as less than their salaries. Notice the

unique predictions of this CID theory account of social comparisons. This account predicts that comparisons to salaries that are moderately lower than her own salary will cause her to evaluate her salary higher than comparisons to salaries that are much lower than her own salary. Likewise, comparisons to salaries that are moderately higher than her own salary will cause her to evaluate her salary lower than comparisons to salaries that are much higher than her own salary. This pattern of biases might not be observable using category rating measures of evaluation (but see Qian & Brown, 2005), but has been observed in my laboratory using recall of values from memory as the dependent measure.

## **DECOY EFFECTS**

Consumers typically have to pick one option among several alternatives. If they are purchasing an airline ticket, they typically have several flight options to choose from and they have to pick one. If they are renting an apartment, they typically have to decide which of several apartments to rent. If they are purchasing pickles, they typically have to decide which of several brands, flavors, and jar sizes to purchase. These options typically vary along multiple attributes and choices between them often involve trade-off between desirable attribute values. One flight option might leave later in the morning allowing the consumer to sleep later but arrive at the destination later and have a longer layover. One apartment might have a shorter commute to work but cost more in rent. To study how people choose among multi-attribute alternatives such as these, researchers have studied the effects of decoys on choice. Decoys are options inserted into a choice set that people typically do not choose. Despite the fact that people do not choose these decoy options, they affect which of the remaining options people chose.

I will discuss two such decoy effects in this section: the asymmetric dominance effect and the phantom decoy effect. Both effects start with two options that typically vary along two dimensions (e.g., two apartment options that vary in rent and the length of the commute between the apartment and work). One option is better on one dimension, but the other option is better on the other dimension. For example, one apartment might have a shorter commute but cost more in rent than the other apartment. In Figure 9, apartment option 1 has a better commute time (25 minutes compared to 45 minutes for option 2), but apartment option 2 costs less in rent (\$575 per month compared to \$700 per month). Researchers try to set these values so that people are indifferent between these options. If the options in Figure 9 were set right (the correct values would differ across different groups and would need to be pretested), the value of the shorter commute from one apartment would be exactly equal to the extra rent one would pay for that apartment so that if those two options were the only ones presented people would be indifference curve.

In the asymmetric dominance effect (Huber et al., 1982), researchers place a dominated decoy in the option set (see Figure 9). This dominated decoy is similar to but worse than one of the other options so that no one would actually choose it. The dominated decoy in Figure 9, for example, is similar to option 1, but it has a longer commute and costs more in rent so that no one would actually choose it if option 1 were available. Nevertheless, its inclusion into the choice set affects the choices people make such that people will be more likely to choose the

option that is similar to and better than the dominated decoy. That is, the inclusion of the dominated decoy in the choice set changes people's preferences so that they are no longer indifferent between option 1 and option 2, pushing option 1 above the indifference curve, pulling the indifference curve down, or both.

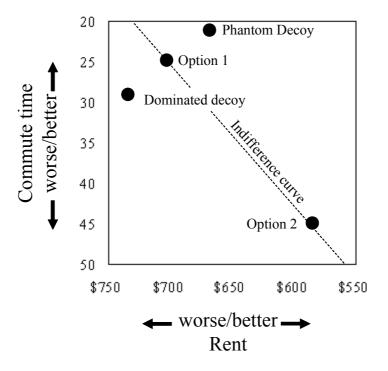


Figure 9. Decoy effects. Options 1 and 2 are designed to be equally attractive (because the benefit of the better value on one dimension is exactly equal to the detriment of the worse value on the other dimension) and so they lie on the indifference curve. Adding a dominated decoy to the option set causes people to choose the option that is similar to the dominated decoy (an effect called the asymmetric dominance effect). Adding a phantom decoy to the option set and marking it unavailable also causes people to choose the option that is similar to the phantom decoy (an effect called the phantom decoy effect).

In the phantom decoy effect (Highhouse, 1996; Pratkanis & Farquhar, 1992), researchers place a phantom decoy in the option set (see Figure 9). This dominating decoy is similar to but better than one of the other options. The phantom decoy in Figure 9, for example, is similar to option 1, but it has a shorter commute and costs less in rent. If this option were available, almost everyone would choose it; but it is marked unavailable, sold out, discontinued, etc. so that it cannot be chosen. Nevertheless, its inclusion into the choice set affects the choices people make such that people will be more likely to choose the option that is similar to the dominating decoy despite the fact that it is not as good as the dominating decoy in the asymmetric dominance effect, changes people's preferences so that they are no longer indifferent between option 1 and option 2. The inclusion of the phantom decoy pulls option 1 above the indifference curve, pushes the indifference curve down, or both.

Choplin and Hummel (2002) argued that CID theory could explain the asymmetric dominance effect; and Choplin and Hummel (2005) discovered a one dimensional analogue of the asymmetric dominance effect and argued that unlike other models in the literature CID theory could account for this one dimensional effect. The basic idea underlying the CID theory account of the asymmetric dominance effect is that verbal comparisons (e.g., better, closer, less expensive, worse, farther, more expensive) between the option that is similar to the dominated decoy (option 1 in Figure 9) and the dominated decoy bias the evaluation of the option that is similar to the dominated decoy. Because the differences between the dominated decoy and the option that is similar to the dominated decoy are typically quite small (i.e., typically smaller than comparison-suggested differences), comparisons bias evaluations apart making the option appear better (and the dominated decoy worse) than it otherwise would and so people choose the dominating option. To simulate these predictions, I modeled the asymmetric dominance effect scenario presented in Figure 9. The choice set included apartment option 1 with a commute of 45 minutes and rent of \$575.00 per month, apartment option 2 with a commute of 25 minutes and rent of \$700.00 per month, and the dominated decoy with a commute of 29 minutes and rent of \$750.00 per month. To model the effects of comparisons on evaluations, I set the comparison-suggested difference (Parameter D in Equation 1) for commute times at 8 minutes (any value larger than 4 minutes would create essentially the same pattern of bias) and the comparison-suggested difference for rent at \$100.00 (any value larger than \$50.00 would create essentially the same pattern of bias). With these values for Parameter D, CID theory predicts an asymmetric dominance effect as long as Parameter c in Equation 3 takes a value greater than 0.173 for commute times (the probability of describing a 4-minute difference in commute times as "approximately the same" is less than 50.0%) and Parameter c takes a value greater than 0.014 for rents (the probability of describing a \$50.00 difference in rent as "approximately the same" is less than 50.0%). These predictions hold regardless of what value the weighting parameter (Parameter w in Equation 2) takes as long as that value is greater than 0.0.

Given this CID theory explanation of the asymmetric dominance effect it might appear at first that CID theory would be unable to account for phantom decoy effects (Highhouse, 1996; Pratkanis & Farquhar, 1992). Since the differences between phantom decoys and the items they dominate are also typically quite small (i.e., they also are typically smaller than comparison-suggested differences), comparison between phantom decoys and options should make the option that is similar to the phantom decoy appear worse. Why then would people choose the similar option?

CID theory could simultaneously account for both the asymmetric dominance effect and the phantom decoy effect only if the unavailability of the phantom decoy were to make people more likely to describe the difference as "approximately the same." Recent empirical results in my laboratory have found evidence that people are more likely to describe options as "approximately the same as" unavailable options than to describe them as "approximately the same as" available options. To simulate these predictions, I modeled the phantom decoy effect described above, the choice set included apartment option 1 with a commute of 45 minutes and rent of \$770.00 per month. Instead of the dominated decoy with a commute of 29 minutes and rent of \$750.00 per month, however, the phantom decoy with a commute of 21 minutes and a rent of \$650.00 was placed in the set, but was not available. To model the effects of

comparisons on evaluations, I used the same values for Parameter D described in the simulation of the asymmetric dominance effect above. With these values, CID theory predicts a phantom decoy effect as long as Parameter c in Equation 3 takes a value less than 0.173 for commute times (the probability of describing a 4-minute difference in commute times as "approximately the same" is greater than 50.0%) and Parameter c takes a value less than 0.014 for rents (the probability of describing a \$50.00 difference in rent as "approximately the same" is greater than 50.0%). Again, these predictions hold regardless of what value the weighting parameter (Parameter w in Equation 2) takes as long as that value is greater than 0.0. If people were not more likely to describe decoys as "approximately the same" as options when decoys were marked unavailable than when they were not, CID theory would not be able to simultaneously account for both asymmetric dominance effects (Huber et al., 1982) and phantom decoy effects (Highhouse, 1996; Pratkanis & Farquhar, 1992) at the same time.

#### CONCLUSION

Simulation results demonstrated that biases created by verbal comparisons are capable of explaining or providing insights into a variety of phenomena in the psychology of judgment and decision making. In particular, simulation results demonstrated that biases created by verbal comparisons might help explain or provide insights into s-shaped evaluation functions (Frederick & Loewenstein, 1999; Kahneman & Tversky, 1979), distribution-density effects (Birnbaum, 1974; Niedrich et al., 2001; Parducci, 1965, 1995; Riskey et al., 1979), anchoring effects (Tversky & Kahneman, 1974), social comparisons (Festinger, 1954), and decoy effect such as the asymmetric dominance effect (Huber et al., 1982) and the phantom decoy effect (Highhouse, 1996; Pratkanis & Farquhar, 1992). Because verbal comparisons are ubiquitous whenever people decide between alternatives, future research may find that biases created by verbal comparisons might help explain or provide insights into other decision-making phenomena as well.

Simulation results demonstrated that biases created by verbal comparisons might help explain why evaluation functions are s-shaped (Frederick & Loewenstein, 1999; Kahneman & Tversky, 1979) when people compare to-be-evaluated values to a standard or reference point value. Consistent with previous research on s-shaped evaluation functions, comparisons to a standard or reference point would cause people to over emphasize the importance of small differences and under emphasize the importance of large differences. In addition, some people might describe small differences as "approximately the same." Describing small differences as "approximately the same" would cause people to under emphasize the importance of small differences. A continuous, monotonic function would be created by averaging across the different types of comparisons people make (i.e., less than, approximately the same, and more than comparisons). This continuous function would be concave downward for positive changes from the reference point and convex (concave upward) for negative changes from the reference point. Affective factors cause people to be loss averse (disliking losses more than they like gains). Biases created by verbal comparisons might mediate some of these loss aversion effects, if affective factors cause people to use different comparison words to describe losses than to describe gains, and these comparisons, in turn, affect evaluations.

Simulation results also demonstrated that reliance on information from verbal comparisons might help explain distribution-density effects (Birnbaum, 1974; Niedrich et al., 2001; Parducci, 1965, 1995; Riskey et al., 1979), if people compare to-be-evaluated attribute values to other values drawn from the distribution (i.e., compare them to smaller values sometimes, larger values sometimes, and frequent values—middle values in normal distributions, small values in positively skewed distributions, or large values in negatively skewed distributions—most frequently of all). Because these comparisons would also cause people to over emphasize the importance of small differences and under emphasize the importance of large differences, comparisons to these other values would cause the evaluation function to be s-shaped for values drawn from a normal distribution, concave downward for values drawn from negatively skewed distributions, inverted s-shaped for values drawn from bimodal distributions, and linear for values drawn from uniform distributions.

Simulation results demonstrated that considering the effects of verbal comparisons on estimates might provide insights into anchoring effects. In particular, if estimates are biased by verbal comparisons then when a participant notes that the unknown value is more than (or less than) the anchor value estimates will not be biased toward the anchor value per se, but rather estimates will be biased toward the comparison-suggested value more than (or less than) the anchor value. This CID theory insight into anchoring phenomena is orthogonal to the insights provided by most current accounts of anchoring effects (Schwarz, 1990; Strack & Mussweiler, 1997; Tversky & Kahneman, 1974; Wilson et al., 1996), but is important in its own right.

Biases produced by verbal social comparisons might also provide insights into social comparison phenomena. In particular, simulation results demonstrate that moderate downward comparisons will produce higher evaluations of the self than will extreme downward comparisons and moderate upward comparisons will produce lower evaluations of the self than will extreme upward comparisons.

Because consumers typically compare alternatives whenever they make a choice, biases produced by reliance on verbal comparisons might help explain some consumer decisionmaking phenomena. In particular, biases produced by verbal comparisons might help explain the effects of decoys on consumer choice. Choplin and Hummel (2002) demonstrated that biases produced by verbal comparisons might explain the asymmetric dominance effect (Huber et al., 1982) and Choplin and Hummel (2005) demonstrated a one-dimensional version of the asymmetric dominance effect and argued that CID theory provided the best account of this phenomenon. Simulation results presented here demonstrated that biases produced by verbal comparisons would be able to simultaneously account for both the asymmetric dominance effect and the phantom decoy effect only if the unavailability of the phantom decoy makes people more likely to describe other options as "approximately the same" as the decoy.

The simulation results presented here demonstrate that the CID theory account of these phenomena is feasible and that biases produced by verbal comparisons are either sufficient to explain phenomena or else provide important insights into phenomena. These simulation results have not shown that biases produced by verbal comparisons are necessary to explain these phenomena nor have they shown that CID theory provides a better account of these phenomena than alternative theories. Arguments that CID theory provides a better explanation of these phenomena than other theories are presented elsewhere. Importantly, however, the fact that CID theory can account for such a large variety of phenomena is itself evidence in favor of CID theory. Far too often, researchers develop theories to account for one phenomenon or class of phenomena, rather than proposing theories that generalize across a wide variety of phenomena. CID theory follows Helson's (1964) adaptation-level theory and Parducci's (1965; 1995) range-frequency theory in proposing an account of attribute evaluation that generalizes across a wide variety of phenomena. Biases in evaluation produced by verbal comparisons are predicted whenever people compare alternatives. Since such comparisons are ubiquitous, this influence is likely to be pervasive.

#### REFERENCES

- Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression formation. *Journal of Experimental Psychology*, 70, 394-400.
- Birnbaum, M. H. (1974). Using contextual effects to derive psychophysical scales. Perception & Psychophysics, 15, 89-96.
- Briesch, R. A., Krishnamurthi, L., Mazumdar, T., & Raj, S. P. (1997). A comparative analysis of reference price models. *Journal of Consumer Research*, 24, 202-214.
- Choplin, J. M., & Hummel, J. E. (2002). Magnitude comparisons distort mental representations of magnitude. *Journal of Experimental Psychology: General*, 131, 270-286.
- Choplin, J. M., & Hummel, J. E. (2005). Comparison-induced decoy effects. *Memory & Cognition*, 33, 332-343.
- Festinger, L. (1954). A theory of social comparison processes. Human Relations, 7, 117-140.
- Frederick, S., & Loewenstein, G. (1999). Hedonic Adaptation. In D. Kahneman, E. Diener & N. Schwarz (Eds.), Well-being: The foundations of hedonic psychology (pp. 302-329). New York: NY: Russell Sage.
- Hagerty, M. R. (2000). Social comparisons of income in one's community: Evidence from national surveys of income and happiness. *Journal of Personality and Social Psychology*, 78, 764-771.
- Haubensak, G. (1992). The consistency model: A process model for absolute judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 303-309.
- Helson, H. (1964). Adaptation-level theory. New York: Harper & Row.
- Highhouse, S. (1996). Context-dependent selection: The effects of decoy and phantom job candidates. *Organizational Behavior and Human Decision Processes*, 65, 68-76.
- Holyoak, K. J., & Mah, W. A. (1982). Cognitive reference points in judgments of symbolic magnitude. *Cognitive Psychology*, 14, 328-352.
- Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9, 90-98.
- Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General, 129*(2), 220-241.

- Irving, L. M. (1990). Mirror images: Effects of the standard of beauty on the self- and bodyesteem of women exhibiting varying levels of bulimic symptoms. *Journal of Social and Clinical Psychology*, 9, 230-242.
- Janiszewski, C., & Lichtenstein, D. R. (1999). A Range Theory Account of Price Perception. Journal of Consumer Research, 25(25), 353-368.
- Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*. New York, NY: Russell Sage.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263-291.
- Kalyanaram, G., & Little, J. D. C. (1994). An empirical analysis of latitude of price acceptance in consumer package goods. *Journal of Consumer Research*, 21, 408-418.
- Lavine, H., Sweeney, D., & Wagner, S. H. (1999). Dipicting women as sex objects in television advertising: Effects on body dissatisfaction. *Personality and Social Psychology Bulletin*, 25, 1049-1058.
- Mellers, B. A., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. Journal of Experimental Psychology: Human Perception and Performance, 8, 582-601.
- Morse, S., & Gergen, K. J. (1970). Social comparison, self-consistency, and the concept of self. *Journal of Personality and Social Psychology*, 16, 148-156.
- Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, 110, 472-489.
- Niedrich, R. W., Sharma, S., & Wedell, D. H. (2001). Reference price and price perceptions: A comparison of alternative models. *Journal of Consumer Research*, 28(3), 339-354.
- Parducci, A. (1965). Category judgments: A range-frequency model. *Psychological Review*, 72, 407-418.
- Parducci, A. (1995). *Happiness, pleasure and judgment: The contextual theory and its applications*. Mahwah, NJ: Lawrence Erlbaum.
- Phelps, L., Johnston, L. S., Jimenez, D. P., Wilczenski, F. L., Andrea, R. K., & Healy, R. W. (1993). Figure preference, body dissatisfaction, and body distortion in adolescence. *Journal of Adolescent Research*, 8, 297-310.
- Pratkanis, A. R., & Farquhar, P. H. (1992). A brief history of research on phantom alternatives: Evidence for seven empirical generalizations about phantoms. *Basic and Applied Social Psychology*, 13, 103-122.
- Qian, J., & Brown, G. D. A. (2005). Similarity-based sampling: Testing a model of price psychophysics. Paper presented at the Twenty-seventh annual conference of the cognitive science society, Stresa, Italy.
- Richins, M. L. (1991). Social Comparison and the Idealized Images of Advertising. *Journal* of Consumer Research, 18, 71-83.
- Riskey, D. R., Parducci, A., & Beauchamp, G. K. (1979). Effects of context in judgments of sweetness and pleasantness. *Perception & Psychophysics, 26*, 171-176.
- Rusiecki, J. (1985). Adjectives and Comparison in English. New York: Longman.
- Schifferstein, H. N. J., & Frijters, J. E. R. (1992). Contextual and sequential effects on judgments of sweetness intensity. *Perception & Psychophysics*, 52, 243-255.
- Schwarz, N. (1990). Assessing frequency reports of mundane behaviors: Contributions of cognitive psychology to questionnaire construction. In C. Hendrick & M. Clark, S. (Eds.), Research Methods in Personality and Social Psychology (Review of Personality and Social Psychology, vol. 11) (pp. 98-119). Beverly Hills, CA: Sage.

- Schwarz, N., & Strack, F. (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*. New York, NY: Russell Sage.
- Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. Science, 237, 1317-1323.
- Sokolov, A., Pavlova, M., & Ehrenstein, W. H. (2000). Primacy and frequency effects in absolute judgments of visual velocity. *Perception & Psychophysics*, 62, 998-1007.
- Stevens, S. S. (1961). To honor Fechner and repeal his law. Science, 133, 80-86.
- Stevens, S. S. (1962). The surprising simplicity of sensory metrics. *American Psychologist*, 17, 29-39.
- Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53, 1-26.
- Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73, 437-446.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1130.
- Volkmann, J. (1951). Scales of judgment and their implications for social psychology. In J. H. Roherer & M. Sherif (Eds.), *Social psychology at the crossroads* (pp. 279-294). New York: Harper & Row.
- Wansink, B. (2004). Environmental factors that increase the food intake and consumption volume of unknowing consumers. *Annual Review of Nutrition, 24*, 455-479.
- Wedell, D. H., & Parducci, A. (2000). Social Comparison: Lessons from basic research on judgment. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 223-252). New York: Plenum.
- Wedell, D. H., Parducci, A., & Roman, D. (1989). Student perceptions of fair grading: A range-frequency analysis. *American Journal of Psychology*, 102(233-248).
- Wedell, D. H., & Pettibone, J. C. (1996). Using judgments to understand decoy effects in choice. Organizational Behavior and Human Decision Processes, 67, 326-344.
- Wilson, T. D., Houston, C. E., Brekke, N., & Etling, K. M. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 125, 387-402.

Chapter 44

# ANALYSIS AND DECISION-MAKING MODELS IN THE PROCESS OF THE ASSIMILATION OF CHANGE AND INNOVATION IN LEARNING SYSTEMS\*

# Baruch Offir

Bar-Ilan University, School of Education, Israel

Research has shown that a relation exists between the level of learning in the schools and universities and a country's strength. A relation also exists between education and the level and quality of life. Education today is a significant factor for ensuring society's normal existence, development and prosperity. However, major cities can afford the student the opportunity to acquire knowledge more than cities found in the periphery. A gap therefore exists between the level of learning in the major cities and the level of learning in the peripheral settlements. Students with high learning abilities who live in the cities can participate in university courses and other learning centers, whereas students with learning abilities who live in the periphery do not have a framework which can afford them knowledge in accordance with their talents and abilities.

This reality was the basis for our research on the integration of technological systems for the advancement of students in the periphery towards academic studies. Our research aims to investigate how technological systems can be used to advance populations of students who live in distant areas, to afford them the opportunity to learn academic courses and to be university students while still learning in high school.

A proper combination of this means during the learning process requires a change in the teaching method. This process of change is very complex, since it must take numerous educational and pedagogical factors which are involved in the process into account. It must recognize the teachers' and students' personal attitudes, must evaluate the student's level, analyze the sociological processes taking place in the classroom, formulate an appropriate teaching method, recognize the teacher's position and status in the classroom, etc. Proper

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

activation of technological systems in order to reduce gaps between populations is a very complex system and its successful implementation depends on the understanding and control of numerous diverse and complex parameters.

The function of the education system today is indeed complex. It must educate towards values and mold each student's behavior, afford the student the ability to crystallize his viewpoint and attitude while concomitantly leading him towards achievements and affording him the tools with which he will be able to learn and acquire a profession so that he will be able to earn a living for himself and his family and will be able to contribute to the society in which he lives. These goals are not identical, and are sometimes not compatible, since the strict and demanding educational framework which accurately evaluates and judges the student's achievements is not necessarily the same educational framework which is soft, encouraging, educating and guiding.

The education system is not static. History shows that its goals, methods of operation and structure have changed during different periods of history, according to sociological influences and society's goals. The goals of education have undergone changes according to the particular period, from a framework whose function was to watch over the children and thus enable the parents to join the workforce, to a framework which affords the student knowledge and skills in order to enable him to become integrated in society and contribute to his environment.

Education systems must develop the ability to change in order to cope with the goals set by society. The education system today needs to prove its effectiveness. Today's society demands that education systems prove that their methods and modes of operation are effective in achieving the goals of education as formulated by the society. These demands, of meeting standards set by the society, require that the education system develop instruments of evaluation and measurement, that it formulate more accurate methods for collecting and analyzing data and that it develop methods for reaching conclusions.

These processes will comprise the basis for a change in the methods of learning, for improvement in existing teaching approaches, for the construction of appropriate tests for evaluation, while affording more accurate and effective instruments for guiding the education system towards the effective achievement of its goals, as currently takes place in other fields such as medicine, engineering, agriculture and economics.

For this purpose the education system must undergo a process of change. Do we know how to generate a change in the education system? Are there mechanisms that can be used to generate change? Does education research afford us with tools to cope with these demands? Do the information and data supplied by education research join the body of knowledge that comprises an instrument for navigating in and improving the education system? Is education research identical to research in the other sciences in the ways in which it collects data and formulates theories, or does education research have unique goals, research methods and ways of reaching conclusions? If so, we must formulate the goals and determine the extent to which these goals are indeed achieved by education research. It is possible that the research approach in the field of education requires change. This requires the consolidation of a method which will enable the systematic achievement of this change.

Three terms which should be considered when referring to the process of consolidating a unique research approach in the field of education are:

- 1. Conceptualization: To what extent are the concepts which we use in the education research process indeed identical concepts?
- 2. Collecting and handling scientific knowledge: How can scientific knowledge be accumulated? How should the knowledge be handled and how should conclusions be reached?
- 3. Rules for constructing a body of knowledge: Can a body of knowledge be constructed? It may be appropriate to carry this out only for defined and limited sections in the field of education, whereas for others we will agree that they have no rules or that other rules apply to them .

It may be assumed that there exist factors in the education system which accelerate processes of change and others which decelerate these processes: Do we know which factors affect the process of change within the education system? What are the mechanisms which enable effective change? Can these changes be measured and evaluated?

#### **EVALUATING A PROCESS OF CHANGE**

We have carried out research on the integration of technological systems in the education system for the past three decades. However, our interest did not focus on the technological systems and their integration in the education system. Rather, we examined processes of change within the education system. The advanced technological systems comprised a tool by whose means we created a research situation that enabled examination of the mechanisms which characterize and accompany processes of change. For example, integration of computers in schools in the 1970's or systems for transmission of information via broadband lines on the internet today, require changes in the education system. The researches we performed monitored these attempts and tried to assess and measure the various variables related to this process and thus to contribute to a better understanding of the process of change: What factors affect this process? How does it occur in the education system? What variables slow the process of change?

A better understanding of the processes of change will make it possible to guide and direct these processes. It will enable assessment of their effectiveness and measurement of the change's contribution to the achievement of the education system's goals.

In this chapter we will deal with the integration of technological systems in the learning process. Integration of technological systems in the learning system is a change. Research on the various variables involved in this change will enable a better understanding of how the process of change takes place in the education system. It will also elucidate the factors and rules that affect this process.

Integration of a technological system must enable the education system to achieve its goal of educating and imparting knowledge to the student. It must therefore take into account the abilities, feelings, attitudes, wishes, personalities and worldviews of all factors involved in the process of change. Integration of technological systems in teaching and education requires making a change. A longitudinal research which will accompany the process of the integration of these systems in teaching and education may therefore afford measurement and evaluation tools and a better understanding of the process of change.

Our occupation with the subject of integrating technological systems is only the framework and research field through which we wish to study and understand processes, mechanisms and variables accompanying the process of change in the education system in order to consolidate ways in which to generate a change in the field of education and learning. The results of the investigations afford tools for generating and guiding change within the education system.

Any change in education which tries to become integrated in the school encounters a teacher who is found in the classroom. The teacher is therefore a most significant factor for the success or failure of the introduction of change into the learning process. The presence of a flesh and blood teacher is essential in some fields of the education system, where the teacher has no substitute. It will be more complicated and less justified to integrate technological systems in these fields. However, there are other fields in the learning process in which it will be easier to integrate technological systems. Technological systems can be used to relieve the teacher of some of his responsibility, especially in the transmission of knowledge. This will enable the teacher to invest more time and to concentrate on the promotion of processes in which his presence comprises an advantage, fields in which he cannot be substituted.

An accepted and consolidated process for analyzing education systems will assist us in evaluating and quantifying various variables that take part in the process of integrating a new technology system. The integration of a technology system in learning necessitates a determination of which measurable components of the learning process can be most easily transferred to the technology system and which subjective non-measurable components should be left to the teacher. We thus wish to design a map of the optimal division of roles between the teacher's contribution, the student's activity and the help afforded by the technology system and other variables in order to achieve a more effective teaching system.

According to this approach we will not measure whole learning systems. Rather, we will divide the super-goal of the system into secondary goals. Some of these goals can be accurately measured and evaluated while others cannot be measured. The Firm Effects Model statistical method enables division between the general investment and a goal-directed investment. In the analysis of a learning system we will also separate between the general learning goals and secondary learning goals which can be measured at different levels. The resultant evaluation is then composed of a measurement of different goals. Some of the secondary goals can be measured accurately, some can be measured less accurately and some can only be evaluated subjectively.

Integrating computers in the school is a process of change which depends on several factors, such as: the learning method, the teacher's role in the classroom, the student's contribution. Research which is directed towards reaching conclusions and understanding processes of change in the school can monitor the processes of integrating computers in the school.

Although many predicted that the integration of computers in teaching will lead to a revolution in the learning system, the appearance of computers within the learning framework did not lead to the expected change. This raised the following questions: Why was this process not realized? Why it did not generate a revolution in teaching? Indeed, expectations were high whereas the changes that took place were negligible.

The process of integrating computers in teaching in the 1980's comprised a research field for our team. It comprised a framework in which to investigate and understand the variables related to the process of change and to elucidate the reason why the expected change did not take place. (Offir, 1987; 1993 Offir etal 1998; 1990a; 1990b; 1995; 1999, Katz & Offir 1988; 1990a; 1990b; 19911; 1993; 1994)

Since the 1990's we have focused on the subject of distance learning. Examination and evaluation of distance learning processes also serve as a framework for understanding processes of change and renewal in the school. Our research on distance learning attempts to assess those factors which promote change and those which interfere and prevent change.

Understanding the variables involved in the assimilation of distance learning systems, which are based on the results of field experiments, may contribute to a more in-depth understanding of the process of change within the teaching and learning frameworks. Evaluating the attitudes of the teachers and the students, the teacher's personality, examining the teaching methods, achievements and other variables and their integration within a comprehensive theory will help us understand the process of integrating advanced technologies in the school. Their integration and evaluation will afford a better understanding of the process of change.

The subject of the process of change, from the 1980's to the 1990's, was the computer and its integration in the teaching process. Since the 1990's, the subject of the process of change has become the distance learning system. Researches dealing in the integration of computers in teaching as well as those dealing in distance learning aim at affording a better description and understanding of the processes of generating a change in the school . We need language. We need concepts that are as accurate as possible, which can be used to produce meaningful information in order to describe and evaluate a process and reach conclusions. The desired direction of development cannot be described without concepts, and research hypotheses cannot be hypothesized or proven. Progress cannot take place without a pool of concepts.

Concepts are formulated by asking questions that are as accurate as possible and affording answers to these questions. Concepts accumulate during generations of ongoing research. One research is the continuation of a previous research. Every research helps make the concept more accurate and differentiates it from other similar concepts.

Scientific research helps us differentiate between concepts which appear similar and defines the differences between similar concepts. A collection of concepts and elucidation of the connections between them affords a body of information which is essential for describing processes in any scientific field. A body of information is effective only when it is constructed of clear and defined concepts, with clear relations between them.

We must assess and measure the change in order to monitor and describe a change. This will enable the production of information on this change. Terms and concepts are necessary in order to define essential elements that comprise the process of change. Thus, scientific research deals in the definition of concepts and their classification. It affords tools for quantifying concepts. This process has been ongoing in scientific research for decades.

Concepts which are defined and clear to everyone are an essential condition in the processes of analyzing and understanding a process. The defined concept has a measurable quantitative value, and can be integrated in the process. Its intensity can be calculated and its effect can be evaluated. Defined concepts enable the achievement of results. They enable evaluation of the effectiveness of a process. Above all, concepts are tools which can be used to explain the past, analyze the present and predict the future. Science defines concepts and thus enables calculations, formulations of hypotheses as well as conclusions pertaining to the effectiveness of the process.

When we deal in learning, the task is much more complex. Not all concepts that are related to the learning process can be defined accurately and measured. The concepts brilliance, creativity, thinking ability and conclusion ability are concepts which are related to learning but are difficult to quantify. In contradistinction, concepts such as "achievement" can be defined, and can be used to judge the success of the process. They can be calculated and quantified more accurately. Measuring and quantifying is even more difficult when discussing and investigating changes in behavior.

Learning and educational theories contain measurable concepts as well as concepts which are difficult to measure. Reaching conclusions will be enabled when the definition of the concepts is more accurate. It will be possible when a larger number of concepts in the process are defined and a more accurate connection between concepts is found.

The need for conceptualization exists in all fields of science. It is, therefore, an essential condition for the development of an education and learning system, as it is essential for the development of any information system. The field of learning, of behavior, has indeed remained a field in which is it difficult to consolidate a paradigm, compared to the natural sciences or even compared to economics, where the result is usually measured by an economic profit. Three main reasons for this are:

The sample size: The sample in the natural sciences is very large. A researcher who bases his conclusions on a milliliter of material actually has a sample with an immense number of molecules. For example, a milliliter of water contains  $3.3 \times 10^{22}$  molecules. The researcher controls this immense sample, and learns from a sample with a large number of items. In contradistinction, the sample in the behavioral sciences is small and limited.

The research population: The research samples in the natural sciences are identical. An experiment can be performed on a chemical substance. This experiment can be repeated and will yield an identical result. The equation in the natural sciences presents the bottom line of the process. In contradistinction, in the behavioral sciences the individual student is also taken into consideration. Understanding the individual is a necessary and essential condition for understanding processes in the behavioral sciences.

The research and measurement tools: The research and measurement tools in the natural sciences are identical during all experimental and information collection processes. The research tools collect objective and measurable data (such as temperature, pressure, etc.). The situation is different in behavior research, where the concepts are subjective. The measurement tools are not identical and the collected data can be manipulated such that their conclusions are not identical.

Theories in the behavioral sciences which referred to the concept "learning" referred to the level of behavior that can be shaped, directed, behavior that can be controlled and evaluated and whose results could be measured. The behaviorist theory, when using the concept "learning", refers to a level of behavior that has rules, learning that can be directed and guided, learning that can be evaluated and judged. However, there exist additional levels of learning, such as discovery, invention, conclusion, etc. (Offir 2000b; 2003a ; 2006 ; Offir etal 2000b; 2005 ; 2002 ;2007 ; 2003b ;2000c ; 2004 )

### THE OCCURRENCE OF LEARNING

Attempts to understand the learning process have been made for centuries, beginning already with Aristotle, Socrates and Plato and on through the philosophical approaches of Hume, Luke and Kant.

How does learning occur? The answer to this question may be obtained by observing man in nature. In ancient times man was a nomad, migrating from place to place in order to find food for his family and livestock. Man then discovered, by asking the right questions and answering them, that there is no need to migrate. He discovered that he could grow his food near his home. He learned to grow wheat in the correct season, he learned to transport water in a canal in order to water this wheat, how to reap the wheat, grind it and bake bread. Man learned to control nature and direct it to his needs. The ability to ask questions, the ability to relate to his environment, the ability to investigate and reach conclusions converted man from a nomad into a wheat grower.

The entire process began when for the first time, by chance, water fell on a pile of wheat grains. The wheat germinated and man was not indifferent to this event. He asked questions: What factors led to the germination of the wheat? Can we ourselves promote the germination of the wheat? Can I grow a field of wheat?

A natural environment is a stimulus-rich environment. Stimuli arouse questions. An environment which supplies the student with questions at the appropriate level is an essential condition for learning. Curiosity, alertness, asking the right questions, logical thinking, an environment that stimulates the asking of questions, trial and error, these are the conditions that enable learning.

This approach emphasizes the importance of the interaction between the student and the learning environment. A student who is naturally curious tries to afford meaning to everything he sees. Knowledge and understanding are processes that are related to the student's tendency to afford meaning, to understand his surroundings. This is learning. The natural environment in which man lives is rich in stimuli and encourages asking questions and answering them. An environment that arouses curiosity is the most suitable environment for learning. Based on this viewpoint, efforts are made to create a learning environment in the classroom which is as natural as possible, an environment that stimulates the asking of questions.

The modern constructivist theory adopted this approach. The constructivist theory believes that knowledge is constructed by the student and is not transmitted by the teacher. Jerome Bruner calls this process "meaning making". This is a process of affording meaning which is the fundament of the constructivist theory. Development of thinking is carried out by thinking. Learning does not occur without understanding. Knowledge is acquired by power of the student as a result of contact with the learning material. Each student's interpretation of the learning material will comprise a stimulus according to his knowledge from the past and according to his intellectual abilities. Learning is the product of activity, response and referral to the material at the student's disposal. Every person has different wishes and therefore a different and complex behavior. Every person needs to be treated uniquely.

The constructivist theory claims that there is limited room for a system of rules in the learning process, the education process and the change in behavior. Every child and every person learns in a method that is unique to him. He learns from the environment, from a collection of stimuli which is appropriate for his ability, his perception and his knowledge. The contact and the relations with the environment, the thinking, the person's ability to cope with the learning material, are essential conditions for the occurrence of learning. Therefore, learning will occur in the classroom only if we supply the student with an environment rich in stimuli which are suitable and significant for his level, interest and knowledge.

Thus, the student does not learn from his teacher. The student learns from the actual thinking process. The student thinks about what he does, what he believes, what others are doing. Thinking is aroused by action, by creating contact. A different action creates different thinking. A different thinking requires use of memory, remembering prior knowledge, reading, learning, experiencing. We learn from experience, from cases, events and activities, from processes in which we were involved. The interpretation of our experience is influenced by what we know. Knowledge is created by thinking and is not transmitted by teaching. Knowledge is acquired by the student's activity and is influenced by his prior knowledge and experience.

Knowledge is influenced and explained by the environment in which it is acquired. The explanation takes place in the student's mind. Problems, questions, deliberations or agreement, striving to acquire knowledge, the student's personality, motivation, all these are intervening variables and influence the knowledge acquisition process. The learning process is, therefore, an individual process. The student acquires knowledge by contact with a rich environment which is meaningful to his prior knowledge. The student acquires knowledge by thinking. The constructivist approach distances itself from any attempt to direct and consolidate regularity in learning.

According to this approach the student must be supplied with a stimulus-rich environment. The learning process must be left to different influences – external influences from the environment or internal influences created by the student's thinking process. These influences are different in every individual and every student absorbs, collects and sorts the meanings which are suitable for him through his contact with the environment.

Direct contact with the environment is the basis for learning according to the constructivist approach. The open school tries to implement this approach. The role of the teacher in the open school is to allow learning to occur by light, careful and positive intervention which will enable the interaction between the student and the environment to take place and develop. The teacher promotes and creates a close and warm atmosphere and turns the student's attention to information which is meaningful to the learning process.

This approach distances itself from any attempt to formulate learning programs and strategies, except for the obligation to supply the appropriate conditions and environment and information which will enable and promote the existence of the learning process. Therefore the teacher's role also changes. It is less fixed and defined. It depends on the teacher's personality and considerations and on the student's feelings. Any action by the teacher is intended to meet the student's thinking and needs.

The constructivist approach aims to create an environment similar to the natural environment in which man discovered how to grow wheat. An education system that affords the student a learning environment which is compatible with his level and with the interest he exhibits in the learning subject, an open learning environment, a learning environment which enables the student to raise questions and find solutions to these questions by himself using a trial and error method, is a more effective method. However, this method is also more expensive, and requires more resources and manpower. A system which enables constant and free interaction with the teacher, with the learning environment and with the knowledge sources is a necessary condition for the realization of the constructivist approach. This is the way for ideal learning, learning which has its own rate and development stage. A learning system in which every student is afforded an environment which is challenging and suitable for him is very expensive, requires many resources that will enable each student to learn at his own rate.

The constructivist approach is based on a theory that claims that the motive driving the learning process is the person himself – the knowledge and interest, the talent and the student's motivation. A learning method according to the constructivist approach will supply every student with the means for expressing his ability. This should be achieved without pressure, without direction and without competition. Every individual will be judged against himself, and not against his friends' achievements. Every student will invest according to his ability, and the environment will stimulate him to express his latent potential.

The student's development is smooth and continuous and takes place one stage after the other, i.e. a continuous growth of knowledge. This method does not claim to understand and control the laws of learning and the process of the acquisition of knowledge by the student. The environment supplies the learning needs and learning takes place.

The constructivist approach comprised the basis for the evolution of teaching methods: active learning, experiential learning and self-regulated learning. These learning processes enable the student to cope with data taken from the situation with which he is faced and creates an interaction with the knowledge he accumulated in prior experiences. The person acquires knowledge via problem solving and shaping. Learning occurs by doing, where the student and the teacher are active. This is learning which involves the students in doing and in thinking about what they are doing. The students are encouraged to experience things and understand them beyond the basic facts. They are encouraged to analyze the ideas and carry out synthesis while working. This led to the development of curricula adapted to the student's fields of interest. The major characteristics of these curricula are:

- The student is involved in doing.
- Less emphasis is placed on transmitting information and more on the development of the student's skills and abilities.
- The student creates a connection between activities by discussion, writing and presenting.
- Emphasis is on the student. The student is afforded the possibility of researching the approaches and values he chooses.

Learning is not effective if it does not include the acquisition of knowledge while developing skills, experience and doing. Learning is a skill that includes training, instruction, trial and correction of errors while learning. These curricula emphasize the acknowledgement of the variance of the individual and his development as a student. An interpersonal relation evolves, which emphasizes the importance of democracy as an existential need in the individual's worldview. The goal is to achieve a balance between the needs of the individual and care for the society. The needs of the individual are expressed in the development of the ability to learn, the development of moral judgment, the development of personal abilities and the development of tools that will enable the individual to act and learn autonomously. The needs of the society are expressed in the development of tolerance, cooperation, social skills and improved interpersonal communication. The principles of learning are:

- Emphasizing human and warm relations between those participating in the learning and emphasizing the student's activity in directing independent learning.
- Developing a flexible curriculum which is adapted to the student.
- Developing and organizing materials and resources in the learning environment which stimulate learning and investigation.
- Using alternative teaching and a diversity of social designs for developing learning experiences.

Meaningful learning occurs via methods which develop activity that combines learning tasks, learning and education experiences that promote in-depth learning. Experience in social roles is very important in the moral and spiritual development of those experiencing the process.

These approaches are based on the assumption that meaningful learning occurs via the manner of experiencing or the personal experience or the teamwork. These teaching methods are based on maximal activity by the student. The teacher's role is to direct and organize the learning experience. It includes affective and psychomotor components which are integrated into a complete whole. This type of learning is suitable for the development of high levels of thinking and problem solving.

In his book *The Third Way*, Anthony Giddens analyzes the trends existing today in the Western World. Today the state passes as much authority and responsibility as possible to its citizens. In order to serve this goal, the education system must educate citizens who will be able to take their fate into their own hands. The goal of the education system is therefore to develop the student's ability, so that he will be able to fulfill his needs. Give the people the ability to solve their problems. Do not supply them with the solutions themselves! The education system thus needs to afford the student with the means and the ability by which he will be able to become integrated in society and will be able to earn a living for himself and his family.

According to this approach we can define the goals of learning, analyze the way and stages by which we will achieve these goals and finally judge the extent to which the goals were achieved. Giddens' approach is based on the assumption that we can direct, guide and lead the students towards goals that we shall define.

A similarity exists between the behaviorist approach and the approach of Hobbes, Luke and Hume. Their approach assumes that man is born a *tabula rasa* (clean slate) and all information reaches him from experiences and that the experience can be measured and directed and influenced. In contradistinction, the philosopher Immanuel Kant claims that a process of thinking can be found behind behavior. The individual's behavior is influenced and directed by the events that take place in his internal world, in his awareness and in his thought. These occurrences are not always understood and are not given to simple evaluation and measurement.

These are two approaches to learning with two methods for activating them. The policy makers must decide which method to use. This is an almost impossible task. It is difficult to use both methods. They are not always complementary and are sometimes even contradictory.

The constructivist approach will dictate freedom in the classroom, an open learning system, a stimulating learning environment, inquiry and experience. On the other hand, the behaviorist approach will demand discipline, presentation of goals, measurements, evaluation, judgment of the teacher and his achievements, constant evaluation of the student, giving grades, giving prizes and encouraging excellence, control of the learning process and making changes in it according to the students' performances.

Which approach will the education system choose? Do we have the tools to decide which method is preferable? Will we afford the teacher or the education system a tool by which they can make decisions in the field of teaching? When should one method be used and not the other? Can the accumulation of knowledge through research in the field of curricula help us improve the manner of collecting information, analyzing and making decisions?

#### JUDGING AND EVALUATING LEARNING PROCESSES

The school system must therefore judge itself. It must evaluate the education and teaching methods. An objective and clear scale which will enable the assessment of processes must be formulated using a systematic and reliable method that will enable measurement, evaluation, control and conclusions.

A system of concepts is required in order to succeed in measuring a process and reaching conclusions about the process. The chances of evaluating and reaching conclusions about a process improve when the process contains more defined and measurable concepts.

The learning process is a complex process. Not all variables that participate in this process are measurable. Not every variable in the field of learning and teaching can be evaluated objectively. Some variables are subjective and cannot be measured. Nonetheless, when discussing these fields we should strive to as clear and accurate a definition as possible of as many variables of the variables which take part in the process as possible.

Data collection, defining concepts and their accuracy is identical in all sciences. This process is essential for increasing knowledge in the field. The concepts become more accurate as more information is supplied by research. This process enables the development of knowledge.

In order to guide a learning process we must define different methods of teaching and the differences between them as well as which goals can be achieved by each method. A series of researches which we carried out during the 1980's helped us differentiate between different methods of activating technological systems in teaching.

The teaching methods are a factor that should be taken into account during the process of planning an effective curriculum (Offir, 2003b; 2004; 2005). A method of learning by computer may be considered along a spectrum. The "traditional method" is at one end of the spectrum and the "open method" is at the other end. The traditional method provides the student with information. The computer directs the student according to his particular level and in response to his answers. The learning processes and stages are well-defined and documented and all possible correct answers are known. In contradistinction, the open method enables the student to operate independently. The student can use different methods and applications in different subjects according to his interests, inclinations and abilities. The

computer serves as an information store, and a broad variety of information can be presented to the student.

The concepts of open and traditional methods can be regarded as being derived from the definition of the open and traditional classroom. There is no consistent way of defining open classrooms. Amabile (1983) cites several definitions or views. Openness can be viewed as "a style of teaching involving flexibility of space, student choice of activity, richness in learning materials, integration of curriculum areas, and more individual or small-group instruction", or one can emphasize the open classroom atmosphere as being conducive to developing curiosity, exploration and self-directed learning. This is in contrast to the traditional classrooms which are characterized by authoritative teaching, examinations, grading, large group instruction and strict adherence to curricula. We adopted Blitz's (1973) definition of the open classroom. This definition includes several fundamental characteristics:

- The open classroom stresses the need to be actively engaged with the environment as a means for achieving meaningful learning.
- The open classroom is tailored to individual interests and activities, and learning must therefore take place at an individual pace and style.
- The open classroom content is relevant to the student's environment, the environment is important in structuring learning.
- The open classroom is designed to be used according to the individual student's particular pace and learning style.
- The open classroom strives to achieve learning which is exciting and enjoyable.
- The open classroom promotes learning which is diagnostic, guiding and stimulating, rather than authoritative.

However, as Blitz points out, additional elements exist which may vary greatly, in addition to the above-mentioned basic characteristics. These elements depend on the teacher's particular philosophy and personality as well as on the teaching facilities.

Research supports the prediction that the open classroom contributes to creativity to a greater extent than the traditional classroom. Amabile (1983), in a review of studies performed on open classrooms and creativity, reports that most of the evidence supports the open classroom:

- The open classroom leads to a consistent and maintained superiority among children.
- Children in the open classroom achieve higher scores on creativity tests.
- The open classroom contributes to higher scores in open-ended tests.

The open method follows the same guidelines as the open classroom. The most frequent activities in the traditional classroom are group reading and mathematics drills. In contradistinction, the most frequent activities in the open classroom are creative writing, group projects and independent reading. The traditional method thus typically consists of drills and practice, whereas the open method usually includes tools for supporting creative writing and exploration.

The open method may include other elements, such as fewer or no constraints on performance, team effort, etc. The open classroom may be designed as such, or may be used in a special way as a standard tool. A database may be introduced as a useful tool in the technological world, and the student can learn to use this database, collect and analyze data (which are traditional applications of the tool). Alternately, the database may be used as a tool for exploration.

According to Schank and Farell (1990), the fastest way to change a traditional classroom into an open classroom is via computer technology. For example, the biggest obstacle to proposing and introducing new and original ideas is the fear of failure. The computer creates an environment which allows for failure and encourages originality. The student asks questions, formulates a hypothesis and examines the hypothesis using a computer simulation or database. The student can make as many mistakes as necessary in order to find the answer, without fear of embarrassment.

Previous research has shown that different variables influence the computer's effectiveness. It is assumed that the presence of certain factors will improve the effectiveness of the open classroom, whereas the presence of others will improve the effectiveness of the traditional classroom. The student's level of creativity should apparently be taken into consideration when using either the open or the traditional method (Crawford, 1988).

We carried out a research with a sample of 140 high school students aged 16 from two socio-economic statuses. Different methods were used as an integral part of their learning process. The different methods were divided according to their degree of openness. The open method was at one end of the scale and the traditional method was at the other end. Two research instruments were used in this study. The first was a 15-item questionnaire designed to examine the students' computer-related attitudes. The items form a general factor, i.e. computer-related attitudes, which explained at least 10% of the variance. The second instrument was the "Torrance Test" for measuring the "level of creativity". Socio-economic information about each participating student was also collected. Statistical analysis of the data did not indicate any significant differences between the "flexibility" and "originality" scores received by students from a low socio-economic status and those from a high status. However, significant differences were found in the scores obtained by the two groups in "fluency". Students from the high socio-economic group received higher scores than those from the low socio-economic status.

Significant differences were also found between the attitudes of these two groups of students towards the methods of using computers in the learning process. The students from the low socio-economic status preferred to work with word processors. In contradistinction, students from the high socio-economic status expressed no preference. The open method was preferred by students with high scores in "originality", whereas the traditional method was preferred by students with low scores in "originality". Thus, differences were found between the creativity level of students from a low and high socio-economic status. Furthermore, the students' level of creativity influenced their attitudes towards different teaching methods in the classroom.

Significant differences were found between three aspects of creativity: "fluency", "flexibility" and "originality". The level of originality was found to be significantly related to the type of computer method preferred by the students. The open method was preferred by students with a high level of originality, whereas the traditional method was preferred by students with a low level of originality. This correlation was found in both groups, i.e. those

with a low and those with a high socio-economic status. Significant differences were also found in the level of fluency of students with a low and a high socio-economic status, with those with a low status exhibiting a lower level of fluency.

The computer stores information in a databank and a wide variety of information may be presented to the student for analysis or for searching for information. The student needs to perform the difficult task of assessing the information and reaching conclusions. It may be concluded that the level of originality is significantly related to the student's attitude towards using technology during the process of learning. As the level of originality increases, so does the preference for an open method of learning. This conclusion is also supported for the level of fluency.

This connection between the level of originality and the preference for an open method of teaching is found only when the level of fluency is high. When this level is low, the influence of originality is not related to the preference for an open teaching method. When the level of originality is low, the influence of other factors, such as the student's attitude towards the subject matter, increases. The research results prove that the concept "open method" is not similar to the concept "traditional method".

Any method proves its effectiveness in different populations in order to achieve different goals. Each method has a different definition. This research helped us define concepts and understand the difference between them. As we define more concepts, we will have more tools to construct a model that will help us make decisions in the process of activating technological systems.

The introduction of instructional technology into the classroom has not always met with success. Technology itself is obviously just another tool with the potential to revolutionize the education system. However, a revolution in education will occur only if the correct educational steps and decisions are taken in order to ensure the successful introduction of technology into the learning and instruction process. Thus, decision making is a major variable that must be taken into account when deciding on the introduction of any instructional technology methodology into the education framework (Offir, 1993; 2000b).

#### THE TEACHER'S ROLE IN INSTRUCTIONAL TECHNOLOGY

The method for operating instructional technology depends mainly on the goal of learning. Instructional technology can contribute towards achieving the learning goal only when the teacher is not essential for learning. Decision making on the part of the teacher demands a clear definition of the different aims of learning and the different methods of operating a technological system in learning.

In the first phase of defining the goal of learning, different psychology theories, such as Bloom's (1956) taxonomy of the cognitive domain, can be used. The lowest cognitive level is knowledge. The students must remember and recall information. The highest level is evaluation. The students must be able to assess the value of a method for attaining a particular purpose.

Guilford (1959) presented a model for representing separate mental abilities which collectively form a map of structure, of intelligence. His theory can also be used for defining the method of teaching. One aspect of his model, called operations, is particularly relevant to

the present discussion. Guilford suggested that retrieval of information from storage in memory basically involves two kinds of operations, divergent production and convergent production. When divergence is high, or the level on Bloom's scale is high (for example, evaluation), the role of instructional technology is minimal since it cannot fully cope with this task given the present state of the art in software. The teacher must take the dominant role in teaching such a subject. However, when the subject matter is low in divergence (convergent) or is mainly a transfer of knowledge, instructional technology can replace the teacher to a higher extent. The teacher's role becomes less effective.

The accuracy of the goal definition increases as more information is collected by the system. Instructional technology can be operated by different methods for establishing different levels of teacher-student interactions. These levels have been found to affect the student's needs for the teacher's presence. These findings directed us in our effort to try and define different methods for operating instructional technology.

#### **OPERATING INSTRUCTIONAL TECHNOLOGY**

The teacher is an important factor in the introduction of instructional technology into the schools. Instructional technology will be used in a school if the teacher considers it to be a more effective method for achieving the aim of teaching. Two extreme methods can be found on the continuum of instructional technology methods. These extremes are the two extremes mentioned above, i.e. the open and the traditional method.

With the traditional method, the material and process of learning are well-defined. All possible answers are known and the aim of the method is solely to transfer information. The "teaching machine", for example, is a traditional method. In contradistinction, the "open method" aims to develop the student's way of thinking and his ability to analyze material, to draw conclusions and present original ideas. The teacher makes a different contribution to the process of learning in these two extreme methods of teaching. In the traditional method the teacher transfers knowledge, i.e. delivers information to the student. This task can be accomplished by lectures, books or other sources of information. In the open method the teacher asks the students open questions, discusses ideas and encourages the students to present their own original conclusions. The teacher's presence in the classroom is more important in the open method. This method of learning is suitable for achieving the aim of "evaluation".

The method of operating instructional technology should be directed by the goals of teaching. In the traditional method, when the aim of learning is mainly the transfer of information, instructional technology can control the learning process. In this situation the student can even obtain information at home via the Internet, television, satellite, book or fax. The effectiveness of the teacher-student interaction in this case is less critical.

In the open method, instructional technology should be operated so that it leaves the teacher in the classroom with enough space for expressing himself. The teacher-student interaction is important in this method. The student should have the opportunity to express his or her ideas and comments, whether the teacher is in the instructional technology studio or in the classroom.

The different systems for operating instructional technology can be classified according to the level of interaction which they enable, with each possessing advantages and disadvantages. Videoconferencing enables full interaction between the teacher and the student. Both can see and hear each other. Learning by satellite enables limited teacherstudent interactions. The students can see the teacher, but the teacher cannot see the students. The students can ask and answer questions mainly in writing. Television and fax enable unidirectional delivery of information. Therefore, the aim of learning will affect the method of learning. The method of learning will affect the type of instructional technology that will be used and this technology will affect the level and quality of the teacher-student interaction.

We found that the teacher-student interaction is an important factor affecting the level of learning. The interaction is important mainly in the open method. However, the instructional technology system is limited in constructing a teacher-student interaction. It may be concluded that when the importance of the teacher-student interaction increases, the importance of the teacher's presence in the classroom also increases.

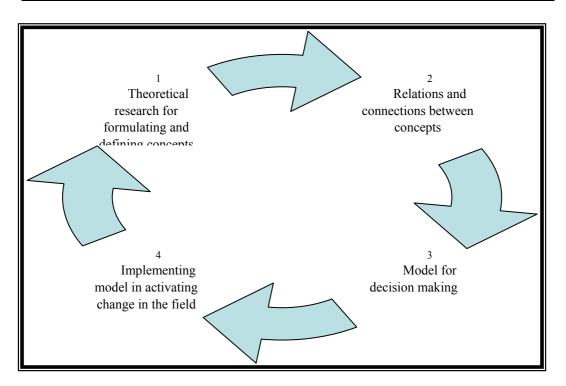
We have constructed a model for analyzing the messages in instructional technology. The validity and reliability of our research tool for analyzing interactions was developed as an integral part of operating our instructional technology system (Offir & Lev, 2000). During the past twenty years our research has directed us towards developing a model for decision making, i.e. where, when and how to use different instructional technology means for achieving defined aims of education (Katz & Offir, 1990; Offir, 1988; Offir & Katz, 1990a,b,c).

The teacher must be taken into account when generating a change in the learning and teaching system. The model which was consolidated enables the analysis of the goals of the learning process which will be achieved by the teacher and those which will be achieved by technological systems: Which teaching method is suitable for achieving a defined measurable and objective aim? Which teaching method is suitable for achieving a less defined and subjective aim?

According to this model, the teacher will concentrate mainly on imparting ability, developing thinking and tasks that involve individual attention. The technological systems will be operated mainly in order to impart information to the student. A combination and integration between technological systems and a flesh and blood teacher will lead to an improvement in the learning and teaching framework and will thus increase their effectiveness.

Research Process for Evaluating the Integration of Technologies in Learning

The process of integrating technological systems in teaching and learning thus requires a unique research strategy which includes four stages, as presented in the following diagram:



The concepts which we use for discussing and making decisions in the field of the "integration of technological systems in learning" are concepts that were created and formulated in various disciplines: psychology, sociology, curricula, business administration, economics, etc. Some of these concepts will be found to be relevant to the research and will be able to help in the process of assimilating and evaluating technological systems in learning. Other concepts will be found unsuitable for the assimilation and evaluation process. The theoretical research must locate the concepts relevant to the field. It affords tools for formulating and defining the goal which we wish to achieve with the learning process.

Defining the concepts and determining their use in the process of assimilating and evaluating technologies is the first stage, i.e. the theoretical research for formulating and defining concepts, some of which will be found relevant to the assimilation process. Some innovative technological systems will be found appropriate, and some will be found unsuitable.

The second stage of the research will deal in defining relations and connections between the concepts. For example, the research methods were defined on a continuum, where the "open method" is at one end and the "traditional method" is at the other end. Determining the methods on this continuum enables reference to methods found between these two extremes, i.e. an open method which also contains fixed objective elements, for example a simulation or model. The concept of "goal" in the second stage of our research can be found to have a relation and connection to different types of "goals". Different theories can be used in order to define the goals, as defined in the previous section.

The third stage will deal in the presentation of a model which will demonstrate complex relations between the various concepts and the relations between the teaching method and the goal which we are planning to achieve. The model will present the relation between the teaching method, the goal, and the teacher's contribution to the teaching process. The model that is constructed in the third stage is three-dimensional and presents relations between various variables.

The validity of the model which is constructed in the third stage will be examined in the fourth stage of the process, i.e. the implementation stage. This is the stage in which we activate the project in the field, make decisions according to the model which we designed and examine the model's effectiveness in improving the discussion and the decision making process. The model's effectiveness is examined by the extent of its contribution to improving the achievement of the project's aims.

# STAGES OF RESEARCH AND IMPLEMENTATION IN THE FIELD OF INSTRUCTIONAL TECHNOLOGY

Thus, the first stage of research is theoretical and attempts to understand the relations and effects between diverse variables. A major part of this research will we found to be non-relevant, i.e. no relations will be found between the variables. However, some variables with significant relations and effects will be found. These variables will comprise the basis for continuing research.

In our research the teacher was found to comprise a significant factor in the process of change. We therefore asked: Does a relation exist between any of the teacher's characteristics, personality or abilities and his willingness to carry out a change in his teaching habits? We carried out investigations in an attempt to discover significant relations between the teacher's willingness to assimilate change and his personality traits. The teacher's personality trait as a risk-taker was found to be significant. The teacher's risk-taker trait can predict his willingness to cope with change.

In the next stage of our research we tried to understand the relation between the level of the teacher's willingness to take risks (risk-taker) and his attitudes towards the use of technological systems in his teaching process. An investigation was carried out with the aim of examining the relations between the concepts and their contribution to the consolidation of a model that will help in the decision making process (Offir, 1990). Two questionnaires were designed especially for this study:

1. Risk-taking questionnaire: This questionnaire was comprised of 12 items and was intended for assessing the subjects' risk-taking level. A principal components analysis with VARIMAX rotation, designed to rotate the initial factors such that each item loads high on one factor but almost zero on others, yielded one significant factor comprising six questionnaire items (this factor had an eigenvalue greater than 1.0, indicating that it consisted of homogeneous content and explained at least 10% of the variance). A higher score in this aspect of the factor indicated that the subject indulged more in high level risk-taking behavior. The second aspect of the factor included these 6 risk-taking experiences where the subjects were asked to predict the number and levels of risk-taking behavior which they thought they would undertake in the future. A higher score on this aspect indicated a higher number and level of future risks

chosen. The compound risk-taking score was based on the cumulative score achieved on both of the risk-taking aspects.

2. Attitude to computers questionnaire: This questionnaire contained one significant factor (eigenvalue greater than 1.0 and explaining at least 10% of the variance) and was comprised of 13 items. The subjects were asked to state their computer-related attitudes on a scale from 1 (low level attitude) to 5 (high level attitude). The final score was based on a summation of the answers to all 13 items.

The subjects were divided into three levels of risk-taking according to their responses in the risk-taking questionnaire. Analysis of variance performed for the risk-taking attitude towards computer variables indicated the existence of a significant difference between the three risk-taking level groups. A *post hoc* Scheffe test, designed to examine the significance of inter-group differences resulting from the significant main effects was performed. The results indicated that high level risk-takers have a significantly more positive attitude towards computers than medium level risk-takers, who in turn have a significantly more positive attitude towards computers than low level risk-takers.

These results can help us make a decision regarding the suitable method for carrying out a change in the education system. Instructional technology is generally accepted by teachers as being an innovation in the field of education and instruction (Katz & Offir, 1988). When technology is introduced into the classroom, the teachers must carry out changes in their instructional and teaching methods in order to effectively accommodate instructional technology in the instructional process. However, any institutional, curriculum or instructional change is only as effective as the teaching staff's ability to implement the change effectively. The effectiveness of any staff development policy can only be considered in light of the manner in which the teachers respond to this policy. Understanding the teachers' behaviors and attitudes is therefore a prerequisite for the implementation of any reform in the teaching and instruction methods.

The relation between change and personality variables is well-known in society in general and in education in particular. When society feels that a fundamental change is necessary, it calls upon innovative leaders to initiate the necessary change. This is also the case in education (Duke, 1987). Technology is a new instructional aid and only those teachers whose personalities are comfortable with innovative teaching methods will respond positively to technology and will make a firm decision to use instructional technology in the classroom (Glasman & Nevo, 1988).

The results of our investigation clearly demonstrate that a positive attitude towards using technology in the classroom is significantly related to the personality trait of risk-taking. High level risk-takers are apparently more likely to be capable of making the transition to using instructional technology than are medium or low level risk-takers. The risk-taking trait appears to be necessary for the adoption of a new and innovative instructional and teaching method and can be used to predict successful introduction of novelty into the classroom. In contradistinction, low level risk-takers appear to reflect traditionalism in the teachers' perceptions of teaching methods. This impairs their ability to accept novel teaching methods in the classroom. These results confirm findings that certain personality traits are necessary for the successful implementation of novel situations in the classroom (Katz, 1984).

When use of instructional technology in the classroom is universally perceived as successful, all teachers (low, medium and high level risk-takers) will agree to implement this technology within their instructional activities. However, if instructional technology is perceived as only about 50% successful, then the low level risk-takers will agree to implement it in only about 10% of their teaching activities. Medium level risk-takers will use this technology in 50% of their instruction, whereas the high level risk-takers will use instructional technology in 80% of their teaching.

These results indicate that schools which are intent on introducing instructional technology into their classrooms should do so only with teachers who are classified as high level risk-takers. After instructional technology is demonstrated as being useful in teaching by these risk-taking teachers, it can be introduced into the classrooms of medium and low level risk-takers. These teachers will then be more inclined to accept instructional technology as a new, but not threatening, teaching method.

Further research should be carried out in order to examine whether the findings of this research are unique to the field of instructional technology.

With the development of the computer in the 1980's, many believed that the computer would become integrated in the teaching process and would comprise an effective tool. Some even predicted that it would replace the teacher. Many believed that the change in the education system would result from a flooding of the school with computers. The model which we consolidated as a result of our researches indicates and directs towards a different method for integrating technological systems in education:

- 1. The system must take the teacher's needs and wishes into account.
- 2. The teacher's personality is an important factor in his willingness to cope with new systems in teaching.
- 3. Success in using instructional technology systems in the past influences the teacher's attitudes towards the change.

A teacher who experienced more successful and tangible success in integrating instructional technology systems in the past will have less resistance to the integration of such systems in the classroom.

The model therefore guides us to the decision to integrate instructional technology systems in slow steps. If we wish to succeed, instructional technology should at first be introduced into a limited number of schools. Concentrating the efforts to a limited number of schools will increase the chances for success. Evidence of success will enable the introduction of more technological systems into the education process. Integration of technological systems in additional schools will be carried out only in places where the chances for success are high. Acting according to this model will lead to greater assurance of success. This method is contradictory to the method that was customary in the past, i.e. "flooding the school with technology" and stems from the research which we carried out and which is based on research results and models that were consolidated in order to make decisions based on these models.

# Assimilating Distance Learning as an Environment for Investigating the Generation of Changes in the Education System

We have been investigating distance learning systems since the 1990's. Activation of a distance learning system within the framework of the classroom is a change and an innovation which evoke responses by the teachers who must cope with this change.

Distance learning today is similar to the integration of computers in the teaching of the 1980's. Both are innovative systems which are supposed to become integrated within the teaching system and are under the authority and influence of the teacher. It is therefore natural that the teacher's attitudes and willingness to cooperate with this process of change are important and dominant variables that must be taken into account in the decision making process when deciding on the most suitable method for integrating technological systems in education.

Our research on distance learning, similarly to our research on the integration of computers in teaching, focuses on the investigation, measurement, assessment and becoming familiar with the process of change generated in the learning systems. It aims to elucidate factors which influence the teacher's place, role and status in the classroom. The evaluation and research process of distance learning systems indicated the importance of the concept "interaction", since in distance learning contact between the teacher and the student is not face to face. Rather, learning is carried out at a distance. The teacher is found in a distant location and the issue of contact with the student is significant, dominant and important.

Teacher-student interactions have always been regarded as a crucial variable for determining learning and attitudinal outcomes in distance learning environments. Many of the recent studies reviewed by Liaw and Huang (2000) are based on the "interaction quality hypothesis" (Trentin, 2000). These studies assume the existence of a direct correlation between the quality of the teacher-student interaction and positive learning outcomes. However, the term "quality" is not always defined. Furthermore, the "interaction quality hypothesis" is not always supported when empirically examining the nature of the relationship that exists between specific types of teacher-student interactions and learning outcomes.

Our interaction research therefore focused on the development and validation of instruments that can be used to analyze the content of verbal and nonverbal teacher-student interactions (Offir & Lev, 1999; 2000). These instruments enabled a comparison of teacher-student interaction patterns in conventional as opposed to distance learning environments. These interaction patterns were used to elucidate how interaction patterns change across contents, using content analysis for establishing empirical links between different types of interactions and student attitudinal outcomes (Offir, Lev & Barth, 2002).

A content analysis instrument which was developed and validated in our studies (Offir & Lev, 2000) was used in order to calculate the frequencies and ratios of different teacherstudent interaction categories. This instrument was based on Henri's (1992) content analysis framework, which was subsequently modified by Oliver and McLoughlin (1996). The instrument's coding scheme contained the following categories:

• Social interaction: Teachers interact with students in order to create social relations and support affective-motivational aspects of the learning process.

- Procedural interaction: Statements containing information about administrative and technical issues related to the lesson or course.
- Expository interaction: Statements which present knowledge content.
- Explanatory interaction: Teachers use the students' reactions to explain content.
- Cognitive task engagement: Teachers present a question or learning task which requires the students to engage in information processing.

Data generated by interaction content analysis was used online to supply the teacher with formative evaluation and data-based recommendations regarding his interaction management. After each lesson, the teacher received a "map" which reflected his use of interactions during the lesson. This objective feedback helped the teacher identify the specific interactions which should be increased or decreased during the next lesson. Ongoing research on the teacher's interactions enabled us to feed information directly "back into the loop" in order to help the teacher make more effective use of interactions.

However, when the number of students in the project increased, it became apparent that the question "effective for whom?" needed further investigation. We had a situation in which students with similar levels of ability (as indicated by their teachers' recommendations) and motivation (self-selection) were producing very diverse learning outcomes. We realized that instead of asking which interaction patterns were effective, we should establish the types of interaction that worked effectively for specific students. We would have to establish what works for which types of students and why. Only then could we use interaction analysis in order to help us design and deliver an effective distance learning program. We began to track the students' performance and learning outcomes in each lesson, so as to understand the diversity in learning outcomes. Multiple criteria were used to collect data on each student's learning outcomes.

As we shall see, when educational research moves from the laboratory to the field, it becomes almost impossible to isolate and control the numerous variables involved. Researchers who focus on content analysis can adopt a systematic and integrative approach in order to examine how person, process or product variables interrelate in a distance learning environment. Online data collection on the students' performance, combined with analysis of the teacher's interactions during each lesson, enable researchers to identify patterns of teacher interactions that support effective learning. Significant correlations between types of interaction patterns and positive learning outcomes can be immediately "fed back into the loop" in order to help teachers manage their interactions more effectively. The objective "map" generated by content analysis helps teachers understand which specific interactions they should increase or decrease. Use of content analysis as a formative evaluation tool, as opposed to summative evaluation at the end of the course, helps both researchers and practitioners maximize the potential of the distance learning environment.

Examination of the data indicated that the category "teacher's explanatory interactions" should be re-divided into the following three sub-categories:

• Learning assistance interactions: The teacher's explanatory interactions which are designed to facilitate the students' comprehension and retention of content. This sub-category contains the teacher's use of advance organizers, overviews and summaries, explicit definition of the lesson's objective and structure,

emphasis on the relevance of the target content and other teaching strategies designed to gain and maintain the student's attention. This sub-category was subsequently defined as a category in order to enable the user to differentiate between expository statements which present content and statements which facilitate students' information-processing (Offir & Barth, 2002).

- Superficial teacher feedback: The teacher's responses which do not contain an informative explanation of why the student's answer or comment is incorrect. Teachers' responses such as: "Incorrect, anyone else?" and "You are in the right direction, try again", can be included in this category.
- In-depth teacher feedback: In-depth explanatory feedback in response to students' questions and comments. Statements in this sub-category supply students with a detailed explanation of why their comments or answers are correct or incorrect. Teacher statements such as: "Your argument leaves out data relating to ..." or "You are assuming that a correlation necessarily implies causality", can be included in this category.

After the category of explanatory interactions was sub-divided, a significant positive correlation was found between the teachers' social interactions and learning assistance interactions. A larger correlation was found at the end of the teaching process than at the beginning.

One of the main aims of interaction content analysis research is to systematically observe and categorize types of teacher-student interactions in order to illuminate interaction patterns that might otherwise be overlooked (Stubbs & Delamont, 1986). Interaction content analysis helps researchers "tease apart" the essential elements of the interaction and investigate which interactions correlate with positive learning and attitudinal outcomes. Empirical examination of interaction patterns which exhibit a significant correlation with positive outcomes facilitates data-based decisions on the quality of the interaction and enables researchers to supply teachers with an effective formative evaluation.

For example, as the course progressed, the correlation between social interactions and learning assistance interactions increased. No such correlation was found at the beginning of the course, whereas later it became significant. The correlation increased even further during the last part of the learning process (Offir 2003a; 2006; 2007). However, no correlations were found between the other categories of teacher interactions. The increased correlation between social interactions and learning assistance interactions was accompanied by a concomitant increase in the percentage of students who confirmed that they understood the content of the unit.

The category of social interactions included all non-content related teacher statements that support motivational-affective aspects of the learning process. This category includes instructors' attempts to increase student confidence and mediate a feeling of competence. For example, statements such as: "Come on guys, this just looks complicated – when you begin using it you will see that you have already mastered much more complicated material", may be classified in this category. When teacher-related findings were correlated with student-related findings, a correlation of r=.9582 was found between the percentage of students who thought they understood the content of the unit and the teacher's social interactions. Furthermore, a correlation of r=.8357 was found between the percentage of students who indicated that they thought they understood the content of the unit and the teacher's learning

assistance interactions. The significant increase in the number of students who confirmed that they understood the content when the teacher's learning assistance interactions were correlated with the teacher's social interactions emphasizes three basic assumptions on which this study was based:

- 1. Non-intellective factors play a key role in determining the extent to which talented students realize their learning potential. The significance of a cluster of non-intellective factors was identified by Terman and Oden (1959) in their thirty year follow-up study of high IQ persons. Their study clearly indicates that traits such as persistence, integration towards goals, self-confidence and freedom from inferiority complexes differentiated between achieving and non-achieving persons. Feuerstein and Tannenbaum (1993) also examined the relationship between non-intellective dispositions and underachievement among highly talented students.
- 2. Students with a high learning potential who live in peripheral areas must be taught non-intellective dispositions together with the subject content. Acquisition of content alone, without these enabling dispositions, will not necessarily empower these students to maximize their full potential. High school students' participation in a university course within the framework of their own school creates a challenging but supportive learning environment that also focuses on the acquisition of enabling dispositions. In our project, one of the main functions of the on-site facilitator is to identify and prevent potential obstacles to effective learning. These obstacles have been reviewed by Tzuriel (1991). They include rapid loss of persistence in the face of failure, interpretation of errors as indicative of insufficient ability and expectation of future failure.
- 3. In a conventional learning environment, effective instructors constantly use verbal and nonverbal messages in order to encourage and reassure their students that they are capable of learning the material. In a distance learning environment, the students do not have access to the teacher's nonverbal expressions and gestures. According to Cookson and Chang (1995), distance learning instructors must compensate for the loss of this visual dimension. Our findings regarding concomitant improvements in students' self-evaluations of content comprehension when the teacher's social interactions correlate with learning assistance interactions reinforce this position.

This field project was designed to identify and enrich students whose psychometric scores and school grades did not reflect their high levels of learning potential. The student sample therefore did not contain the full range of ability levels. The findings regarding the correlation between types of teacher interactions and students' evaluation of content comprehension reinforce our previous findings (Offir & Lev, 2000). However, our conclusions remain limited to samples consisting of students with relatively high ability levels. Furthermore, our findings regarding the impact of teacher interactions remain limited to students' subjective feelings and attitudes towards content comprehension. Future research should focus on clarifying when and under what conditions the teachers' interactions significantly affect the students' objective scores.

Our findings regarding the use of objective feedback to modify teaching behaviors reflect previous research on feedback and behavior modifiability as reviewed by Mory (1996). Thus, our future research will focus on teacher-related variables which affect the modifiability of teacher interactions. Further research is also necessary in order to establish the extent to which teacher interactions influence student outcomes when the content's difficulty level is increased. We hope to include additional courses in future studies in order to investigate how teacher interactions affect student outcomes in a distance learning environment across subject-content areas and across varying levels of student ability.

In the first stage of our research we always try to define and examine concepts. This is how we defined the concept of interaction. The research helped us differentiate between different types of teacher-student interactions. Definition of the different types of interaction enabled investigation of the differences between the interactions that exist in a distance learning environment and interactions which exist in the traditional classroom. This will enable us to assess which interactions should be left to the teacher in the classroom and which interactions can be operated in a distance learning environment. The results of the research will help in the process of training teachers. The teacher in the classroom will fulfill the same tasks and will help achieve those goals of the curriculum which are difficult or impossible to achieve via distance learning systems. Defining the concepts helps us understand a model for operating a distance learning system as an auxiliary aid and will help the teacher's activity in the teaching and learning processes. The results of this investigation direct us to those interactions which are beneficial in different teaching situations.

#### CONCLUSIONS

The education system must achieve two main goals. One is to shape the student's behavior, to impart habits, to shape his attitudes, to increase his willingness to integrate in the society in which he lives and to contribute from his abilities and skills, to develop the student's thinking and motivation and to educate him to become an independent learner. The other goal is to impart as much knowledge as possible to the student. Some of these goals can be achieved by individual education and direction. When our goal is to transmit knowledge we will ask: What are the efforts? What is the investment which we are willing to make and which it is correct to invest in order to reach a defined achievement? In such systems the investment will be measurable and the products quantifiable.

When we activate a distance learning system we turn to a large population of students. Hundreds of students who are dispersed over the entire country can participate in a distance learning lesson. Here we will measure: What is the investment, the input? To what achievements have we led the students? What is the profit gained from activating the distance learning system, i.e. the output?

When referring to large populations of students, the considerations systems are not measurable. In such cases we refer to the mean achievements of the population. Decision making models are therefore models that measure and characterize quantitative products.

The overall goals of learning cannot be achieved solely by measurable and objective processes. There exist subjective variables such as the teacher's contribution, education to moral behavior, attitudes, thinking and learning ability, motivation and interest. These concepts are not measurable and cannot be judged in terms of objective measures. The cooperation, integration between the distance learning system and a flesh and blood teacher are parameters which must be taken into account when constructing and designing an education and teaching system. The quantitative concepts cannot be used solely to appropriately and accurately describe the learning process, and neither can the qualitative concepts.

Our research began with an attempt to define concepts taken from the field of the social sciences (psychology, sociology). Our aim was to quantify and define these concepts as accurately as possible. These concepts comprise basic elements for a model which helps us in decision making and deliberations processes.

The learning process can be described on a continuum which begins at a particular point and ends when the learning goals have been achieved. Parts of this continuum can be described using objective concepts. Parts of this continuum can be quantified. However, such a model does not afford an accurate picture since some of its parts cannot be defined in objective terms. This fact does not absolve us of attempting to consolidate models and trying and make them as accurate as possible. The most accurate picture is described using objective concepts. However, where objective concepts are limited, subjective concepts will be added in order to understand and describe the picture. This combination of objective and subjective concepts will present the true picture. Accurate mathematical models cannot describe and explain the decision making process. This process includes objective, concepts and try to define them as accurately as possible. We then try to use these concepts in order to predict processes.

This chapter detailed three fields which have developed during the years of this ongoing research. The first field enables definition of the method of operating technological systems according to the teaching goals. Different teaching goals were achieved using different methods (Offir, 1995; 2000). The second field helps determine a method for the assimilation of innovative systems within the education system. This method indicated that the preferable method for activating a "change" is to first activate it within a limited environment where it can be controlled and where its non-failure can be ensured. Success of this change reduces resistance to the change to a minimum and enables its growth and expansion (Offir, 1990). The third field in which our research concentrates is the field of teacher-student interactions. First we invested an effort in an attempt to define various teacher-student interactions. Then we attempted to examine which interactions are suitable for face to face teaching in the classroom and which interactions can be completed by distance learning (Offir, 1999; 2000c; 2002; 2003a, 2003b; 2004; 2005; 2006; 2007).

The results of this research help define the division of roles in the teaching process between the teacher and the technological system. It defines which teaching goals will be achieved by a flesh and blood teacher and which will be achieved by the technological system. The research process in these fields began with an attempt to define concepts and variables which are relevant to the learning process via technological systems. We then tried to find connections and influences between these variables and to present a model which will comprise a basis for analysis and decision making processes in this field.

There is general agreement that the field of learning and education is a most important field in modern society. Few will disagree with the assumption that a relation exists between the level of learning and the effectiveness of the education systems and the strength of the society in which we live. Shaping the behavior of tomorrow's citizens is carried out within today's schools. However, the learning and education systems must undergo a process of change in order to achieve these goals.

The fields of business administration, economics and sciences have models for collecting and classifying information. These models are directed towards and aid in the analysis of processes and making decisions. However, in spite of the great importance of the field of learning and education, this field has not adopted and assimilated these ways of thinking, discussion and analysis: How are changes generated in the learning systems? What are the mechanisms which promote change? What are the dominant variables in the process of consolidating models for evaluation and measurement of change in learning systems? These are subjects that have been investigated by our research team for approximately twenty-five years.

In contradistinction to economic and business systems, judgment and evaluation in the field of education and learning are mainly intuitive. Analysis of research data demonstrated that the integration of computerized systems within the school framework is a process of change. However, this process of change is not identical to processes of change in economical systems and therefore cannot use the same models.

A change in the field of learning systems must take many data that belong to the field of education and learning into account, including attitudes and the teachers' and students' ability and personality, psychological and sociological processes that accompany a process of change, the teaching method, the learning goals as well as cost versus effectiveness (Offir, 1999). However, these concepts are not accurate and it is difficult and complex to quantify them. The researches that were carried out and published by our research team deal in this issue. The complexity of the issue increases because economic models as a sole means are not effective. Nonetheless, knowledge on learning and education is also not sufficiently quantified to comprise a basis for analysis and decision making. We must find a method that will use an accurate quantitative model but will at the same time leave some room for subjective considerations.

These conclusions were reached by our team already in the 1980's, in light of the fact that computers were not generating the expected revolution in the education system. Since the 1980's the team has investigated and published articles dealing in the issue of the integration of technological systems in learning. The aim of these researches is to consolidate models for the assimilation of change in learning and education systems.

Computers and distance learning systems are only means which bring the change to the learning frameworks. Integration of these systems within learning systems creates a situation of change. Evaluation and research on the processes of assimilation of innovative technologies enables evaluation research and elucidates the process of change in learning systems.

The research method which was used is the "ongoing research". The experimental fields in which data are collected are the schools that decided to join the research and assimilate a change. The collected data comprise the basis for decisions for improving methods of operating the project in the following stages of its development. As the project progresses, the decision making model becomes more accurate and effective. The research process thus aspires to:

- Formulate a definition that is as accurate as possible of psychological and sociological concepts that appear to be relevant to the evaluation of the process of assimilation of innovative systems in teaching, according to the literature.
- Examine connections between variables and the accuracy of their definition by researches carried out within the framework of the project in the schools.
- Find relations between the variables. Relations and connections between the variables may help present a holistic model which will contribute to the decision making process.
- Confirm the model, i.e. examine the accuracy of the model in the decision making process.

The teaching and learning system cannot waive the teacher's contribution, the personal human contact between the student and the teacher. The research results indicate that the teacher's contribution is not fixed and changes according to the learning goals. The model defines those goals for which the teacher can use teaching means.

The research team is interested mainly in increasing the effectiveness of the teaching and learning system. The teacher is mainly a mediator. His contribution is mainly in those complex fields of education, development of thinking ability, crystallization of attitudes and moral behaviors in which the teacher's contribution as having a personal-human approach is critical. There is no technological system that can exchange the teacher and which can replace the personal human teacher-student contact. The fact that technological systems fulfill some of the teacher's functions (such as the transmission of information) enables the teacher in the classroom to pay greater attention to individual education in small groups. This enables the teacher to contribute from his ability and uniqueness as a person and as a teacher. With this approach the level of teaching and the transmission of information is carried out in a more professional and strict manner.

The process of the confirmation of the objectivity of the model is a process that has been taking place for years. It is based on the relation between the investment and the product. The model is more effective when the investment is smaller and the product is greater. Our aspiration is, of course, to make a small investment and achieve a high output. Investment when implementing a model in an educational learning environment is measured by concepts which have a varying value in terms of being quantifiable. For example, the investment includes the type of teacher involved in operating the system, the effort invested by the teacher, the teacher's willingness and ability, etc. On the other hand, the products include, for example, the student's motivation, his willingness to contribute to the environment in which he lives, his ability, his skill, the knowledge he acquired and the achievements he attained.

#### REFERENCES

Amabile, T.M. (1983). *The Social Psychology of Creativity*. New York: Springer Verlag.
Blitz, B. (1973). *The Open Classroom: Making it Work*. Boston: Allyn and Bacon.
Bloom, B.S. (1956). *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain*. New York: Longman.

- Cookson, P.S. & Chang, Y. (1995). The multidimensional audio conferencing classification system. *American Journal of Distance Education*, 9 (3), 18-35.
- Crawford, R. (1988). Computer and curriculum in 1999 computers in education. In: F. Lovis & E.D. Tagg (eds.) *Proceedings of the IFIP TC3, European Conference on Computers in Education*, pp. 645-650.
- Duke, D.L. (1987). School Leadership and Instructional Improvement. New York: Random House, pp. 40-47.
- Feuerstein, R. & Tannenbaum, A.J. (1993). Mediating the learning experiences of gifted underachievers. In: B. Wallace & H.B. Adams (eds.) Worldwide Perspective on the Gifted Disadvantaged. Natal: AB Academic Publishers.
- Giddens, A. (1999). *The Third Way: The Renewal of Social Democracy*. Melden MA: Polity Press.
- Glasman, S.L. & Nevo, D. (1988). *Evaluation in Decision Making*. Boston: Kluwer Academic Publishers, pp. 71-112.
- Guildford, J.P. (1959). Three faces of intelligence. American Psychologist, 14, 469-479.
- Henri, F. (1992). Computer conferencing and content analysis. In: A.R. Kaye (ed.) Collaborative Learning through Computer Conferencing: The Najaden Papers. New York: Springer, pp. 115-136.
- Katz, Y.J. & Offir, B. (1988). Computer oriented attitudes as a function of age in an elementary school population. *Computers in Education*. Elsevier Science Publishers, pp. 371-373.
- Katz, Y.J. & Offir, B. (1990a). Computer assisted instruction and students' social orientations. Proceedings of the 5<sup>th</sup> Jerusalem Conference on Information Technology. Los Alamitos CA IEEE: Computer Society Press, pp. 660-664.
- Katz, Y.J. & Offir, B. (1990b). Learning and teaching with computers: psychological and counseling considerations. *Educational Counseling*, 1 (2), 124-130.
- Katz, Y.J. & Offir, B. (1991). The relationship between personality and computer related attitudes of Israeli teachers. *Education and Computing*, 7, 249-252.
- Katz, Y.J. & Offir, B. (1993). Computer assisted learning and cognitive abilities: hardware and software planning implications. In: A. Knierzinger & M. Moser (eds.) *Informatics* and Changes in Learning. Linz, Austria: IST Press, pp. 11-13 (section I).
- Katz, Y.J. & Offir, B. (1994). Computer games as motivators for successful computer enduse. In: J. Wright & D. Benzie (eds.) *Exploring a New Partnership: Children Teachers and Technology*. Amsterdam: Elsevier Science Publishers, pp. 81-87.
- Liaw, S. & Huang, H. (2000). Enhancing interactivity in Web-based instruction: a review of the literature. *Educational Technology*, 40, 41-45.
- Mory, E.M. (1996). Feedback research. In: D.H. Jonassen (ed.) Handbook of Research for Educational Communications and Technology. New York: Simon and Schuster Macmillan, pp. 919-956.
- Offir, B. & Cohen-Fridel, S. (1998). Psychological factors in conducting an interactive distance learning system. *Teleteaching*, Edited by the Austrian Computer Society, pp. 779-787.
- Offir, B. & Katz, Y.J. (1990a). Learning-curve as a model for analyzing the cost-effectiveness of a training system. *Education and Computing*, 6 (1-2), 161-164.

- Offir, B. & Katz, Y.J. (1990b). Computer oriented attitudes as a function of risk-taking among Israeli elementary school teachers. *Journal of Computer Assisted Learning*, 6, 168-173.
- Offir, B. & Katz, Y.J. (1995). The teacher as initiator of change: fact or fiction. *Curriculum and Teaching*, 10 (1), 63-66.
- Offir, B. & Lev, J. (1999). Teacher-learner interaction in the process of operating D.L. (distance learning) system. *Educational Media International*, 36 (2), 132-138.
- Offir, B. & Lev, J. (2000a). Constructing an aid for evaluating teacher-learner interaction in distance learning. *Educational Media International*, 37 (2), 91-98.
- Offir, B. (1987). Application of psychology theory in computer-based instruction. *Educational Technology*, XXII (4), 47-49.
- Offir, B. (1993). C.A.I. as a factor in changing the self image of pre-school children. *Computer in Education*, Pedagogical and Psychological Implications, UNESCO Publishers, pp. 68-74.
- Offir, B. (2000b). Map for decision making in operating distance learning research results. *Educational Media International*, 37 (1), 8-15.
- Offir, B. (2003a). Analyzing of teacher-students interaction. Proceedings of the Fourth International Conference on Teacher Education – Teacher Education as a Social Mission – A Key to the Future. Achva College of Education, Mofet Institute.
- Offir, B. (2006). Influence of a distance learning environment on university students' attribution of success and failure factors. *Computers in Education Journal*, 16 (1), 82-94.
- Offir, B., Barth, I., Lev, J. & Shteinbok, A. (2005). Can interaction content analysis research contribute to distance learning. *Educational Media International*, 42 (2), 161-171.
- Offir, B., Bezalel-Rosenblatt, R. & Barth, I. (2007). Introverts, extroverts, and achievement in a distance learning environment. *The American Journal of Distance Education*, 21 (1).
- Offir, B., Golub, M.R. & Cohen-Fridel, S. (1993). Attitude towards courseware as a function of high schools' creative level. *Information and Changes in Learning*, Elsevier Science Publishers, 7, 211-218.
- Offir, B., Katz, Y.J. & Schmida, M. (1991). Do universities educate towards a change in teacher attitudes? The case of computer attitudes. *Education and Computing*, 7, 289-292.
- Offir, B., Lev, J., Harpaz, Y. & Barth, I. (2002). Using interaction content analysis instruments to access distance education. Special issue of the journal *Computer in Schools*, 18 (2/3), 27-42.
- Offir, B., Lev, J., Lev, Y., Barth, I. & Shteinbok, A. (2003b). Teacher-student interactions and learning outcomes in a distance environment. *Internet and Higher Education*, 6 (1), 65-75.
- Offir, B., Lev, Y. & Lev, Y. (2000c). Matrix for analyzing verbal and non-verbal teacherlearner interactions in distance learning. In: D. Benzie & D. Passey (eds.) *Educational* uses of Information and Communication Technology. Publishing House of the Electronics Industry, pp. 319-326.
- Offir, B., Lev. Y., Lev, Y., Barth, I. & Shteinbok, A. (2004). An integrated analysis of verbal and nonverbal interaction in conventional and distance learning environments. *Journal of Educational Computing Research*, 31 (2), 101-118.
- Oliver, R. & McLoughlin, C. (1996). An investigation of the nature and form of interaction in live interactive television. *ERIC Document* Number 396738.

- Schank, R.C. & Farell, R. (1990). *Creativity in Education: New Standard for Teaching with Computers*. New Haven: Yale University.
- Stubbs, E & Delamont, S. (1986) Explorations in classroom Observation, in . P. Croll (ED) Systematic classroom Observation (London . The falmer Press).
- Terman, L.M. & Oden, M.H. (1959). *Genetic Studies of Genius: The Gifted Group at Midlife*. Stanford CA: Stanford University Press.
- Trentin, G. (2000). The quality-interactivity relationship in distance education. *Educational Technology*, January/February, 17-27.
- Tzuriel, D. (1991). Cognitive modifiability, mediated learning experience and affectivemotivational processes: a transactional approach. In: R. Feuerstein, P.S. Klein & A.J. Tannenbaum (eds.) *Mediated Learning Experience: Theoretical, psychological and Learning Implications*. London: Freund Publishing House.

Chapter 45

# AN ERROR-MINIMIZING RULE FOR ANIMAL DECISION-MAKING UNDER UNCERTAINTY<sup>\*</sup>

## J.A. Devenport<sup>1</sup> and L.D. Devenport<sup>2</sup>

<sup>1</sup>University of Central Oklahoma, Edmond, Oklahoma, USA <sup>2</sup>University of Oklahoma, Norman, Oklahoma, USA

## ABSTRACT

Decision-making under uncertainty is to be expected in natural environments. The greatest source of uncertainty comes from the passage of time, because time-and the environmental variability it allows to proceed-discounts the reliability of information on which decisions are based. Information is most reliable if it can be acted on immediately, but as time passes, an average of past values of the alternatives is the best estimate of current value, since this subjective process matches the objective tendency of biological variables to regress to their means. The most optimal strategy, therefore, would be to flexibly shift from tracking the most recent outcomes to averaging across them. A model, the temporal weighting rule (TWR), accomplishes this transition. The output of TWR is a dynamic average whose rate matches the rate of environmental change. We review empirical studies showing the wide range of species that make dynamic foraging decisions consistent with TWR, the special predictions the model makes and their accuracy, its ecological relevance, and the memory mechanisms it appears to rely on. We conclude that this quantitative model and its accompanying decision rule, or something very similar to it, solve the one of the commonest problems animals face in their variable environments. TWR minimizes decision error made under uncertainty.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

### INTRODUCTION

In the natural world there are many reasons for incomplete information but one of the most important is variability. Change is inevitable and time is the fundamental cause. The quality of resources may degrade with time, and the quantity or availability of resources changes as competitors discover and utilize them, or natural factors, e.g., wind-dispersed seeds, alter conditions. So food resources that were once good will in time get worse but the reverse is also true; resources that were once poor will improve in time, as patches regenerate (Devenport, J. & Devenport, L. 1994).

Decisions made with incomplete information are said to be made under uncertainty. Most uncertainty results from not knowing the current probability or likelihood of a naturally varying outcome (reviewed in Dall, Giraldeau, Olsson, McNamara, & Stephens, 2005; Inglis, 2000). A decision-maker has incomplete information because it is uncertain whether the current state matches the previously sampled state.

Much of what we know about decisions under uncertainty comes from investigations of foraging decisions in animals. When an animal is foraging for food it often visits locations or patches that it has sampled in the past and therefore has information about the quality or quantity of food found in those patches. However, because natural resources vary, the difficulty of knowing the *current state* of a patch contributes to the uncertainty of information about it. We will argue that natural selection has equipped animals with a cognitive adaptation to deal with this problem.

Just as time leads to change, it necessarily discounts the reliability of information about the comparative value of alternatives, so uncertainty is dynamic. The degree of uncertainty is correlated with the amount of time that has passed. For example, a bird that finds a ripe fruit tree is likely to find fruit is still available an hour later, but as returns are postponed, the probability of fruit remaining declines progressively owing to harvesting by competitors. Correspondingly, the reliability of information about that alternative is diminished and should be discounted. This relationship between time and information reliability obtains for all resources or options that have some variability, which in the natural world is the rule (Stephens, 1989). What differs is the rate of change of resources (Devenport, L., & Devenport, J., 1994).

Regardless of the exact rate of change, choices made shortly after the information is obtained are (almost) completely deterministic. If a squirrel finds a nut tree and returns an hour later, it is a near certainty that there will still be nuts to harvest. A honey bee returning to a flower patch in 15 min will still find nectar or pollen. In such cases, where information is fresh, keeping track of where you have just been and repeating that choice provides the best outcome, provided the animal has not itself depleted the resource. There are many examples of animals tracking resources in both natural and laboratory settings (Bateson, 2003; Cowie, 1977; Stephens & Anderson, 2001). But this strategy, termed *tracking*, only works when information is recent and therefore predictive of current conditions. Delineating what is recent or fresh information is difficult because this is a *relative* not absolute measure and varies with the type of resource and prevailing conditions.

Information that is not recently acquired is more uncertain and less reliable because there has been a greater opportunity for change to occur, but knowledge acquired in the past still has worth because the long-term average value of the alternative can be used to predict its

current condition, albeit with some added error. This strategy, termed *averaging*, is based on the assumption that variability has regularity or form, namely regression-to-the-mean (Devenport, L., 1998; Stephens, 1989). For instance, a dandelion plant produces flowers and seeds in succession so at any given time, the presence of seeds may be uncertain (depending on when the last assessment occurred). However, larger or healthier plants or those growing in more favorable microclimates will produce more seeds more rapidly. Ideally, a forager visiting that plant will be able to valuate the quality of that plant or *patch* as higher on average than another.

Because resources vary with time, good patches become bad, and bad ones become good. Even though the last visit to a patch might have resulted in no food, the forager could assume that the plant will regress to its mean (good) value in the long run so that when information has not been recently up-dated, choosing the alternative with the higher average value will give better returns than decisions based on tracking or random foraging. There is evidence that many foragers use averaging to assist with their foraging decisions (Cuthill, Haccou, & Kacelnik, 1994; Reboreda & Kacelnik, 1992), but just as with tracking, exclusive use of an averaging strategy will not result in the best use of information because even a typically poor patch can be temporarily good. This implies that individuals who use a combination of both strategies would have an advantage. However, flexibly deploying two different strategies requires a mechanism for transitioning from one to another.

Given the importance of decision-making and the inevitable effect of time on virtually all resources, it would be surprising if selective forces had not adapted animals to optimize decisions under uncertainty (Inglis, 2000). Such an adaptation would need to give an output of estimated patch values corresponding to the rate of environmental change. This would allow for a match between the reliability of information and the temporal changes in the real world. Because change unfolds in time, time itself could be the mechanism that guides the decision-making strategy.

We suggest that the best regulator of the progression from non-probabilistic to probabilistic outcomes would be an average obtained from patch experiences weighted by their temporal recency, i.e., how much time has passed since the visit. Such a weighted average is dynamic and would permit flexible shifts from tracking (using the most recent information) to averaging (regression-to-the-mean). Because time is compared as a ratio (Gibbon, 1977), the relative recency of information about alternatives could act as a weighting factor, automatically giving less weight to less recent experiences. Such an adaptation would require interval time perception and a decision rule.

Time is a basic element of the world and therefore perception of temporal durations or intervals is common throughout the animal kingdom. Interval time perception has been demonstrated in many species including birds, mammals, insects, fish and at least one reptile (reviewed in: Lejeune & Weardern, 1991; Matell & Mech, 2000) and some basic timing processes exist in many if not most animal species (reviewed in Gallistel, 1990).

#### Dynamic Averaging

We have developed a dynamic averaging model based on the functional properties of the environment and the psychological mechanisms of timing. The temporal weighing rule (TWR) model proposes that the value of an option whose objective value is known but variable, is weighted by the relative recency of information about that option and thus when information is recent, it is most heavily weighted, returning a high probability in favor of the last best alternative (Devenport, L., 1998). When information is not recent, the weighted average regresses to the mean and returns a value equivalent to the absolute long-term average. TWR is the only decision model that accomplishes this transition.

More formally, the output of the model is a weighted average of the subjective quality,  $Q_i$ , of an alternative weighted by it recency 1/T, which is the time between previous experiences and when the decision is being made. The temporally weighted value of a given patch, e.g., x, would be represented as:

$$\boldsymbol{\mathcal{V}}_{wx} = \sum_{n=1}^{n=i} \left[ \boldsymbol{\mathcal{Q}}_i \left( \frac{1}{T_i} \right) \right] \div \sum_{n=1}^{n=i} \left( \frac{1}{T_i} \right)$$

which is the sum of the *n* set of x patch quality experiences,  $Q_i$ , weighted by their respective recencies and then divided by the sum of the recencies. When choices need to be made between two alternatives, e.g., patches A and B, the weighted averages would be compared as a ratio. To illustrate, let the individual values of patch A =  $v_{wa}$  and patch B =  $v_{wb}$ . The *relative* weighted value of patch B ( $V_{wB}$ ) is simply a proportion.

$$V_{\rm wB} = v_{\rm wb}/(v_{\rm wa} + v_{\rm wb})$$
, where  $V_{\rm wA} = 1 - V_{\rm wB}$ .

If the two weighted values are not discriminably different, choices for the two patches should be evenly divided. If the relative value of one patch is substantially greater than the other, choices should strongly favor that alternative (see Appendices in Devenport, L.,1998 and Devenport, L., & Devenport, J., 1994 for more details).

#### **GENERAL EXPERIMENTAL APPROACH**

The TWR model is based on inevitable resource variability and delays between when information is gathered and acted on. To test the model, animals were given foraging experiences in two experimental food patches that varied in availability and quality, such that animals experienced one patch that was good (full) when the other was bad (empty) and then after some time, conditions reversed. In some cases, the two patches were equal in overall (average) value and in some cases, they differed. When they differed, the patch with the higher value was experienced first. A third patch, that was never baited, was often provided to assess memory (see *Memory Mechanisms* below). Tests occurred either when information was fresh or after a delay, using independent groups. Patches never contained food during the choice test.

As an example of our methods and procedures, the initial observations and data for development of the model came from a field study of wild populations of least chipmunks and golden-mantled ground squirrels. We erected feeder stands in open meadows where these ground-dwelling rodents forage for seeds. We permitted free-ranging, tagged animals to visit a feeder that contained sunflower seeds. After a predetermined time, conditions changed so that the previously baited feeder was empty and another feeder contained seeds. Simultaneously, animals were given an equivalent amount of experience visiting the second feeder (patch) that varied inversely. When the first patch was good, it was bad and vice versa. We then imposed a delay to simulate normal interruptions in foraging that lasted either 1 or 48 hr, followed by a single test trial where the original feeders were again presented but were empty. Other studies were carried out with different animals in the field or laboratory, with variations in most key variables, but this example is typical and the simplest.

### **SPECIES SIMILARITIES IN DECISION-MAKING**

One of our goals has been to investigate the generality of cognitive adaptations for decisions under uncertainty, especially those caused by temporal variation. Variability is one of the few constants in nature. Regardless of habitat or phyla, most animals are very likely to experience variable resources and interruptions in the flow of information and the resultant uncertainty. Therefore, adaptations for dealing with uncertainty should be common among species.

In our initial study we found no species differences between least chipmunks, *Tamias minimus* and golden-mantled ground squirrels, *Spermophilus lateralis* (Devenport, L., & Devenport, J., 1994). Both species made decisions consistent with TWR, choosing the most recently good patch when information was recent, but averaging when information was not recent. The species similarity could have been attributable either to relatedness (same family, *Sicuridae*) or similar selective pressures from a shared environment. We and others have subsequently studied several unrelated species that do not share evolutionary histories or similar environments.

Two laboratory species, rats (*Rattus norvegicus*) and pigeons (*Columba liva*), make an especially interesting comparison because, reared in stable environments, both have had no previous exposure to the extreme fluctuations in resources that might produce sensitivity to variability or uncertainty. Nevertheless, both species showed dynamic decisions (Devenport, L., Hill, Wilson, & Ogden, 1997; Mazur 1995; 1996) that were indistinguishable from those of wild populations of sciurids. They favored the last good outcome when choices were relatively recent, but progressively relied more on patch averages with increasing delays.

One reason for the consistency among these divergent species could be the similarity in type of diet. All primarily consume non-renewable foods that are high in energy content and spatially and temporally variable. The similarity in diet might especially favor a cognitive adaptation to deal with uncertainty, so we extended our investigation to species with very different types of diets: carnivores and herbivores.

Although most domestic dogs do not experience a strictly carnivorous diet, they have retained many of the characteristics of their wolf ancestors, especially foraging behaviors (Fox, 1965). We tested sporting and working breed pets in their home yards. We found that choices were time-dependent, with all dogs choosing the most recently good patch when it immediately followed training experiences, but with only half of the dogs choosing the last good patch when neither experience was recent, and weighted averages would have been equivalent (Devenport, J., & Devenport, L., 1993).

Herbivores have diets even more dissimilar from those of all the previously mentioned species; they consume foods low in caloric value, more homogeneously distributed, and that renew rather quickly, factors that could affect the role of uncertainty in foraging decisions. Despite these differences, herbivores make dynamic choices (Bailey & Sims, 1998). Cattle (*Bos spp.*) given daily trials foraging in an 8-arm radial maze baited with food (alfalfa) of moderate quality, followed by trials where two arms contained either preferred (grain) or non-preferred (straw) food, learned the locations and visited grain arms first and avoided straw arms. Following a 30 day interruption, visits to the non-preferred arms recovered as the higher values from the first trials were averaged in, consistent with TWR.

Similarly, quarter horses given variable patch experiences in a familiar pasture or arena found either patches of equal or unequal average value (independent groups). As predicted, horses made time-dependent decisions. When there was fresh information, regardless of the absolute value, all chose the most recently good patch but when tests were delayed so that no information was recent, patch choices matched unweighted patch averages (Devenport, J., Patterson, & Devenport, L., 2005).

Taken together, these studies show a remarkable similarity in decision-making across of a wide variety of species and all are consistent with predictions of TWR. (Also see Applications and Implications below for an insect example).

### **PREDICTIONS OF THE TWR MODEL**

If a model is to functionally represent the environment, it should make accurate predictions based on the relevant properties of the environment. Several different variables have been studied, including the rate of patch change, the effects of different patch values, causes of variability in patch values, the nature of the transition from good to poor patch values, and the influence of non-food variables. We then compared decisions to TWR model predictions

#### Rate of Patch Change

It is obvious that different resources change at different rates. Some, e.g., pooled rainwater, can change quickly whereas others, e.g., nut trees, change slowly. The rate at which patch conditions change should alter the rate of regression from tracking to averaging and thus effect foraging decisions. We and others have assessed a wide range of rate changes in a number of ways.

When patches change in minutes to hours, the regression from tracking to averaging should shift relatively quickly. Calculations using of the midpoint of the durations of the good and bad states of patches were used to determine test times when weighted averages for the last good patch were very high or equivalent to the unweighted average. Animals overwhelmingly chose the patch with the higher recent value (tracking) at the early test and distributed choices equally between the two alternatives (averaging) at the delayed test (Devenport, L., & Devenport, J., 1994).

Using longer time-frames, independent groups experienced much slower rates of change, for example, where each patch was good for 4 hours and bad for 4 hours, separated by 32 hr. Using model predictions, we selected test times for high, medium, and low weighted averages for the last baited patch and found dynamic choices, with preference for the last, most recently good patch extended about 20 times longer than when the two states were temporally more contiguous and changes occurred more quickly (as described above).

If resources change very rapidly, e.g., on a scale of seconds to minutes, weighted averages should regress very quickly to unweighted average values. We tested this in the laboratory where animals made repeated visits to variable patches, and the pace of trials was controlled so that each patch state of the experiment was 20 min in length and independent groups were tested at a range of delay intervals. Animals made choices that were consistent with the weighted averages: when tested soon after experiencing the baited patches, all chose the most recently baited patch, but after a 4 hr delay, test choices mirrored unweighted averages (Devenport, L., et al., 1997).

#### Transition From Tracking to Averaging

The studies described above confirm another prediction of the model: the transition from tracking recent trends to relying on an average should be a smooth progression. We systematically chose test times that generated weighted averages that were high to very low and found that when relative values were above 0.6, groups of animals overwhelmingly choose that alternative, whereas when weighted estimates were less that 0.4, animals overwhelmingly preferred the other alternative. For patches with weighted values between 0.4 - 0.6, animals were indifferent (Devenport, L., et al., 1997). A smooth transition from tracking to averaging mirrors resource changes which are typical of the natural environment.

#### **Different Patch Values**

In nature there is considerable variability in quality between patches. We noticed for example, that most patches that chipmunks encounter are spatially and temporally variable but that some have higher productivity than others. Optimal foraging theory predicts that animals should select the patch with the absolute higher value (Charnov, 1976; Cowie, 1977), but as we have shown, time alters objective value. The TWR model predicts that patch choices should favor the most recent patch at short delays, *regardless* of absolute values in patch quality, but that at delayed tests there can be a complete preference reversal depending on whether the two patches being considered were of equal or unequal average quality.

In experiments where unweighted patch values were equal, choices split 50-50 when there was no recent information, paralleling the similar unweighted averages. However when the two patches had different unweighted values, e.g., one yielded for a longer time (Devenport, L., et al., 1997) or animals got more food at the first of two patches (Devenport, L., 1998; Devenport, J., et al., 2005) there was a complete reversal of choices as the patches regressed to different unweighted means. It is important to note that we are referring to groups of animals tested at different delays, and not individuals reversing preferences.

#### Quantification of Patch Averages

Natural patches can vary in a number of ways. Some patches, e.g., fungi, are edible only for a short time before they begin to degrade, making duration of yield the metric of patch value. Some patches vary in the amount of food an animal can extract on each visit. For instance, hoarding animals make multiple trips to a patch to cache food (reviewed in VanderWall, 1990). In that case, the run of good visits is a measure of patch value. Fast changing patches, e.g., seeds in wind shadows, can vary in the probability that conditions are good when the animal visits.

We have tested the importance of these patch metrics in a series of studies and found that all produce dynamic decisions. Whether we varied the amount of food available on each trial but held duration of yield constant (Devenport, J., et al., 2005) or held the amount of food available at each patch constant and varied the number of visits to a patch (Devenport, L., et al., 1997), choices were always time-dependent.

Probabilistic changes in patches have been investigated in operant paradigms where responses were reinforced on variable-interval schedules and the probability of reward was equal for several sessions, followed by several sessions where one key "patch" was substantially better than the other. In other conditions, the sequences were reversed. Regardless, choices were time-dependent, regressing to the unweighted average of past sessions following the 24 hr interruption between daily sessions but returning to the more recent averages within sessions as new information was acquired. (Mazur, 1995; 1996).

These results from probabilistic rewards mirror those from a field study where two patches were probabilistically baited for five days and one patch had a substantially higher probability of yield. Test choices were again dynamic, with animals choosing the last good patch irrespective of its unweighted mean at the early test, but reverting to averaging at the late test (Devenport, L., & Devenport, J., 1994).

#### TWR and Non-foraging Variables

Decisions are rarely made solely on the basis of one dimension and are usually a trade-off between two or more conflicting variables. Predation risk is one of the most important non-food related influences on foraging behavior (reviewed in Lima & Dill, 1990) and many foraging decisions reflect a compromise between patch quality and predation risk. This raises the question of whether variables like risk are treated like patch quality and could figure into a dynamic average.

Winterrowd & Devenport (2004) tested the predictions of TWR in a field study by giving animals a choice between patches that varied, but that had equivalent average food quality. Half of the patches (safe) were close to experimental rock piles where animals could find cover and half were in the open (risky). Perceived safety was independently quantified in groups that did not experience variable patches. Animals treated patch safety in much the same way as patch food value. Most chose the risky patch if it had yielded food relatively recently but when experience was not recent, risk took precedence and most chose the safe patch even though average food values had been equivalent. The TWR model easily accommodated safety as an added dimension of patch quality. In much the same way, TWR has been used to predict sampling and exploration. Model simulations that incorporated hunger, patch location uncertainty, reward expectation, and reward uncertainty found TWR to be the best predictor of when a forager should seek information, i.e., sample, vs. spend time foraging (Inglis, Langton, Forkman, B., & Lazarus, 2001).

Grazing distribution patterns can also be predicted using a modification of TWR that incorporates physical and mechanical factors such as rate of turning, speed of movement, and abiotic factors such as slope of ground and location of other important resources, e.g., water or shade (Bailey, Gross, Laca, Rittenhouse, Coughenour, Swift, & Sims, 1996). Taken together these studies demonstrate that TWR can flexibly accept several different variables thought to be important influences on decisions, without modifying the basic structure of the model.

## THE ECOLOGICAL VALIDITY OF TWR

Adaptations are the result of selective pressures for solving specific "problems" and in the case of foraging, interruptions and variable food resources are among them. If interruptions are common, foragers should be able to adjust to them. It would also be adaptive if foragers not only kept track of the rate at which external factors brought about change in food resources but were aware of the effect of their own behavior on resource variability. Below we review evidence showing that these conditions have a substantial impact on foraging behavior and discuss their importance in dynamic averaging models like TWR.

#### Interruptions in Foraging

The importance of time in understanding foraging is well recognized. From the first, optimal foraging models incorporated time into the search or discovery phase, in calculating the rate at which food was harvested and thus the giving up time (Charnov, 1976), and in the time cost involved in processing or handling prey items which determined whether prey items should be included in the diet or avoided (reviewed in Bateson, 2003; Shettleworth, 1998). But one aspect of time that has been largely ignored is that as animals forage, they often encounter unexpected interruptions in their foraging activities. Those interruptions arise from a variety of factors including unexpected events such as weather changes, the appearance of competitors, the need to escape or hide from predators, or to give parental care. There are also predictable interruptions in foraging, such as daily fluctuations of darkness or tides that suspend activities. For most animals interruptions in foraging probably occur very frequently. Whether anticipated or predictable, familiar or novel, interruptions all have the same effect on the reliability of information as discussed above. Models, like TWR, that incorporate interruptions are likely to be better predictors of foraging behavior.

#### Other Sources of Change

All of the studies described thus far have involved changes in the environment caused by forces external to the forager but the forager itself also contributes to changes in patch values by depleting patches as it forages and the objective value of the patch diminishes as prey are harvested. In order to forage optimally, it seems likely that animals should be able to take account of those changes in their patch valuations.

We tested whether animals were able to represent changes in patch value resulting from their own foraging by allowing chipmunks to partially deplete one of two buried seed patches. The following day all of the animals chose the patch from which fewer seeds had been taken. Even when we controlled the amount of depletion each animal was able to accomplish, all still chose the patch that they had least depleted. Animals refused to revisit fully depleted patches and would only do so if given an external signal of patch renewal (Devenport, L., Humphries, & Devenport, J., 1998). These experiments demonstrate that least one species does take account of their own foraging activities and discounts patch values based on its own behavior, representing resources in terms of future rather than past value. The only model to take self-depletion effects on future value into account is TWR.

#### **MECHANISMS AND ALTERNATIVE INTERPRETATIONS**

#### Memory Mechanisms

The temporal weighting rule incorporates three main variables: 1) patch item value, *what* type and quality of food items were discovered, 2) instances of patch quality changes, i.e. *when* patches yielded best, moderately, or worst, and 3), spatial location, *where* patches were found. It might be wondered if animals are capable of acquiring and retaining such rich representations of their foraging experience.

That many animals use sophisticated spatial representations is no longer doubted (Jacobs & Liman, 1991; Kamil & Balda, 1985; MacDonald, 1997). This generalization also applies to the wild animals we have used extensively in our studies (Devenport, J., & Devenport, L.,1994; Devenport, L., & Devenport, J., 1994). Of course, memory for the comparative quality and quantity of food items has long been established in the "preference" and "choice" literature and is taken for granted in the design of studies aimed at investigating higher order cognitive processes, such as "chunking" (Cook, Brown, & Riley, 1985), "transitivity" (Bond, Kamil, & Balda, 2003; Roberts & Phelps, 1994) and selective cache recovery (reviewed in Vander Wall, 1990).

To this list, TWR adds patch variability, because stability in nature is considered to be exceptional (Devenport, L., 1983; 1989; Stephens, 1989). Olton (1979) was among the first of many (reviewed in Shettleworth, 1998) to show that the absence or removal of reward is at least as memorable as the presence of reward. In fact, animals can transition flexibly between using (retrospective) memory of patches that have been visited and emptied, to using (prospective) memory of patches that still remain good in order to efficiently solve a harvesting problem involving 8-19 separate patches (Cook et al. 1985; Zentall, Steirn, & Jackson-Smith, 1990). The transition from the use of empty vs. full patches to guide choices

is evidently based on which requires fewer cognitive resources (i.e., retention of fewest items in working memory). This research obviously bears on temporal memory as well. TWR, as we have formulated it, requires retrospective and, possibly, prospective memories.

Further evidence for the spontaneous encoding and use of time-based memories is now well documented (Cook, et al., 1985; Emory & Clayton, 2001) Research in the last quarter of the 20<sup>th</sup> century worked out important details of interval timing and the types of memory i.e., working and reference—that enable it (Church & Gibbon, 1982; Gibbon, 1977). Recent studies (Nagshbandi & Roberts 2006; Raby, Alexis, Dickinson & Clayton, 2007) answer lingering skepticism about how long retrospective memories are retained and whether an animal's prospective representation of time could in any sense be used to guide behavior farther into the future than a few seconds or minutes (Roberts, 2002; Suddendorf & Busby, 2003). For example, after having learned that they will be restricted to a particular area for their morning foraging sessions, western scrub jays selectively provisioned this area during a free-food session on the preceding afternoon (Raby, Alexis, Dickinson & Clayton, 2007). Animals not only selectively pre-stocked this area with extra food when it otherwise promised to be empty, they assured a variety of food by caching a second type of food item when only a single type would ordinarily be present. Importantly, this "planning" behavior appeared in the absence of a motivational state specific to the plan (i.e., preparing for the eventual absence of food when it was still present), assuring that the behavior was not reflexive (Raby, Alexis, Dickinson & Clayton, 2007).

The remaining question is whether or not animals can simultaneously incorporate what, when and where information into their patch valuations and choices—i.e., use declarative memory (Tulving, 1983). Work by Clayton and her colleagues bears directly on this question. They have shown that scrub jays use time, place, and item type to decide where to search for food. For example, birds choose to harvest first a preferred food type (meal worms) over a somewhat less preferred type (peanuts) that they have previously cached. This in itself is not surprising. But if the opportunity to forage is considerably delayed, a different pattern appears: the animals now prefer to search in areas where the peanuts were previously cached and the meal worm caches are avoided (although none of the patches actually contained food items at test). This selective change in preference over time is related to the more rapid spoilage of the insects compared to the seeds. This is one in a series of studies showing that jays readily incorporate what (meal worms vs. peanuts), when (delay between discovery and subsequent test trials), and where (the spatially unique, hidden patches where the two food items are found) information into their foraging decisions (Clayton, Yu, & Dickinson, 2003; de Kort, Dickenson & Clayton 2005). Further evidence for the flexible use of declarative memory is found in the reaction of caching birds (Emory & Clayton, 2001) and rodents (Preston & Jacobs, 2001) to potential robbery by onlookers and in other complex decisionmaking tasks (Bird, Roberts, Abroms, Kit, & Crupi, 2003; Grosenick, Clement, & Fernald, 2007; Zentall, 2005). We think that similarly rich representations of experience are probably involved in the implementation of TWR.

#### Other Mechanisms

TWR is primarily a rule that that guides optimal decision-making over a continuum of near-certain to uncertain information. Because time-, quality-, and place-tagging of item

information in memory appears to be an automatic and relatively effortless process (reviewed in Gallistel, 1990; see earlier discussion above), TWR was developed with those processes in mind. It will be recalled that there are no memory variables in the model, only patch quality, time, and place. Thus, the quantitative model of TWR is silent on the question of how it is implemented. The output of the model, which we have referred to as a dynamic average, could be executed by mechanisms other than declarative memory without altering the validity of the model. In this regard, "spontaneous recovery" phenomena are similar to dynamic averaging. Spontaneous recovery has historically been attributed to non-cognitive mechanisms (reviewed in Rescorla, 2004). It is therefore worth considering if this mechanism might qualify as an explanation of dynamic averaging, or if a more cognitive interpretation better explains both dynamic averaging and spontaneous recovery.

Spontaneous recovery reliably appears after an animal has been rewarded to perform some behavioral response and is then extinguished, i.e. by withholding reward until it no longer responds. Placed back into the training context at some later time, animals almost always begin performing again. For example, an animal running down a long alley for food does so vigorously for several trials, until food is withheld. After a few no-food experiences, the animal refuses to run. In a few hours or days, the running behavior can be expected to "recover"; i.e., reappear. This phenomenon is highly reliable, can be seen in virtually any training context, and is plainly time-dependent (reviewed in: Devenport, L., 1998; Robbins, 1990; Rescorla, 2004; Rescorla & Cunningham, 1978).

Spontaneous recovery can be seen as an instance of patch variability, but in this case variability in a single patch rather than the multiple patches we have used for our choice studies of TWR. The patch valuation function of TWR can, and has been (Devenport, L., 1998), applied to such instances in which the "choice" is between whether to perform the response or not. Early on, the weighted average TWR gives for a patch that was good and then bad will continue to be bad for some period of time, the length of which depends on the temporal pace of training and extinction, but the patch will progressively rise in subjective value and finally converge on the average of all training trials. At some point in this process, the increasing value will brighten the animal's prospects for obtaining food and it will once again perform the response. At least this is the explanation we prefer.

Current thinking in the animal learning literature favors a different interpretation, involving differential trace-decay mechanisms. The gist of the idea is that spontaneous recovery involves two associative memories (or "traces"), and that "inhibitory" traces (thought to be established by non-reward) decay more quickly than "excitatory" traces (thought to be established by reward). Thus, the inhibitory trace at first predominates, but soon becomes weaker than the more slowly decaying excitatory trace and thus trained behavior re-emerges (Bouton, 1993; 1994; Kraemer & Spear, 1993). Although somewhat adhoc, and not formalized into a quantitative model, this differential forgetting account seems to capture the essential features of spontaneous recovery and could conceivably be applied to more complex choice situations.

We therefore incorporated tests of memory into most of our experiments to determine if significant decay of any sort could be detected, and have found surprisingly little. In one paradigm (Devenport, J., & Devenport, L, 1993; Devenport, J., et al. 2005), three patches were used, but one was always empty. When two variable patches eventually converged on the same value (as given by the quantitative model of TWR), animals appeared to choose between them indiscriminately. This lack of discrimination could be attributed to the random

behavior of animals who have forgotten what happened at the patches or it could be attributed to indifference—they remember what happened, but each patch seems as good a prospect as the other. The evidence falls in favor of the latter interpretation. Despite their indifference about the two variable patches after a long delay, virtually no responses were directed to the never-rewarded, never-extinguished patch (Devenport, J., & Devenport, L, 1993; Devenport, J., et al. 2005). Even when allowed a second choice, animals selected the other variable patch over the neutral one, indicating that they remembered the behavior of each patch.

Another way of assessing memory factors is by creating patches of unequal value. We usually did this by placing more food items in one patch when in its good state, than in the poorer patch when in its respective good state. Both patches were variable and went through identical states of non-reward. When sufficient time passes since the patches were sampled, a cognitive interpretation of TWR predicts that choices should be made on the basis of average patch quality, not the last state the patches were in. This requires a memory not only of which patch yielded rewards at some point in the past, but which one yielded the most when in its good state. As expected from our optimality perspective, animals chose the patch with the best average—a complete reversal from the choices animals made when much less time had elapsed (Devenport, J., et al. 2005; Devenport, L., 1998). From a functional perspective, this implies that patch memories remained intact; otherwise choices would have been random.

The memory issue can also be addressed by requiring a highly improbable response. Using a T-maze, animals not only experienced variable quality outcomes in the patches found in each of the two arms, one arm had a better overall average than the other. To make the correct response more difficult, an additional pair of choice points at the end of each arm required further spatial discrimination. Only one member of the sub-pair was baited when a particular arm was in its good state. As expected, animals at first preferred the arm (and subchoice) that was most recently baited, but with the passage of time, the arm (and subchoice) with the higher average was preferred. Now, if those final choices were made because the details of the experience with the alternative arm were forgotten, then when allowed to make a second choice and run into the non-preferred arm, the subchoice should have been little better than random. This was not the case. Animals accurately remembered the details of where food was once available within the non-preferred arm (Devenport, L., et al, 1997). In a follow-up experiment, animals were trained in the same way and tested at short or long delays but at test no arm choice was allowed. The door was closed to what would have been the preferred arm, forcing the animals to go into the less-preferred arm and choose between the two sub-options. Again, the animals performed almost flawlessly, indicating that their previous behavior was not guided by the selective absence of memory.

In one of our earliest experiments (Devenport, L., & Devenport, J., 1994), we varied the probability that a patch would yield. Animals were given repeated experience with three patches and then one of the patches was baited for a final time. This is a case in which, according to associative theory and research, the least frequently baited patch should be the most resistant to extinction, since it yielded on a leaner schedule. Thus, predictions from a cognitive account of TWR and a trace-decay account about how animals should select among the patches after a long delay were exactly opposite. The choices clearly fell in favor of the cognitive interpretation, indicating that an active choice was made among equally salient memories, and not just the appearance of a "choice" springing from a weak or absent competing memory.

One other observation bears on the memory issue. Memory decay is considered to proceed at an unspecified but fixed rate, perhaps because of steadily declining synaptic strength. TWR's cognitive account, on the other hand, predicts that patch valuations will not change at a fixed rate. Instead, they should be directly tied to the pace of change experienced at the patches. We demonstrated this experimentally, showing that preference shifts across train-test intervals were hastened or delayed proportionally by the previously experienced rate of patch change (Devenport, J., & Devenport, L., 1993; Devenport, L., & Devenport, J., 1994), consistent with the use of intact memories and inconsistent with the trace-decay hypothesis. At present, we consider that TWR relies on declarative memory-type processes, not on forgetting.

#### **Computational Considerations**

Early optimal foraging theory was developed with overly optimistic assumptions about the cognitive architecture of animals. It concerned itself with the economics of time and energy budgets, hoping that a forager's cognitive adaptations would match the sophisticated economic formulations, which often required an exhaustive memory of experience. Some of the early results springing from this approach were disappointing (reviewed in Krebs, Houston, & Charnov, 1981). Animals did not always choose "optimally". Given the benefit of hindsight and historical developments, we can see that, more often then not, the formulations themselves were non-optimal. Still, the reaction by many in the animal behavior community was to swing in the opposite direction and assume that animals retain very little about their everyday experience and instead might rely on a set of (presumably innate) "rules-of-thumb" that would suffice to keep an animal competitive while making smaller computational demands (Cuthill, Kacelnik, Krebs, Haccou, & Iwasa, 1990; Dow & Lee, 1987; Kacelnik et al., 1987; Reboreda & Kacelnik, 1992). Several such "rules" have been suggested, and could probably serve the animal well enough, if not optimally, in some instances.

A more moderate position about animal cognition would admit that animals like humans, collect and process information selectively and do not remember irrelevant details of everyday life, nor do they retain an exhaustive record of relevant details. But, such limitations do not preclude flexible cognitive adaptations to improve decision-making. To illustrate, consider an animal that has visited a small set of patches that have varying histories. After attending to other concerns, it is now ready to resume foraging and needs to decide among these patches. There are only three simple rules of thumb available: 1) visit the patch that was good the last time, 2) visit the patch that has the best average, or 3) forage randomly. Rules 1 and 2 will often hold an advantage over random foraging, but not if the two patches are of equal value and perhaps not the same as other patches in the area. Under these conditions random foraging allows animals to sample the environment and discover unexpected opportunities. When Rules 1 and 2 do perform better, they will still be suboptimal because both fail to conform to the forces of patch variability that will tend to intervene before the animal returns. TWR handles this variability and provides a systematic transition from the predominant use of Rule 1 to the use of Rule 2, while transitioning through Rule 3.

How much more complex are the computations required by TWR? First, rules-of-thumb are not as simple as they seem. Remembering which patch was the last best one requires the time-tagging of experiences so that recency can be discriminated. The alternative is to erase records of earlier experiences so that only a memory of the last best patch remains. Of course, this option strips the forager of important knowledge, since the location and behavior of patches in its home range are surely among the most important things it can know. Rule 2, averaging, requires additional computation, but there is no doubt that animals carry out such operations. Faced with temporal variability (as in variable schedules of operant reinforcement) or variable quality of reward (as when magnitude or duration of reward varies), animals have no difficulty comparing and selecting the richest option, provided that the mean differences are of discriminable magnitude (Herrnstein,1970; Gallistel, 1990; Reboreda & Kacelnik, 1992). Although animals can compute them, exclusive reliance on averages would leave an animal unable to adapt to change, e.g., when typically poor patches produce temporary bonanzas. The output from TWR passes through a relatively wide region of indifference, where sampling and exploration of patches would be encouraged, similar to (but not in fact) like random foraging.

The cognitive demands of TWR require little more than averaging, and rely on what might even be considered a computational weakness. TWR considers all patch valuations to be averages. But since they are weighted by recency, the averaging function is not apparent until time passes and the relative recency of each patch experience becomes almost equal. The progression from highly weighted to nearly unweighted averages is the work of the computational "weakness". Animals, including humans, compare time, distance, amount, and so forth by division, not subtraction (reviewed in: Gallistel, 1990; Gibbon, 1977; Reboreda & Kacelnik, 1992), as described by Weber's Law. If the comparison were by subtraction (as a human trained in mathematics might prefer), then something that happened 2 seconds ago would always be 5 seconds more recent than something that happened 7 seconds ago. But temporal perception computes the difference as a ratio. At first the last-visited patch is 7/2 =3.5 times more recent than the other. As time passes, this difference vanishes and automatically converges on 1.0. For example, after 5 min, the last-visited patch is only 307/302 = 1.02 times more recent. Thus, no special weighting mechanism is required by TWR, just the timing mechanism that is already in place, Nor are excessive cognitive loads required since animals appear to use only a sample when computing an average based on large numbers of temporal events stored in memory (Gibbon, Church, Fairhurst, & Kacelnik, 1988).

Considering the gains to be made by conforming to a general rule of nature, i.e., regression-to-the-mean, and in light of the poorer rules of thumb available, which nevertheless require considerable computation, the cognitive resources needed for implementing TWR would, at least in theory, appear to be more than repaid in foraging success.

## **APPLICATIONS AND IMPLICATIONS**

Many types of decisions are made under uncertainty and therefore dynamic averaging might be useful for several types of decisions. For example, reproductive decisions such as where to lay eggs can affect offspring size and the sex ratio of eggs laid (reviewed in Ode & Heinz, 2002). Females have to decide whether to accept or reject an ovposition site and past information may be useful.

Parasitic wasps (*Diglyphus isaea*) exposed to increasingly larger host sizes laid significantly more female eggs than did those who experienced invariant host sizes, whereas those exposed to decreasing host sizes decreased the proportion of female eggs laid. TWR produced estimates of the relative weighting of experiences, in this case favoring recent events which would be expected given the rapid rate of change (Ode & Heinz, 2002). In addition to broadening the settings in which dynamic decisions are useful, this is the first demonstration of time-dependent decisions in an invertebrate.

Time-dependent decisions could also be useful for less important choices. There is evidence that humans make dynamic choices for non-biological alternatives. Participants given a simple problem solving task ("Build A Stick"), where correct solutions varied from trial to trial, chose a strategy based on how well the solution matched the current problem and the history of success of that strategy. The more successful a strategy had been in the past, the more likely it was to be chosen, regardless of its current utility (Lovett & Anderson, 1996). Despite substantial differences in the type of decision, a dynamic average appeared to be used in assigning weights to past and current history (Lovett, 1998). In a study directly testing TWR, human participants played a game that involved finding hidden tokens in one of three boxes followed by an equivalent number of trials locating tokens in a second box. Test trials were given at different delays following the second set of trials, in independent groups. As predicted, decisions were time-dependent, with 65% choosing the last box where tokens were found when tested immediately but only 20% choosing the last box at the delayed test. When independent groups were given a 4:1 difference in token densities, 80% chose the last baited, lower density box at the immediate test whereas only 10% chose it at the delayed test (Devenport, J., unpublished observations). These results show that dynamic decisions are made in a variety of situations and temporal weighting is useful whenever there is variability in the state of one or more alternatives.

## CONCLUSION

In conclusion, many decisions made by animals in the real world involve some degree of uncertainty, resulting from the passage of time. For many species that uncertainty appears to be minimized by a cognitive adaptation that takes time into account and weights the reliability of information by its relative recency. Dynamic decisions as specified by TWR have been demonstrated in a wide range of species, to aid decisions on a number of dimensions. The consistency in behavior seen in wild, domestic, and laboratory-bred species encourages us to think that TWR could serve as a flexible adaptation to assist decision-making under uncertainty. The variety of experimental settings, from field to laboratory, across long and short time spans, gives us confidence that dynamic decisions are easily implemented in real-world conditions and underscores the validity of the TWR model.

## REFERENCES

- Bailey, D.W., Gross, J.E., Laca, E.A., Rittenhouse, L.R., Coughenour, M.B., Swift, D.M., & Sims, P.L. (1996). Mechanisms that result in large herbivore grazing distribution patterns. *Journal of Range Management*, 49, 386-400.
- Bailey, D.W., & Sims, P.L. (1998). Association of food quality and locations by cattle. Journal of Range Management, 51, 2-8.
- Bateson, M. (2003). Interval timing and optimal foraging. In Meck, W. H. (Ed.), *Functional and neural mechanisms of interval timing* (pp. 113-141). Boca Raton, FL: CRC Press.
- Bird, L.R., Roberts, W.A., Abroms, B., Kit, K.A., & Crupi, C. (2003) Spatial memory for food hidden by rats (*Rattus norvegicus*) on the radial maze: studies of memory for where, what, and when. *Journal of Comparative Psychology*, 117, 176-187.
- Bond, A.B., Kamil, A.C., & Balda, R.P. (2003). Social complexity and transitive inference in corvids. *Animal Behaviour*, 65, 479-487.
- Bouton, M.E. (1993). Context, time and memory retrieval in the interference paradigms of Pavolvian learning. *Psychological Bulletin*, 114, 80-99.
- Bouton, M.E. (1994). Conditioning, remembering, and forgetting. *Journal of Experimental Psychology: Animal Behavior Processes, 20,* 219-231.
- Charnov, E.L. (1976). Optimal foraging: the marginal value theorem. *Theoretical Population Biology*, 9, 129-136.
- Church, R.M., & Gibbon, J. (1982). Temporal generalization. *Journal of Experimental Psychology: Animal Behavior Processes, 8,* 165-186.
- Clayton, N.S., Yu, K.S., & Dickinson, A. (2003). Interacting cache memories: evidence for flexible memory use by western scrub jays (*Aphelocoma californicia*). Journal of Experimental Psychology: Animal Behavior Processes, 29, 14-22.
- Cook, R.G., Brown, M.F., & Riley, D.A. (1985). Flexible memory processing by rats: use of prospective and retrospective information in the radial maze. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 12-20.
- Cowie, J. (1977). Optimal foraging in great tits (Parus major). Nature, 268, 137-139
- Cuthill, I., Haccou, P., & Kacelnik, A. (1994). Starlings (*Sturnus vulgaris*) exploiting patches: response to long-term changes in travel time. *Behavioral Ecology*, *5*, 81-90.
- Cuthill, I.C., Kacelnik, A., Krebs, J.R., Haccou, P., & Iwasa, Y. (1990). Starlings exploiting patches: the effect of recent experience on foraging decisions. *Animal Behaviour*, 40, 625-640.
- Dall, S.R.X., Giraldeau, L-A., Olsson, O., McNamara, J.M., & Stephens, D.W. (2005). Information and its use by animals in evolutionary ecology. *Trends in Ecology and Evolution*, 20, 187-193.
- de Kort, S.R., Dickinson, A., & Clayton, N.S. (2005). Retrospective cognition by western scrub-jays. *Learning & Motivation*, *36*, 159-176.
- Devenport, J.A., & Devenport, L.D. (1993). Time-dependent decisions in dogs. Journal of Comparative Psychology, 107, 169-173.
- Devenport, J.A., & Devenport, L.D. (1994) Spatial navigation in natural habitats by grounddwelling sciurids. Animal Behaviour, 47, 747-749.
- Devenport, J. A., & Devenport, L. D. (1998). Foraging in squirrels: Cognitive adaptations. In

- M. Haraway & G. Greenberg (Eds.), *Comparative psychology: A handbook* (pp. 492-498). New York: Garland Press.
- Devenport, J.A., Patterson, M.R., & Devenport, L.D. (2005). Dynamic averaging and foraging decisions in horses (*Equus callabus*). Journal of Comparative Psychology, 119, 352-358.
- Devenport, L.D. (1983). Spontaneous behavior: inferences from neuroscience. In R.L. Mellgren (Ed.), Animal cognition and behavior (pp. 83-126). Amsterdam: Elsevier.
- Devenport, L. (1998). Spontaneous recovery without interference: Why remembering is adaptive. *Journal of Experimental Psychology: Animal Behavior Processes, 26*, 172-181.
- Devenport, L.D., & Devenport, J.A. (1994). Time-dependent averaging of foraging information in least chipmunks and golden-mantled ground squirrels. *Animal Behaviour*, 47, 787-802.
- Devenport, L., Hill, T., Wilson, M., & Ogden, E. (1997). Tracking and averaging in variable environments: A transition rule. *Journal of Experimental Psychology: Animal Behavior Processes, 23*, 450-460.
- Devenport, L.D., Humphries, T W., & Devenport, J.A. (1998). Future value and patch choice in least chipmunks. *Animal Behaviour*, 55, 1571-1581.
- Dow, S.M., & Lea, S.E.G. (1987). Foraging in a changing environment: simulations in the operant laboratory. In M.C. Commons, A. Kacelnik, & S.J. Shettleworth (Eds.), *Quantitative analysis of behavior Vol. 6: foraging* (pp. 89-113). Hillsdale, NJ: Lawrence Earlbaum.
- Emory, N.J., & Clayton, N.S. (2001). Effect of experience and social context on prospective caching strategies by scrub jays. *Nature*, 414, 443-445.
- Fox, M.W. (1965). Canine behavior. Springfield, IL: Charles C. Thomas.
- Gallistel, C.R. (1990). The organization of learning. Cambridge, MA: MIT Press.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, 84, 279-325.
- Gibbon, J. Church, R.M., Fairhurst, S., & Kacelnik, A. (1988). Scalar expectancy theory and choice between delayed rewards. *Psychological Review*, *95*, 102-114.
- Grosenick, L., Clement, T.S., & Fernald, R.D. (2007). Fish can infer social rank by observation alone. *Nature*, 445, 429-432.
- Herrnstein, R.J. (1970). On the law of effect. Journal of the Experimental Analysis of Behavior, 13, 243-266.
- Inglis, I.R. (2000). The central role of uncertainty reduction in determining behaviour. *Behaviour*, 137, 1567-1599.
- Inglis, I.R., Lanagton, S., Forkman, B., & Lazarus, J. (2001). An information primacy model of exploration and foraging behaviour. *Animal Behaviour*, 62, 543-557.
- Jacobs, L.F., & Liman, E.R. (1991). Grey squirrels remember the locations of buried nuts. *Animal Behaviour, 41*, 103-110.
- Kacelnik, A., Krebs, J.R., & Ens, B. (1987). Foraging in a changing environment: an experiement with starlings (*Sturnus vulgaris*). In M.L. Commons, A. Kacelnik, & S.J. Shettleworth (Eds.), *Quantitative analysis of behavior Vol. 6: foraging* (pp. 63-87). Hillsdale, NJ: Lawrence Earlbaum.
- Kamil, A.C., & Balda, R.P. (1985). Cache recovery and spatial memory in Clark's nutcrackers. Journal of Experimental Psychology: Animal Behavior Processes, 11, 95-111.

- Krebs, J.R., Houston, A.I., & Charnov, E.L. (1981). Some recent developments in optimal foraging. In A.C. Kamil & T.D. Sargent (Eds.), *Foraging behavior: ecological, ethological, and psychological approaches* (pp. 3-18). NY:Garland Press.
- Kraemer, P., & Spear, N.E. (1993). Retrieval processes and conditioning. In T.R. Zentall (Ed.), Animal cognition (pp. 87-107). Hillsdale, NJ: Erlbaum.
- Lejeune, H., & Wearden, J.H. (1991). The comparative psychology of fixed-interval responding: some quantitative analyses. *Learning and Motivation*, 22, 84-111.
- Lima, S. L., & Dill L.M. (1990). Behavioral decisions made under the risk of predation: a review and prospectus. *Canadian Journal of Zoology*, 68, 619-640.
- Lovett, M.C. (1998). Choice. In J.R. Anderson & C. Lebiere (Eds.), *The atomic components* of thought (pp. 255-296). Mahwah, NJ: L. Erlbaum.
- Lovett, M.C., & Anderson, J.R. (1996). History of success and current context in problem solving: combined influences on operator selection. *Cognitive Psychology*, *31*, 168-217.
- MacDonald, I.M.V. (1997). Field experiments of duration and precision of grey and red squirrel spatial memory. *Animal Behaviour*, 54, 879-891.
- Matell, M.S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *BioEssays*, 22, 94-103.
- Mazur, J.E. (1995). Development of preference and spontaneous recovery in choice behavior with concurrent variable-interval schedules. *Animal Learning & Behavior, 23*, 93-103.
- Mazur, J.E. (1996). Past experience, recency, and spontaneous recovery in choice behavior. *Animal Learning & Behavior, 24*, 1-10.
- Naqshbandi, M., & Roberts, W.A. (2006). Anticipation of future events in squirrel monkeys (Saimiri sciureus) and rats (Rattus norvegicus): test of the Bischof-Kohler hypothesis. Journal of Comparative Psychology, 120, 345-357.
- Ode, P.J., & Heinz, K.M. (2002). Host-size-dependent sex ratio theory and improving massreared parasitoid sex ratios. *Biological Control*, 24, 31-41.
- Olton, D.S. (1979). Mazes, maps, and memory. American Psychologist, 34, 583-596.
- Preston, S.D., & Jacobs, L.F. (2001). Conspecific pilferage but not presence affects Merriam's kangaroo rat caching strategy. *Behavioral Ecology*, 12, 517-523.
- Raby, C.R., Alexis, D.M., Dickinson, A., & Clayton, N.S. (2007). Planning for the future by western scrub-jays. *Nature*, 445, 919-921.
- Reboreda, J.C., & Kacelnik, A. (1992). Risk sensitivity in starlings: variability in food amount and food delay. *Behavioral Ecology*, *2*, 301-308.
- Rescorla, R.A. (2004). Spontaneous recovery. Learning and Memory, 11, 501-509.
- Rescorla, R.A., & Cunningham, C.L. (1978). Recovery of the US representation over time during extinction. *Learning & Motivation*, 9, 373-391.
- Robbins, S.J. (1990). Mechanisms underlying spontaneous recovery in autoshaping. *Journal* of Experimental Psychology: Animal Behavior Processes, 16, 235-249.
- Roberts, W.A. (2002). Are animals stuck in time? Psychological Bulletin, 128, 473-489.
- Roberts, W.A., & Phelps, M.T. (1994). Transitive inference in rats: a test of the spatial coding hypothesis. *Psychological Science*, *5*, 368-374.
- Shettleworth, S. (1998). *Cognition, evolution, and behavior*. Oxford: Oxford University Press.
- Stephens, D.W. (1989). Variance and the value of information. *American Naturalist, 134*, 128-140.
- Stephens, D.W., & Anderson, D. (2001). The adaptive value of preference for immediacy:

when shortsighted rules have farsighted consequences. *Behavioral Ecology*, *12*, 330-339. Suddendorf, T., & Busby, J. (2003). Mental time travel? *Trends in Cognitive Science*, *7*, 391-396.

Tulving, E. (1983). Elements of episodic memory. Oxford: Oxford University Press.

VanderWall, S.B. (1990). Food hoarding in animals. Chicago: University of Chicago Press.

Winterrowd, M.F., & Devenport, L.D. (2004). Balancing variable patch quality with predation risk. *Behavioural Processes*, 67, 39-46.

Zentall, T.R. (2005). Animals may not be stuck in time. Learning & Motivation, 36, 208-225.

Zentall, T.R., Steirn, J.N., & Jackson-Smith, P. (1990). Memory strategies in pigeons' performance of a radial-arm-maze analog task. *Journal of Experimental Psychology: Animal Behavior Processes*, *16*, 358-371.

Chapter 46

# **DEVELOPMENTAL TRENDS IN DECISION MAKING: THE CASE OF THE MONTY HALL DILEMMA\***

## Wim De Neys<sup> $\dagger$ </sup>

University of Leuven, Tiensestraat 102 3000 Leuven, Belgium

## ABSTRACT

The Monty Hall Dilemma (MHD) is a notorious brain-teaser where people have to decide whether switching to another option in a game is advantageous. Most adults erroneously believe that chances of winning remain equal or that they should stick to their original choice. The present study tested the impact of cognitive development on MHD reasoning to examine possible differences in the nature of the erroneous intuitions. Twelve to seventeen year old high school students were presented the MHD and selected one of three responses (switch, stick, or chances equal). Results showed that whereas maturation decreased adherence to the erroneous "stick with your first pick" belief, the "chances are equal" belief became more dominant with increasing age. Consistent with predictions, children who selected the latter response also scored better on a syllogistic reasoning task. Results further showed that twelve year old eighth graders selected the correct switching response more frequently than senior high school students. Implications for popular reasoning and decision making theories are discussed.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: Lab Experimentele Psychologie, K.U. Leuven Tiensestraat 102 3000 Leuven Belgium, Email: Wim.Deneys@psy.kuleuven.be, Homepage : http://ppw.kuleuven.be/reason/wim/, Fax: +32 16 326099 Tel: +32 16 326143

## DEVELOPMENTAL TRENDS IN MONTY HALL DILEMMA REASONING

Imagine you're the final guest in a TV quiz. Monty, the show host, is asking you to choose one of three doors. One of the doors conceals a BMW sports car but the two other doors only contain a bunch of toilet paper. If you choose the right door you will be the proud owner of the fancy BMW. However, after you finally select one of the doors, host Monty does not open it immediately. First, he opens one of the doors you did not choose to reveal it contained toilet paper. Monty now actually offers you the possibility to chance your mind and pick the other unopened door. What should you do to have most chance of winning the car of your dreams? Stay with your first choice, switch to the other unopened door, or doesn't it matter whether you switch or not?

The above switching problem is known as the "Monty Hall Dilemma" (MHD) after the host of the American TV show "Let's make a deal" where it was introduced. Contrary to most people's intuition the correct answer is that switching to the other door will actually increase the chance of winning. The counter-intuitive solution hinges on the crucial fact that Monty will never open the door concealing the prize, and obviously, he will not open the door the guest initially picked either. The probability that you initially select the correct door is one out of three. In this case it would be better not to switch. However, in the other two thirds of the cases the non-chosen closed door will hide the prize and switching is advantageous. Hence, switching yields a 2/3 chance of winning.

If you failed to solve the problem correctly you may find some comfort in the fact that you're in good company. Empirical studies show that typically less than 10% of educated adults give the correct switching response (e.g., Burns & Wieth, 2004; Friedman, 1998; Granberg & Brown, 1995; Krauss & Wang, 2003; Tubau & Alonso, 2003) and even ace mathematicians do not seem to be immune to MHD errors (e.g., Burns & Wieth; vos Savant, 1997). Most people have the strong intuition that whether they switch or not the probability of winning remains 50% either way. Research indicates that this powerful intuition is based on the so called number-of-cases heuristic ("if the number of alternatives is N, then the probability of each one is 1/N", see Shimojo & Ichikawa, 1989, and Falk, 1992). Since only two doors remain people will automatically assign a 50% chance to each door and fail to take the "knowledgeable host" information into account. The "equal chance" heuristic is so self-evident that it will literally dominate our thinking.

Over the last decades numerous studies have demonstrated that similar intuitive responses are biasing people's performance in a wide range of reasoning tasks (e.g., Evans & Over, 1996; Kahneman, Slovic, & Tversky, 1982). Influential dual process theories of reasoning and decision making have explained this "rational thinking failure" by positing two different human reasoning systems (e.g., Epstein, 1994; Evans, 2003; Evans & Over, 1996; Sloman, 1996; Stanovich & West, 2000). Dual process theories come in many flavours but generally they assume that a first system (often called the heuristic system) will tend to solve a problem by relying on intuitions and prior beliefs whereas a second system (often called the analytic system) allows reasoning according to normative standards. The heuristic default system is assumed to operate fast and automatically whereas the operations of the analytic system would be slow and heavily demanding of people's computational resources. Although the fast and undemanding heuristics can provide us with useful responses in many daily situations they can bias reasoning in tasks that require more elaborate, analytic processing

(e.g., Sloman, 1996; Stanovich & West, 2000; Tversky & Kahneman, 1983). That is, both systems will sometimes conflict and cue different responses. In these cases the analytic system will need to override the automatically generated intuitive response. Since the inhibition of the heuristic intuitions and the computations of the analytic system draw heavily on people's limited cognitive resources, most people will be tempted to stick to mere intuitive reasoning. Therefore, correct analytic reasoning would be characteristic of those highest in cognitive capacity (Stanovich & West, 2000).

De Neys and Verschueren (2006) recently examined the relation between people's cognitive capacities and Monty Hall reasoning. Over 200 participants were presented the MHD and a test to measure their working memory capacity. Consistent with Stanovich and West's (2000) dual process predictions participants who did manage to solve the MHD correctly were specifically those highest in working memory span. Experimentally burdening the cognitive resources with a secondary task also decreased the rate of correct switching responses. These findings supported the basic claim that the analytic override of heuristic thinking draws on people's cognitive working memory resources.

Although the dual process framework has been quite influential in the reasoning and decision making community it has also been criticized severely (e.g., Gigerenzer & Regier, 1996; see commentaries on Evans & Over, 1997, or Stanovich & West, 2000, for an overview). One of the critiques focuses on the the framework's postulation of a single heuristic system that unitarily handles all intuitive processing. It has been argued that different kinds of heuristics, with a different processing nature, need to be differentiated (e.g., De Neys, 2006a; Gigerenzer & Regier, 1996; Moshman, 2000; Newton & Roberts, 2003). De Neys and Verschueren's (2006) study pointed to an MHD trend that is specifically interesting in this respect. Note that one can distinguish different erroneous MHD responses. Whereas the vast majority of reasoners is biased by the number-of-cases heuristic and beliefs that switching or sticking does not matter, a small group of reasoners is convinced that they should stick to the initially selected door to win the prize. These people are biased by the general belief that when making a decision one should always stick to one's first choice. Such a bias has long been noted in responses to multiple-choice exams (e.g., Geiger, 1997). Gilovich, Medvec, and Chen (1995) clarified that this "stick with your pick" intuition would be based on an anticipation of regret.

De Neys and Verschueren (2006) found that reasoners in the group of erroneous responders who selected the "equal chances" response tended to have a slightly higher working memory span than those who believed that "sticking" was advantageous. Moreover, the secondary task load specifically boosted the rate of "sticking" responses whereas the selection of "chances equal" responses was hardly affected. Although the trends were not significant, De Neys and Verschueren noted that the two types of heuristic beliefs might indeed have a different processing nature. The number-of-cases heuristic would be based on a cognitive probability estimation (albeit a simple one) whereas the "stick with your pick" heuristic would be computationally more complex than the "stick with your pick" heuristic. Therefore, the more elementary, least demanding "stick with your pick" response would be the preferred answer under conditions of cognitive load.

De Neys and Verschueren (2006) suggested that in their sample of university students the trends presumably failed to reach significance because the basic number-of-cases computation would be completely automatic for educated adults. Consequently, the

distinction with the alleged less demanding "sticking" responses would be blurred. The present study adopts a developmental approach to clarify the issue. The MHD was presented to twelve to seventeen year old adolescents. A first prediction concerned the selection rates of the different MHD responses in the different age groups. Younger reasoners have smaller cognitive resource pools and still lack an important part of the mathematical training that familiarized university students with fractions and probabilities. Therefore, one can expect that computation of the number-of-cases heuristic will be less automated and more demanding in the younger age groups. This should result in a less frequent selection of "equal chances" responses by younger reasoners. Under the assumption that the "stick with your pick" heuristic is more basic one predicts that younger reasoners will show a stronger preference for the "stick" responses. The alleged differential demands of the "stick" and "equal" heuristics should thus result in a specific developmental trend: The rate of "chances equal" responses should increase with age whereas the "sticking" response should decrease. Given the high computational demands of the correct switching response for adult university students, correct MHD reasoning in the younger age groups was not expected.

A second prediction concerns the overall relation between young adolescents' responses and their cognitive capacity. After the students had solved the MHD they were also presented a specific syllogistic reasoning task where they had to inhibit automated responses. Previous studies established that this task is a good marker of children's general cognitive ability (e.g., Kokis, Macpherson, Toplak, West, & Stanovich, 2004). De Neys and Verschueren (2006) argued that the completely automated computation of the "equal chances" response in their sample of university students blurred the distinction with the alleged undemanding "sticking" responses. The less automated and more demanding nature of the number-of-cases heuristic for the younger reasoners (vs. adults) should show a clearer distinction. Therefore, it is predicted that youngsters who manage to give the "chances are equal" response will score better on the syllogisms than those who believe that mere sticking is the best strategy.

The developmental MHD trends will help clarifying possible differences in the nature of different heuristics. In addition, examining children's MHD performance will also have interesting implications for a fundamental controversy concerning the development of reasoning and decision making itself. The dual process framework and many developmental theories (e.g., Case, 1985; Inhelder & Piaget, 1985) share the assumption that children's reasoning becomes less heuristic and more logical across the lifespan. Hence, traditionally it has been assumed that analytic thinking simply replaces heuristic thinking with cognitive maturation. Recent developmental reasoning studies have argued against this so-called "illusion of replacement" (e.g., Brainerd & Reyna, 2001; Reyna & Ellis, 1994). Klaczynski (2001), for example, showed that whereas in some tasks (e.g., reasoning about sunk costs) the heuristic appeal indeed decreased with age, other tasks showed the opposite pattern and indicated that heuristic reasoning remained constant or increased with age (e.g., denominator neglect in statistical reasoning). Such findings are already hard to reconcile with the traditional view. The present MHD study allows a direct validation of the replacement claim within one and the same task. Based on the traditional view one would simply expect that both types of erroneous MHD responses will decrease with age and will be replaced by a higer number of correct responses. Evidence for the predicted differential developmental "equal chances" and "sticking" trends will cut the ground under the replacement view.

## Method

## Participants

A total of 132 high school students in grades 8 to 12 participated in the study. Forty-two students attended eighth grade (mean age = 12.9, SD = .91), 25 attended ninth grade (mean age = 14.64, SD = .64), 20 students were in grade 10 (mean age = 15.75, SD = 1.02), 16 were in grade 11 (mean age = 16.56, SD = .63), and 29 in grade 12 (mean age = 17.66, SD = .69). Students in grades 9/10 and grades 11/12 were collapsed in two age groups for the analyses. All participants were recruited from the same suburban school with socially mixed catchments areas. All spoke Dutch as their first language and had no known behavioral problems or learning difficulties.

## Material

*Monty Hall Dilemma*. Students were presented a version of the MHD based on Krauss and Wang (2003, see De Neys & Verschueren, 2006). The formulation tried to avoid possible ambiguities (e.g., the random placement of the prize and duds behind the doors and the knowledge of the host were explicitly mentioned). As in Tubau and Alonso (2003), participants could choose between three answer alternatives (a. Stick – b. Switch – c. Chances are even). The complete problem format is presented below:

Suppose you're on a game show and you're given the choice of three doors. Behind one door is the main prize (a car) and behind the other two doors there are dud prizes (a bunch of toilet paper). The car and the dud prizes are placed randomly behind the doors before the show. The rules of the game are as follows: After you have chosen a door, the door remains closed for the time being. The game show host, Monty Hall, who knows what is behind the doors, then opens one of the two remaining doors which always reveals a dud. After he has opened one of the doors with a dud, Monty Hall asks the participants whether they want to stay with their first choice or to switch to the last remaining door. Suppose that you chose door 1 and the host opens door 3, which has a dud.

The host now asks you whether you want to switch to door 2. What should you do to have most chance of winning the main prize?

- a. Stick with your first choice, door 1.
- b. Switch to door 2.
- c. It does not matter. Chances are even.

The MHD was presented on computer. Participants were instructed to carefully read the basic problem information first. When they were finished reading they pressed the ENTER-key and then the question and answer-alternatives (underscored text) appeared on the screen (other text remained on the screen). Participants typed their response (a, b, or c) on the keyboard. Instructions stated there were no time limits.

*Syllogistic Reasoning Task.* The syllogistic reasoning task was based on Sá, West, and Stanovich (1999). Participants evaluated four syllogisms taken from the work of Markovits and Nantel (1989) whereby the logical status of the conclusion conflicted with its believability (e.g., a valid but unbelievable conclusion like 'All mammals can walk. Whales are mammals. Therefore, whales can walk'). Thus, in order to give a correct response the reasoner has to override the tendency to judge the conclusion based on its believabilty. Individual differences studies (e.g., De Neys, 2006b; Stanovich & West, 2000; Newstead, Handley, Harley, Wright, Farrelly, 2004) and developmental research (Kokis et al., 2004) suggest that performance on this task is a good marker of cognitive capacity as measured with classic intelligence or working memory tests<sup>1</sup>.

#### Procedure

Participants were tested in groups during a course break. The MHD was presented before the syllogistic reasoning task.

## RESULTS

Figure 1 shows the selection rates of the three possible MHD responses in the different age groups. For completeness, rates are shown for every grade separately. Performance of university students in De Neys and Verschueren's (2006) study is also plotted for comparison.

As Figure 1 shows, twelve year old eighth graders have no clear preference for the "chances are equal" or "stick with your pick" response. Both are selected by about 45% of the eighth graders. As expected, however, the number-of-cases heuristic is becoming more and more dominant in the older age groups. The preference for the "stick" response, on the other hand, decreases with age. To test these trends statistically students in grades 9/10 and 11/12 were collapsed in two age groups. Selection rates in these two groups were compared with selection rates in grade 8. This resulted in an approximately equal *n* in the three age groups. Cochran's Q test showed that the increase in the proportion of "chances are equal" responses from grade 8, over grade 9/10, to grade 11/12 was significant, Q(2)= 21.44, p < .0001. The decrease in "sticking" responses over the three age groups also reached significance, Q(2)= 16.91, p <.001. These findings are consistent with the claim that the "stick with your pick" belief is based on a more basic and computationally less complex heuristic than the "chances are equal" belief. Cognitive maturation and mathematical training during the high school years seem to be especially boosting the number-of-cases heuristic.

<sup>&</sup>lt;sup>1</sup> Time-constraints and school policy prevented the administration of more classic ability tests.

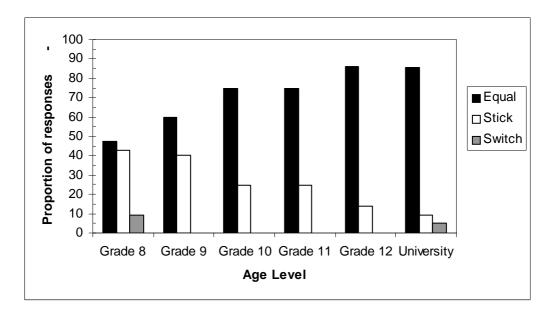


Figure 1. Proportion of the three possible Monty Hall Dilemma responses in different age groups.

None of the students in grades 9 to 12 managed to give the correct switching response on the standard MHD. However, the youngest participants in our study, the twelve to thirteen year old eighth graders, showed about 10% switching responses. The decrease in switching responses over the three high school age levels also reached significance, Cochran Q(2) = 10, p < .01. The finding that twelve year olds outperform senior high school (and even university) students might seem surprising at first sight. However, we believe that the higher rate of correct responding for the eighth graders can be described as a "Less Is More Effect". Adults' reasoning is impeded precisely because of the automated nature of the "50%" response. For adults the "equal chance" heuristic is so self-evident that it literally dominates their reasoning. For a twelve year old the conclusion that since two doors remain both will have a 50% chance of hiding the prize will be less evident and will require active, resource demanding computation. Bluntly put, for younger reasoners the switching response will be less counter-intuitive than for adults. In this sense children thus actually benefit from a lack of knowledge. This issue is further addressed in the General Discussion.

Students' performance on the syllogistic reasoning task was taken as a measure of their general cognitive ability. It was predicted that for less mature reasoners computation of the "chances are equal" response would be less automated than for adult, university students. Therefore, in the present younger age group the two response types should show a better differentiation in terms of cognitive capacity. The "chances are equal" heuristic is assumed to be more demanding than the "stick with your pick" intuition. Consequently it was predicted that children with a larger resource pool would be more likely to arrive at the "equal" response. This hypothesis was confirmed. Overall, students who selected the "equal" response solved 43% of the syllogisms correctly (mean score = 1.70, SD = 1.27) whereas students who believed that sticking was beneficial only solved 28% correctly (mean score = 1.12, SD = 1.08), t(125) = 2.50, n1 = 86, n2 = 41, p < .015, d = .48.

The eighth graders were the only students who gave switching responses. For completeness we compared the syllogistic reasoning performance of eighth graders who gave

correct and incorrect responses. There were no significant differences between the two groups. Interestingly, contrary to the university students in De Neys and Verschueren's (2006) study, eighth graders who selected the switching response actually tended to score lower on the measure of executive capacities (19% correct conclusions) than the eighth graders who erroneously selected the "sticking" or "chances equal" responses (28% correct conclusions). This is consistent with the idea that selecting the switching response for eighth graders is based on an intuitive rather than an executive resource demanding reasoning process.

## **GENERAL DISCUSSION**

The present study explored developmental trends in children's reasoning about the Monty Hall Dilemma (MHD). By comparing the MHD performance in different age groups, possible differences in the nature of erroneous intuitions were examined. Previous MHD studies distinguished at least two types of erroneous MHD reasoning (e.g., De Neys & Verschueren, 2006; Gilovich et al., 1995; Tubau & Alonso, 2003). Participants could indicate they believed that switching and sticking had both a 50% chance of winning or they could indicate they believed that sticking to the first chosen door was advantageous. Based on trends in the work of De Neys and Verschueren it was hypothesized that the "stick with your pick" response would be more elementary and computationally less demanding than the "chances are equal" response. Consistent with this claim results showed that maturation over the high school years boosted the dominance of the "chances are equal" heuristic. Children who selected this response also scored higher on an indirect measure of cognitive capacity than those who preferred the "sticking" response. These findings support the idea that the two heuristic responses differ in computational complexity. For educated adults, however, the computation of the "50%" heuristic would become fully automated. Indeed, it is precisely the automated, self-evident, and intuitive nature of this heuristic that is supposed to be impeding university students' reasoning.

Results further showed that the youngest reasoners in the study outperformed senior high school or university students. Twelve and thirteen year old eighth graders showed about 10% correct responses. It is interesting to note that such apparently counter-intuitive developmental findings have also been observed with other classic reasoning tasks. Jacobs and Potenza (1991), for example, studied children's performance on the notorious base-rate neglect problems (e.g., the lawyer-engineer problem, Kahneman & Tversky, 1973). In these tasks, salient, stereotypical information is pitted against more reliable statistical base rate information. When a person is described as an engineer, adults will conclude it is an engineer although they were told that the person was drawn from a sample were there were twice as many lawyers than engineers. The vast majority of educated adults typically neglect the statistical base rate information. Jacobs and Potenza observed that the base rate neglect decreased with decreasing age. As Kokis et al. (2004) argued, the finding that younger children err less frequently on these problems is not surprising because stereotype knowledge is typically less developed for children. Since children lack knowledge of many social stereotypes, they may seem to be using base-rate information more, simply because the potentially biasing information is unavailable to them. As with the present findings, the point is not that twelve year olds are actively computing the correct base rate or switching response but rather that twelve year olds are less tempted by the heuristics that are impeding adults' reasoning.

The present MHD trends support the rejection of the "illusion of replacement" in the developmental literature (e.g., Brainerd & Reyna, 2001). Traditionally, reasoning and decision making researchers have characterized cognitive development as a process whereby children's reasoning becomes less heuristic and more in line with logical standards. Recent developmental studies started cutting the ground under this view (e.g., Klaczynski, 2001). The present study strengthens this critique by showing that different heuristics can follow a different developmental path. This further dismisses the "illusion of replacement" as a simplistic idealization.

The work of De Neys and Verschueren (2006) stipulated that limitations in working memory resources cannot be neglected as cause of erroneous MHD reasoning. For educated adults, sidestepping salient intuitions and selecting the correct switching response will be compromised by a lack of cognitive resources. These findings fitted well with dual process theories' characterization of the *analytic* reasoning system. The present study clarifies that at the same time one must bear in mind that more extreme resource shortages for less mature or gifted reasoners might also compromise the automated cueing of erroneous heuristics itself. Moreover, the developmental findings seriously question the framework's characterization of the *heuristic* system as a unitary system that handles all intuitive processing uniformly (e.g., De Neys, Schaeken, & d'Ydewalle, 2005). The present MHD results stress that different kinds of heuristics, with a different processing nature, need to be differentiated. One and the same heuristic might be differentially computed by different groups of participants: A heuristic that is automatically triggered for one group of participants might require active cognitive resource demanding computations for a younger and/or less gifted group of reasoners. In line with recent critiques, dual process theories need to take this diversity into account in order to fully characterize the processing specifications of the two reasoning systems.

#### ACKNOWLEDGMENTS

Wim De Neys is post doctoral fellow of the Fund for Scientific Research-Flanders (FWO-Vlaanderen). I would like to thank Ellen Gillard and Deborah Everaerts for their help testing the high school students. Parts of this study were presented at the 27<sup>th</sup> Annual Conference of the Cognitive Science Society, Stresa, Italy.

#### REFERENCES

- Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory : Dual processes in memory, reasoning, and cognitive neuroscience. In H. W. Reese & R. Kail (Eds.), *Advances in child development and behavior* (Vol. 28, pp. 41-100). San Diego: Academic Press.
- Burns, B. D., & Wieth, M. (2004). The collider principle in causal reasoning: Why the monty hall dilemma is so hard. *Journal of Experimental Psychology: General, 133,* 434-449.

Case, R. (1985). Intellectual development: Birth to adulthood. New York: Academic Press.

- De Neys, W. (2006a). Automatic-heuristic and executive-analytic processing in reasoning: Chronometric and dual task considerations. *Quarterly Journal of Experimental Psychology*, 59, 1070-1100.
- De Neys, W. (2006b). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, 17, 428-433.
- De Neys, W., & Verschueren, N. (2006). Working memory capacity and a notorious brain teaser: The case of the Monty Hall Dilemma. *Experimental Psychology*, *53*, 123-131
- De Neys, W., Schaeken, W., & d'Ydewalle, G. (2005). Working memory and counterexample retrieval for causal conditionals. *Thinking & Reasoning*, 11, 123-150.
- Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist*, 49, 709-724.
- Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. Trends in Cognitive Sciences, 7, 454-459.
- Evans, J. St. B. T., & Over, D. E. (1996). Rationality and reasoning. Hove, UK: Psychology Press.
- Evans, J. St. B. T., & Over, D. E. (1997). Rationality in reasoning: The problem of deductive competence. *Current Psychology of Cognition*, 16, 3-38.
- Falk, R. (1992). A closer look at the probabilities of the notorious three prisoners. *Cognition*, 77, 197-213.
- Friedman, D. (1998). Monty Hall's three doors: Construction and deconstruction of a choice anomaly. American Economic Review, 88, 933-946.
- Geiger, M. A. (1997). Educators' warnings about changing examination answers: Effects on students perceptions and performance. *College Student Journal*, *31*, 429-432.
- Gigerenzer, G., & Regier, T. (1996). How do we tell an association from a rule?: Comment on Sloman (1996). *Psychological Bulletin, 119*, 23-26.
- Gilovich, T., Medvec, V. H., & Chen, S. (1995). Commission, omission, and dissonance reduction: Coping with regret in the "Monty Hall" problem. *Personality and Social Psychology Bulletin*, 21, 185-190.
- Granberg, D., & Brown, T. A. (1995). The Monty Hall dilemma. Personality and Social Psychology Bulletin, 21, 182-190.
- Inhelder, B., & Piaget, J. (1958). The growth of logical thinking from childhood to adolescence. New York: Basic books.
- Jacobs, J. E., & Potenza, M. (1991). The use of judgment heuristics to make social and object decisions: A developmental perspective. *Child Development*, 62, 166-178.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgement under uncertainty: Heuristics and biases. Cambridge, MA: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.
- Klaczynski, P. A. (2001). Framing effects on adolescent task representation, analytic and heuristic processing, and decision making: Implications for the normative/descriptive gap. *Applied Developmental Psychology*, 22, 289-309.
- Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology*, 83, 26-52.

- Krauss, S., & Wang, X. T. (2003). The psychology of the Monty Hall Problem: Discovering psychological mechanisms for solving a tenacious brain teaser. *Journal of Experimental Psychology: General*, 132, 3-22.
- Markovits, H., & Nantel, G. (1989). The belief bias effect in the production and evaluation of logical conclusions. *Memory and Cognition*, 17, 11-17.
- Moshman, D. (2000). Diversity in reasoning and rationality: Metacognitive and developmental considerations. *Behavioral and Brain Sciences*, 23, 689-690.
- Newstead, S. E., Handley, S. J., Harley, C., Wright, H., Farrelly, D. (2004). Individual differences in deductive reasoning. *Quarterly Journal of Experimental Psychology*, 57A, 33-60.
- Newton, E. J., & Roberts, M. J. (2003). Individual differences transcend the rationality debate. *Behavioral and Brain Sciences*, 26, 530-531.
- Reyna, V. F., & Ellis, S. (1994). Fuzzy-trace theory and framing effects in children's risky decision making. *Psychological Science*, 5, 275-279.
- Sá, W., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, 91, 497-510.
- Shimojo, S., & Ichikawa, S. (1989). Intuitive reasoning about probability: Theoretical and experimental analysis of the "problem of three prisoners". *Cognition, 32*, 1-24.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3-22.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, 23, 645-726.
- Tubau, E., & Alonso, D. (2003). Overcoming illusory inferences in a probabilistic counterintuitive problem : The role of explicit representations. *Memory & Cognition*, 31, 596-607.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293-315.
- Vos Savant, M. (1997). The power of logical thinking. New York: St. Martin's Press.

Chapter 47

# CHARTING THE COURSE OF SELF EVALUATIONS AND SOCIAL COMPARISONS OVER TIME\*

## Janet A. Sniezek<sup>†1</sup>, Phanikiran Radhakrishnan<sup>2</sup> and Tahira M. Probst<sup>3</sup>

<sup>1</sup>(posthumous) University of Illinois at Urbana-Champaign, Champaign, Illinois, USA <sup>2</sup>University of Toronto at Scarborough, Scarborough, ON, Canada M1C1A4 <sup>3</sup>Washington State University, Vancouver, Washington, USA

## ABSTRACT

Self-evaluations of performance are important in theory and practice. In contexts with multiple persons performing the same task, the evaluation of one's own performance is expected to be a process involving judgments about the performance of others, and comparisons between one's own and others' performance. We conducted a longitudinal study tracking 79 participants' evaluations of their own and others' performance on five repetitions of a task over a four-month period. Three temporal factors that Radhakrishnan, Arrow, and Sniezek (1996) identified as influences on self evaluations of performance were examined: Temporal Perspective, Time Horizon, and Experience. In the present study, we investigated in more detail, the role of these factors, on judgments evaluations at multiple time-points before and after each task performance event. Results show that in general, evaluations of own and others' performance as well as on social comparisons. Participants made evaluations at multiple time-points before and after each task performance event. Results show that in general, evaluations of own and others' performance and social comparisons both had a positively leniency bias. This bias in self evaluations and social comparisons decreased when estimates were made (a) after performance than before; (b) closer to the performance event than farther away from it; and (c) with increasing experience. However, evaluations of only one's own performance

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Economics, Business and Finance*, edited by Klaus P. Hofmann published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Address Correspondence to: Phanikiran Radhakrishnan, Management Department, University of Toronto at Scarborough, 1265 Military Trail, Scarborough ON Canada M1C1A4, Fax 416 287 7363. phanira@utsc.utoronto.ca. www.utsc.utoronto.ca/~phanira. Voicemail 416 287 7319.

were more variable with changes in the temporal factors. Further, the increase in bias with longer time horizons was reduced considerably with increasing experience. Changes in inter- and intra-individual validity followed those for bias. Interestingly, changes in solo evaluations over time were similar to those for social comparisons.

Charting the Course of Self Evaluations and Social Comparisons over Time Evaluative responding is assumed to be an automatic part of human behavior (e.g., Markus and Zajonc, 1985; Jarvis and Petty, 1996). Festinger (1954) made the astute observation that the evaluative process tends to involve social comparison when others are present. The act of comparing some aspect of oneself to another is meaningful, and may change behavior. For example, it can fulfill a wide variety of needs (Helgeson and Mickelson, 1995; Sedikides and Strube, 1997; Wood, 1989). And when the comparisons to others concern performance, they can change one's performance, for example by cueing competition (Seta, 1980). Thus, it should be natural and useful for people to evaluate their own performance in comparison to that of others performing the same or similar tasks.

Several empirical studies indicate that social comparison is indeed important to those in social performance settings. For example, Levin and Levin (1973) gave students the choice of obtaining either one of two pieces of information about their performance on a test: percent correct or score in percentiles. In general students preferred the social comparison feedback. In a study conducted in a similar setting, Suls and Tesch (1978) found that students were particularly interested in learning the average class score on the test.

Not only are people interested in social comparison information, they are affected by it in ways that may alter subsequent behavior concerning the task. For example, perceived competence has been shown to be affected more strongly by information about one's competence relative to others than information about absolute performance quality (Sansone, 1986). People are particularly motivated to compare their performance to that of others when there is a mix of knowledge of feedback and uncertainty (Brickman and Berman, 1971)-a condition that characterizes most performance settings.

But what do these social comparisons of performance look like? How do people make them? In this work, we seek to understand more fully the way in which people evaluate their performance in social contexts. We examine social comparisons of performance, and compare them to solo evaluations, i.e., judgments of one's performance without explicit reference to any other performers. In addition, we include the study of social performance judgments, the appraisals of the performance of others that are inherent to social comparisons of performance. The picture of self evaluation that we attempt to provide illustrates the relationships among these three types of judgments-solo evaluations, social performance judgments, and social comparisons. It also shows the correspondence of the judgments to actual performance levels.

To obtain a more complete picture of the way in which people evaluate their performance, we ask how social comparisons of performance-as well as self evaluations and social performance judgments--change over time as performance itself changes. We turn next to a discussion of the role of time in shaping self evaluations of performance.

# **SELF EVALUATION OVER TIME**

The importance of comparing one's performance to that of others prior to the task is illustrated in a line of research initiated by Spence and Helmreich (1983). They identified differences in achievement motivation that manifest themselves in terms of either mastery or performance goals. Whereas the relevant standard with mastery goals is internal, it is external with performance goals. People with performance goals desire to be better than others. While those with such a competitive personality may enjoy superior performance (Harackiewicz, Barron, Carter, Lehto, and Elliot, 1997), they also suffer particularly at lower levels of perceived competence (Dweck and Legget, 1988). This research area affirms the importance of the performance of others as a comparison standard before performing the task. But it does not reveal whether those with such goals have a good sense of how others will perform, and how they will stand in comparison. Further, there is no indication of how their views will change as the performance event draws near, passes, and reappears.

Few studies provide thorough data on social comparisons of performance, and none that we have found examine social comparisons of performance over time. In those studies where the standard of comparison is social performance, time is typically ill defined. That is, the performer is asked about his or her relative standing with respect to the average or entire distribution of co-performers for some indefinite period (e.g., Klar and Gilladi, 1997). It is not customary to define the relevant period of performance to the participant, or to report elapsed time between the task and the social comparison judgment in research articles.

Yet, time can make an enormous difference in the evaluation of one's own performance. Although neither their theory nor data concern social comparisons in particular, Radhakrishnan, Arrow, and Sniezek (1996) address the role of time in detail. They present a model proposing that the accuracy of self-evaluations of performance varies with temporal variables due to shifts in motivational and informational influences. Specifically, Radhakrishnan et al. hypothesize that the performer's goals and information available to the performer vary with two time variables. These are (a) *Temporal Perspective*, or the direction of time between evaluation and task performance (i.e., before vs. after task performance) and (b) *Time Horizon*, or the amount of time between evaluation of performance and task performance.

According to the Radhakrishnan et al. model, information and motivational states shift as the performer moves from a future to a past perspective. Prior to the task, the performer has some uncertainty about task requirements, yet wants to maximize performance. Both information and motivation change once the task has been completed. Then the performer has information about the task but can no longer make a difference in the quality of performance. One salient motive at this point is to defend one's self esteem. These shifting motivations and information levels combine to make people more overconfident about their performance evaluations prior to the task, and much less confident once the task has been completed. Radhakrishnan et al. report an empirical study that shows support for the effects of Temporal Perspective predicted by their model. Students overestimated their score on the first quiz before taking it, and while still overconfident in their estimates, became significantly more accurate afterwards.

As for the role of Time Horizon, Radhakrishnan et al. (1996) hypothesized that the magnitude of optimistic bias increases with time horizon. That is, self-evaluations should be most optimistic long before and long after the time of the task-when the least information is available from the environment or memory. Their data show support for their prediction: the students were most optimistic well in advance of the performance event with the amount of bias decreasing as the event approached. Data on related phenomena point to similar conclusions. Nisan (1972) reports that people were more risky in their predictions and had higher expectations of success when they made predictions four weeks before, than immediately before task performance. In a study examining the effects of time horizon on prospective and retrospective judgments of performance on hypothetical tasks, Gilovich, Kerr and Medvec (1993) found that evaluations were most optimistic as time to, and form a performance event increased. A final prediction from the work of Radhakrishnan et al., (1996) was a dampening of the effects of Temporal Perspective and Time Horizon with repeated task experience. This was expected to occur as long as experience allowed one to become well-informed about task demands as well as one's performance. This too was supported in their empirical work; the negative effects of time horizon and temporal perspective on accuracy were strongest for the first of three quizzes and weakest for the third quiz.

Two limitations of the previous study (Radhakrishnan et al., 1996) are noteworthy. First, time horizons for self-evaluations before and after the quizzes were not symmetrical. Consequently time horizon and temporal perspective were confounded. Second, there was no manipulation of time horizon following the task. But considering the parallel performance of multiple people, perhaps the most important restriction of our study was that the self-evaluations were made with respect to only a performance scale, and not to the performance levels of others. Although participants were asked to estimate the average for the sample, they never provided comparisons of their performance to this average. This is particularly unfortunate given that social comparisons were likely to have been automatically generated and quite important to the participants. It is therefore unclear how similar solo evaluations of performance are to social comparisons in terms of the effects of time horizon and temporal perspective. The present research is designed to provide an empirical answer. Next, we turn to theoretical bases for differences between these self evaluations.

# SOLO EVALUATIONS VS. SOCIAL COMPARISONS

There are four reasons to suspect that the two types of self-evaluations- solo evaluations and social comparisons -- differ. The first concerns the availability of a standard for comparison; the second concerns motives regarding one's performance in comparison to the standard. Social comparison, by definition, involves only comparative judgment. Absolute judgment is possible only with non-social self-evaluation, as when one assigns a value to performance to represent the number of products produced. It is well known that absolute and comparative judgment differ (Biernat, Manis, and Kobrynowicz, 1997). Thus we can assume that the self-evaluation process is unique when solo evaluation results from absolute judgment. Of course this is not always the case; a comparison process is inevitable when one has defined performance goals with the same metric. But direct comparison is inhibited by vague representations of goals (Scheier and Carver, 1983), thus comparative judgment is not necessarily part of self-evaluation for tasks with great personal importance. In other words, solo evaluation of the number of books read results from a comparative judgment process if one's goal is to "Read 25 books", but not if one's goal is to "Read a lot".

Nevertheless, the processes leading to solo evaluation are different from those leading to social comparisons even when both kinds of judgments are comparative in nature. This is

because, with a solo evaluation, behavior regarding the comparison is directed at reducing the discrepancy between the goal and judged performance levels. From a control theory perspective on self-regulation, performing above or below the standard triggers a change in behavior directed at achieving the standard (Carver and Scheier, in press). In contrast, there is no inherent motive regarding self-other discrepancies. Seta (1982) reports data showing how one can be indifferent to assorted discrepancies: performance levels of coactors had no effect on one's own performance if the level was inferior, identical, or very superior. Alternatively one could strive to achieve a large, or small discrepancy. That is, one might prefer to be like everyone else, to be much better, or to be much worse.

A third reason to consider that the processes of solo evaluations and social comparisons may be distinct concerns one's degree of control over the comparison standard. One's goal is under one's control while the performance of the referent group is not (except, perhaps, under extraordinary conditions). Finally, there is a general difference in uncertainty surrounding the comparison standard: One knows one's goal but one must estimate the performance (or the goal) of the referent group. To summarize the differences between solo evaluations and social comparisons, the latter always involve comparative judgment while the former may or may not. Even if comparative judgments are involved in some types of solo evaluation, it is still a distinct process from social comparison. Specifically, preferences and behaviors concerning the self and a standard, and the degree of control and uncertainty regarding the standard differ greatly for solo evaluations and social comparisons. Thus, there are reasons to expect that different pictures of self evaluation of performance will result from solo evaluation and social comparison.

Social Performance Judgments. Another set of research questions concerns social performance judgments, or the performer's appraisal of the co-actors' performance. In the present study, we compare solo evaluations with social performance judgments. As Prentice (1990) commented, it is surprising how few social psychological studies directly compare judgments about the self and judgments about others. Here we define a solo evaluation as estimate of one's own score and a social performance judgment as an estimate of the average score for the class. Our general task is to identify differences between participants' solo evaluations and social performance judgments. Specifically, we examine how these judgments pertaining to the self and others vary with Temporal Perspective, Time Horizon, and Experience. Although Radhakrishnan et al. (1996) showed that these factors could combine to influence solo evaluations, there are no comparable data for social performance judgments.

# **CURRENT STUDY**

The current research effort attempts to understand more fully the pattern of changes in self-evaluations of performance overtime and in a social context. Whenever one has coperformers, there are really two social components in the process of evaluation: How one judges the performance of others, and how one compares one's own performance to that of others. Thus, we obtain both social performance judgments (evaluations of the performance of the reference group in the form of estimates of the class average) and social comparisons (evaluations of one's performance relative to that of others in the form of estimates of one's percentile rank). The social judgments and social comparisons are in addition to evaluations of one's own performance with respect to an absolute standard. Data for all three types of judgments make it possible to determine whether social comparisons change with time and experience due to perceived changes in judgments about the performance of oneself, the reference group, or both. For example, suppose I decide that I will do more poorly than the average member of the sample will. This might be because I now think less of my performance, or because I now judge the sample even better than I previously thought.

The first goal of this research is to report systematic variations judgments of performance within the period surrounding the task performance event (i.e., temporally close vs. temporally farther away from the event and before vs. after the event) as well as over repetitions of similar tasks. Put another way, our purpose is to describe and explain the three types of performance judgments -solo evaluations, social performance judgments, and social comparisons as a function of Time Horizon, Temporal Perspective, and Experience. Of special interest are differences in the patterns for solo evaluations, social comparisons, and social performance judgments.

Another goal is to evaluate the quality of these three types of judgments with respect to actual performance. Of course, this requires an objective criterion measure of performance for the entire sample of performers. Given a criterion measure, it is possible to describe the correspondence a judgment and the criterion in two different ways. First is by calculating the accuracy or distance between the judgment and criterion. The second way is to assess validity by computing the correlation between the judgment and criterion. Although correlational and distance measures are related, they are not the same (Sniezek and Reeves, 1986) and will not necessarily be affected by the same variables. Thus we use both to assess the correspondence between each performance judgment and the appropriate criterion.

Our study was conducted in a 15-week advanced course in psychology. Several researchers (Carver and Scheier, 1994; Campion and Lord, 1982; Radhakrishnan et al., 1996) have noted the advantages of conducting longitudinal research in an academic, classroom setting. A university course provided an ideal setting in the case of the present study because it involved a large number of persons repeatedly performing identical tasks in parallel.

More importantly, the classroom setting permits broad generalization. Consider how natural settings in which coactors perform identical tasks independently but concurrently tend to offer feedback, whether formal, as in statistics in sports performances, or informal, as in observations or personal communications. The course used for this study provided substantial feedback about participants' own quiz scores, and the distribution of quiz scores for the sample the participant was a part. This feedback promoted learning about the task as well as about the quality of evaluations. Both absolute and relative standing contributed to the determination of valued outcomes for participants' place in the distribution were all meaningful to participants. A final useful feature was that the classroom context allowed for social interactions among participants over the time-period during which the study was conducted. In sum, the classroom setting made social comparisons and social performance judgments meaningful and possible for participants.

Participants made three types of performance estimates - solo evaluations, social performance judgments, and social comparisons. These judgments were made for each of five quizzes at multiple points over the duration of the course. This setting made it possible for us to vary Temporal Perspective, Time Horizon, and Experience. We varied Temporal Perspective by asking participants to make estimates of their performance before and after each task performance event, but before receiving feedback about that particular task

performance. In addition, Time Horizon, or the amount of time elapsed between the point of evaluation and the performance event was varied by having participants make judgments about their performance several weeks, three weeks, a few days, and a few min. before, and after, each performance event. The third factor, Experience, concerned systematic changes in evaluations of performance over five discrete performance events of similar tasks.

The performance task was an in-class, non-cumulative quiz consisting of 20 multiplechoice items with four-alternatives each. The task was highly involving for the studentparticipants enrolled in the course who were presumably motivated to maximize performance. Although a multiple choice quiz was not entirely novel for upper-class university students, the combination of a new course, a new instructor, a new teaching-assistant, and novel content produced sufficient uncertainty about task performance, made performance evaluations non trivial, and feedback about task performance informative. Five regularly scheduled quizzes provided multiple performance trials so that we could investigate the dynamic aspects of performance evaluations. A special advantage of this task was that quiz scores provided an objective external criterion for evaluating the quality of performance evaluations, and offered accurate and unambiguous feedback that promoted favorable learning conditions. This in turn enabled subsequent evaluations more accurate, less biased and more valid.

## Method

#### Participants

The 79 undergraduates enrolled in the upper-level course made multiple evaluations of their performance on each of the five in-class quizzes as partial fulfillment of course requirements. Participants were informed that their responses would have no impact on their grades and that their evaluations would be kept sealed until the course was completed and final grades were submitted to the registrar. This information was provided to ensure, and assure participants of the confidentiality and consequences of their evaluations. All student-participants consented to the use of their evaluations for research purposes. Participants were assured that providing consent to the use of their evaluations for research purposes was not part of the class requirement. The course was taught by one of the authors. However, a research assistant unrelated to course administrative duties collected students' performance evaluations.

#### Procedure

Several estimates were collected before and after performing each quiz. The usual sequence was to obtain the initial judgments for a quiz two-three weeks before (on the date of the previous quiz, with the exception of the first quiz). Additional evaluations were made two days before, a few min. before on the day of the quiz, a few min. after the quiz, and finally, again two days after the quiz. Note that all evaluations made after the quizzes were done prior to receiving feedback.

#### Measures

*Elicited Measures.* Participants provided three types of performance estimates at each time-point of data collection, by writing privately, on a form containing instructions for each type of estimate. One performance estimate was the solo evaluation, an estimate of one's own raw score, or specifically, the number correct out of 20. A second type was the social performance iudgment, the estimated class average or specifically, the average number correct out of 20 for the entire class. A third type was the social comparison iudgment, an estimate of one's percentile rank, or the percentage of class members scoring below the individual's estimated raw score.

*Derived Measures.* To describe fully the quality of each type of judgment, we computed the accuracy and validity for each of the three types of estimates we elicited.

Accuracy. Both the absolute level of accuracy and the direction of inaccuracy are of interest in describing the quality of performance judgments. We defined accuracy to include Mean Absolute Deviation (MAD) and Mean Bias (MB). For a set of judgments  $\{J_1, J_2, ..., J_i, ..., J_k\}$  from k individuals with criterion values  $\{Y_1, Y_2, ..., Y_i, ..., Y_k\}$ :

MAD = 1/k	$\frac{K}{\sum  J_i - Y_i }$	[1]
	i=1	

Mean Absolute Deviation (MAD) quantifies the magnitude of error in a set of judgments. It is useful to show the extent to which estimates deviate from the criterion. However, it does not yield any information about the direction of the deviation.

In contrast, Mean Bias indicates whether the judgments overestimate or underestimate the criterion. It is equal to the mean signed deviation between the judgment and criterion:

K		
Mean Bias = $1/k \sum  J_i - Y_i $		
i=1	[2	]

MAD and Mean Bias were computed for solo evaluations (estimates of own scores) and social comparisons (estimates of percentile ranks). The criterion variables were, respectively, actual own score and actual percentile rank. For class averages, MAD and MB were obtained by taking the absolute difference between the estimates of the class average and the obtained class average and by taking the signed difference between the two respectively. Mean bias allowed us to examine the direction of inaccuracy and its changes over time and judgment type are examined which we report in detail in this paper.

*Indirect social comparisons.* We define indirect social comparison as the discrepancy between a solo evaluation and a social performance judgment. Whereas the performer makes a direct comparison with social comparisons, we make the comparison with indirect social comparison. We computed both the signed deviations and absolute differences between solo and social performance judgments. These discrepancies provided an alternative perspective on differences in judgments of self vs. others. In addition, they allowed us to diagnose whether social comparisons were inaccurate due to poor social judgment or poor comparison ability.

*Validity.* There are two meaningful types of validity coefficients. The correlation between the judgment and criterion over a set of persons, called *inter-individual validity*, tells whether the members of a sample can differentiate among their actual performance levels. As such, it is an index of quality at the level of the sample. An alternative, *intra*¬*individual validity*. is assessed by the correlation between a single individual's judgments and that individual's actual scores over multiple tasks. This provides information about quality at the individual level -- whether an individual's judgment can capture true changes in his/her criterion performance over repeated evaluations of a performance event and over repeated performance of similar tasks.

We assessed *inter-individual validity* by correlating performance estimates with actual performance scores over persons within each trial. Inter-individual validity was computed for solo evaluations (by correlating estimated and obtained raw scores for the individual) as well as for social comparisons (by correlating estimated and actual percentile ranks). It is not possible to calculate an analogous validity coefficient for social judgments because the actual class average does not vary within a trial. For *intra-individual validity* coefficients, we computed a correlation between solo evaluations and the corresponding actual scores for each individual.

# RESULTS

#### **Overview of Analyses**

Our analyses are organized around the three main questions motivating this study. While all the questions pertain to the effects of three factors--Temporal Perspective; Time Horizon, and Experience--on self-evaluations, they differ in terms of the particular self evaluation judgments of interest. The first question asks how solo evaluations (estimates of raw scores) and social performance judgments (estimates of class averages) varied in accuracy and validity with changes in Temporal Perspective, Time Horizon, and Experience. We analyzed solo evaluations and social performance judgments as two levels of a within-subjects factors because we were interested in whether the three factors affected the accuracy and validity of appraisals about the performance of oneself and that of others differently.

The two remaining sets of analyses help us identify when individuals see themselves as most different from, and most similar to their peers. Specifically, they address questions about the way in which Temporal Perspective, Time Horizon, and Experience influence social comparisons. The second question addresses the effects of these factors on the accuracy and validity of direct social comparisons (estimates of percentile ranks) while the third question concerns their impact on indirect social comparisons (discrepancies between solo evaluations and social performance judgments). The combination of these two analyses allows us to determine the extent to which social comparisons are inaccurate due to poor social judgment or due to poor comparison ability. In addition to the third question about the accuracy and validity of solo and social performance judgments, we were interested in the discrepancy between these two judgments as a dependent variable in and of itself. This discrepancy reveals the extent to which people judge their own performance to be different from that of the average other person. Thus, we also performed separate analyses on the discrepancy between solo and social performance judgments.

# *Question 1. Accuracy and Validity of Solo Evaluations and Social Performance Judgments*

*Changes in Bias.* To examine the separate and interactive role of Temporal Perspective, Time Horizon, and Experience, we conducted three different sets of analyses on measures of mean bias'. First, we tested the role of a short Time Horizon and Experience on the mean bias in solo evaluations and social performance judgments. We used type of Estimate (Two Levels: Solo vs. Social Performance) x Time Horizon (3 Levels: 1 Quiz Before vs. 2 Days Before vs. Directly Before) x Experience (Five Levels: Quizzes 1 through 5) [6] This resulted in a 2 x 3 x 5 repeated measures ANOVA factorial design.

In addition, we tested the role of a long prospective Time Horizon and Experience on the bias of solo evaluations and social performance judgments [7]. This analysis had the same two types of Estimates (Solo vs. Social Performance), but a larger range of Time Horizon (4 Levels: 2 Quizzes Before Target Quiz vs. 1 Quiz Before Target Quiz vs. 2 Days Before Quiz vs. Directly Before Quiz) and a smaller range of Experience (3 Levels: Quiz 3 vs. Quiz 4 vs. Quiz 5). This was a 2 x 4 x 3 repeated measures ANOVA [8].

Finally, we investigated the roles of Temporal Perspective, Time Horizon, and Experience. We conducted 2 x 2 x 2 x 4 repeated-measures ANOVA on the bias of two types of Estimates (Solo vs. Social Performance) with Temporal Perspective (2 Levels: Pre-task vs. Post-task), Time Horizon (2 Levels: 2 Days Away from the Quiz vs. Directly Away from the Quiz) and Experience (4 Levels: Quizzes 1, 2, 4, and 5) [9]. Due to different combinations of missing data, the three sets of analyses have different sample sizes. The design examining the effects of a short Time Horizon has 30 participants, the design exploring the effects of a long prospective Time Horizon has 38, and finally the design examining the effect of both Time Horizon and Temporal Perspective has 23 participants with complete data.

*Temporal Perspective.* Recall our prediction that post-task evaluations would be less biased than pre-task evaluations. As can be seen in Figure 1, both solo and social performance judgments had higher mean bias when they were generated before the task was performed (M=.92) than after it was performed (M=.50, F (1,21)=31.69, p<.001).

We predicted that solo and social performance judgments generated at points that were temporally closer to the performance event would be less biased than those made at temporally distal points. Again we found support for our predictions. Figure 2 shows that with a shorter time horizon (i.e., when evaluations were made at the three time-points before the performance event) mean bias decreased as individuals approached the task performance event, F (2,56)=57.85, p<. 0001. Similarly, Figure 3 shows that with a longer time horizon (i.e., when evaluations were made at four timepoints before the performance event) mean bias decreased much more dramatically, F (3,108)=51.21, p<. 0001. Figure 1 also shows that mean bias decreased with changes in Time Horizon even when evaluations were made before versus after the task performance event but before performance feedback from an objective source, F (1,21)=20.61, p <. 0001. In general, mean bias decreased as evaluations were made at points that were temporally closer to the performance event than those that were made at points that were made before to the performance event than those that were made at points that were temporally closer to the performance event than those that were made farther away from it.

We predicted that evaluations made after experience and feedback (i.e., in later quizzes) would be less biased. We found that as students gained experience evaluating their

performance and received feedback about each performance event, mean bias in their solo evaluations and social performance judgments decreased. Further, mean bias decreased with increasing Experience when the Time Horizon was short, F (4,112)=30.08, p< .0001 (see Figure 2), when the Time Horizon was long, F (2,72)=21.73, p< .0001 (see Figure 3) and when Temporal Perspective varied with Time Horizon and Experience, F (3,63)= 10.93, p< .0001 (see Figure 1).

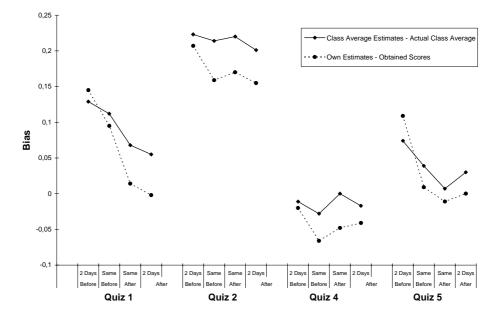


Figure 1. Bias in solo evaluations and social performance judgments as a function of Temporal Perspective and brief Time Horizons.

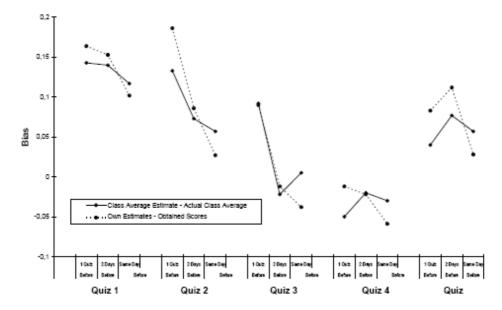


Figure 2. Bias in prospective solo evaluations and social performance judgments over short Time Horizons.

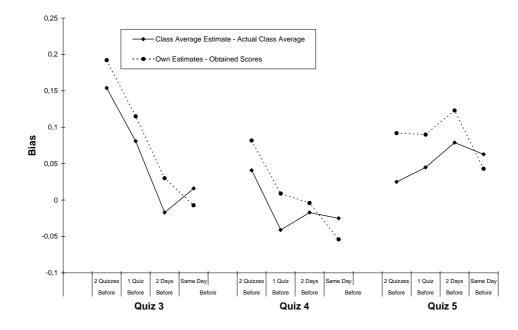


Figure 3. Bias in prospective solo evaluations and social performance judgments over long Time Horizons.

We predicted and found that the effects of Time Horizon on the bias of evaluations attenuated as individuals gained experience evaluating their performance (i.e., Temporal distance by Experience interaction). This was true for both solo and social performance judgments. Further, we found this pattern when the Time Horizon was short, F (8,224)=9.97, p<.0001 and when the Time Horizon was long, F (6,216)=16.77, p< .0001. Figures 2 and 3 show that at Quizzes 4 and 5, there was no significant difference (p> .05) in the mean bias of evaluations that were made farther away from the performance event than those that were made closer to the performance event. In contrast, at earlier quizzes (i.e., 1,2,3) there was a significant difference (p< .05) in the mean bias of evaluations that were made temporally closer, when compared to those that were made temporally farther away from the performance event.

In addition, we found a three-way interaction between Time Horizon, Experience and Temporal Perspective, F (3,63)=3.63 p < .02. As can be seen in Figure 1 post-hoc analyses showed that although increasing experience attenuates the bias caused by increasing by temporal distance, the bias does not completely disappear - it reappears when retrospective judgments are made. This suggests that memory distortion may also play a role in shaping evaluations, despite increasing experience with evaluating performance and receiving feedback regarding the quality of one's evaluations.

We found that across the different Time Horizons, solo evaluations had larger and more consistent decreases in their mean bias than did social performance judgments. This was supported by a significant interaction between Type of Estimate and Time Horizon when the Time horizon was short, F (2,56)=22.71 p<. 0001, when the Time horizon was long, F (3,108)=19.38 p<.02 (see Figure 1).

It is of interest to note that in general, people over-estimated their performance rather than under-estimated it (see Figures 1, 2, and 3). An exception is that overestimation disappeared at points shortly in advance of Quizzes 3 and 4 (see Figure 3). One possible explanation for this is the difficulty of Quiz 2; the class average for Quiz 2 was lower than that of the other four quizzes (p< .05). Therefore, students may have under-estimated their performance on Quizzes 3 and 4 following feedback from Quiz 2 in anticipation that Quizzes 3 and 4 would be as difficult as Quiz 2. However, once Quiz 3 was completed and feedback was received, the expectation was revised.

In addition, we also found that the relative degree of bias in solo evaluations and social performance judgments changed with Time Horizon. The dominant finding prior to the quizzes was for more bias in solo evaluations than in social performance judgments. This pattern was evident when estimates were made two days before the quiz, one quiz before, and two quizzes before (see Figures 1, 2, and 3). However, directly before each quiz, the pattern reverses: solo evaluations show lower levels of positive bias than did estimates of the class average. This was true when for the relative differences in the bias of solo and social performance judgments analyzed with a short time horizon, <u>F</u> (2,56)=22.71, <u>p</u><. 0001 (see Figure 2), a long time horizon, F (3,108)=19.38, p<.0001 (see Figure 3) and in the context of temporal perspective F (1,21)=6.63; p<.02 (see Figure 1) [10]

*Changes in Validity.* To test our hypotheses about changes in the *inter-individual validity* we conducted a pattern analysis using binomial probabilities. We tallied the number of times the hypothesized pattern of correlations matched the obtained pattern of correlations by making multiple pair-wise comparisons. For example, if the correlations between estimates (and actual scores) made before Quiz 1 were .33, .21, .23 and those made after Quiz 1 (but before receiving feedback on Quiz 1) were .35, .29, .20, we concluded that five out of the six comparisons matched our predicted pattern. To test our hypotheses about the changes in the *intra-individual validity*, of solo evaluations we computed for each individual, a correlation coefficient between that person's solo evaluations and actual scores. These estimates were ones made at three different time horizons before the quiz: one quiz before the target quiz, two days before the target quiz, and immediately before the target quiz." We converted these coefficients to z scores and conducted a repeated measures ANOVA using Time Horizon as the 3-level within-subjects variable. For both intra and inter-individual validity, we used only those solo evaluations that were made before receiving feedback about actual performance [12].

We wanted to examine the effects of Temporal Perspective, Time Horizon and Experience on the *inter-individual validity* of solo evaluations. We hypothesized that the validity of estimates made after performing the task should be greater than that of estimates made before performing the task. As predicted, post-task solo assessments were highly correlated with actual performance than were pre-task predictions. The obtained correlation coefficients matched the expected pattern of correlations 41 times out of the 47 possible comparisons, resulting in a binomial probability of p < .0001,  $\binom{47}{C_{41}}(.5^{41})(.5^{6})$ . We also found that mean pre-task coefficients were lower than post-task coefficients. Using a Fisher r to z transformations we found that the mean pre-task correlation for solo evaluations was lower (r= .30) than the post-task correlation (r=.45).

We hypothesized that the predictive validity of self-evaluations should increase with decreasing Time Horizon. That is, estimates made temporally closer to the time of task should be more predictive of actual performance than should estimates made at temporally distant points. We expected that when estimates were made temporally closer to the quiz (in either direction), the correlation between those estimates and performance should correspondingly increase. We found that correlations between solo evaluations and actual scores (aggregated over Quizzes) decreased

as Time Horizon from the Quiz increased in either direction. These correlations followed the predicted pattern 50 out of the 63 comparisons, p < .0001,  $\binom{63}{C_{50}}(.5^{50})(.5^{13})$ .

We also predicted that as individuals gained experience predicting their performance, their estimates of performance should become increasingly valid. Our hypothesis regarding the effects of Experience on participants' ability to align their predictions of performance to their actual performance was also supported. Correlations across the five quizzes increased. Solo evaluations for later quizzes were more highly correlated with actual scores than were solo evaluations for earlier quizzes. The data for solo evaluations matched the predicted pattern 34 out of 48 times, p < .01,  $\binom{48}{C_{34}}(.5^{34})(.5^{14})$ .

*Intra-individual validity* coefficients were computed for each solo evaluation generated at three different time points (i.e., pre-performance, pre-feedback time points). There were no significant differences between the three time horizons. The validity of solo evaluations made one quiz before the target quiz was -.20, that of evaluations made 2 days before the target quiz was -.25 and the validity of solo evaluations made immediately before the target quiz was -.17.

#### Question 2. Social Comparisons

*Changes in Bias.* As for solo and social performance judgments, we conducted three sets of repeated measures ANOVAs on mean bias of social comparisons. One design examined the effects of Time Horizon (3 Levels: 1 Quiz Before vs. 2 Days Before vs. Directly Before) and Experience (Five Levels: Quizzes 1 through 5) on the bias of percentile ranks. The second examined the effects of a longer Time Horizon (4 Levels: 2 Quizzes Before Target Quiz vs. 1 Quiz Before Target Quiz vs. 2 Days Before Quiz vs.

Directly Before Quiz) and a smaller range of Experience (3 Levels: Quiz 3 vs. Quiz 4 vs. Quiz 5). The third examined the role of Temporal Perspective (Pre vs. Post Task) a short Time Horizon (2 Levels: 2 Days Away from the Quiz vs. Directly Away from the Quiz) and Experience (4 Levels: Quizzes 1, 2, 4, and 5) on the bias of percentile ranks.

We predicted that percentile rank estimates made after task-performance should be less biased than those made prior to task-performance. We found support for this hypothesis. Figure 4 illustrates the general finding that students were less biased in evaluating their performance after task performance than before task-performance, F (1,21)=10.86 p < .005. In addition, we also found two significant interactions. Figure 4 also shows that the biggest difference in bias between pre- and post-task percentile ranks was in the first quiz. This was supported by a significant interaction between Experience and Temporal Perspective, F (3,63)=6.69 p < .002. When social comparisons are generated after task performance, mean bias in percentile ranks does not change much with increasing time horizon. This pattern was supported by a significant interaction between Temporal Perspective and Time Horizon, F (1,21)=10.25, p< .005.

We predicted that percentile rank estimates made at points that are temporally closer to the performance event should be less biased than those made at temporally distant time-points. We found support for this hypothesis. In general, when individuals estimated their percentile ranks at points that were temporally closer to the time of task performance they were less biased. We found that bias in percentile ranks decreased when Time Horizon was short F (2,56)=12.98, p<.001 (see Figure 5), long F(3,108)=14.12, p<.001 (see Figure 6), and when estimates of percentile ranks were made prospectively or retrospectively, F(1,21)=4.62, p<.05 (see Figure 4).

We predicted a reduction in bias as individuals gained more experience estimating percentile ranks. However, we did not find the predicted effects for Experience. We predicted that with increasing experience, the effects of Time Horizon on the bias of percentile ranks should attenuate. Again, we did not find the predicted interaction between Experience and Time Horizon [14].

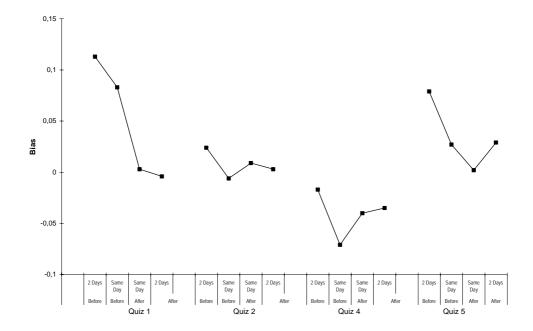


Figure 4. Bias in social comparisons as a function of Temporal Perspective and brief Time Horizons.

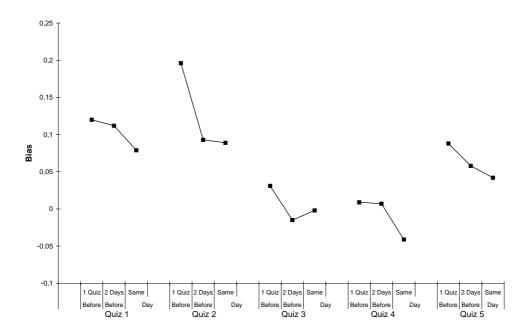


Figure 5. Bias in prospective social comparisons over short Time Horizons.

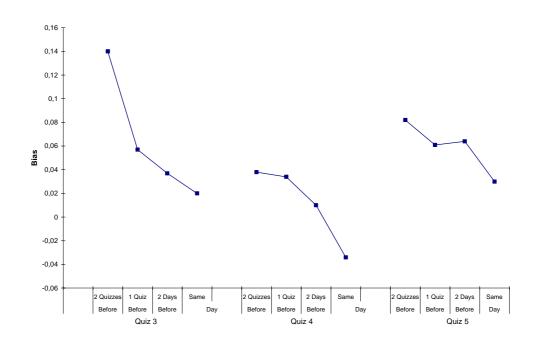


Figure 6. Bias in prospective social comparisons over long Time Horizons

Changes in Validity. As for solo evaluations, we conducted similar analyses to examine the changes in inter and inter-individual validity of percentile ranks. In general, our results supported our hypotheses for changes in the inter-individual validity. To summarize, the changes in the validity coefficients for solo evaluations and social comparisons were similar (see Figure 7). Temporal Perspective affected the validity of social comparisons. Estimates of percentile ranks made before task performance had lower correlations with actual percentile ranks than did estimates of percentile ranks made after task performance. Obtained correlations matched the hypothesized pattern of correlations 33 times out of the possible 46 comparisons, resulting in an binomial probability of  $\underline{p} < .002$ ,  $\binom{46}{C_{33}}$  (.5<sup>33</sup>) (.5<sup>13</sup>) [15]. Similarly, the Fisher r to z transformation showed that the pre-task correlation for percentile rank estimates was lower (r= .33) than the post-task correlation (r= .39). Correlations between estimated and actual percentile ranks also increased as Time Horizon became shorter. The obtained pattern of correlations followed the pattern 40 times out of the 64 possible comparisons, p < .02,  $\binom{64}{C_{40}}(.5^{40})(.5^{24})$ . Thus, the validity of percentile ranks increased with decreasing Time Horizon. Correlations between estimated and obtained percentile ranks increased with greater Experience. The validity coefficients for percentile ranks matched the predicted pattern 32 out of 49 times, p < .02,  $(^{49}C_{32})(.5^{32})(.5^{17}).$ 

The changes in intra-individual validity coefficients for percentile ranks across the three Time Horizons were marginally significant, F (2,1,04)=1,9 p=.15. In general, there was an increase in the validity of percentile ranks as time horizon decreased. The validity of percentile ranks generated one quiz before the target quiz (M=-.35) were lower than those made 2 days before the target quiz (M=.34). A post-hoc t-test suggested that the validity of percentile ranks generated immediately before the target quiz (p<.05) higher than from those made one quiz before the target quiz.

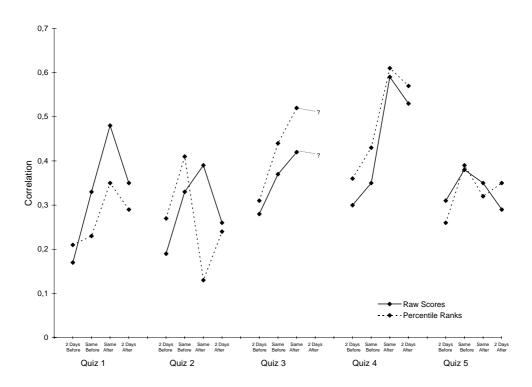


Figure 7. Inter-individual validity coefficients for solo evaluations and social comparisons as a function of Temporal Perspective and brief Time Horizons.

Question 3. Discrepancies between Solo and Social Performance Judgments

*Changes in Bias.* Here again, we conducted three sets of repeated measures ANOVAs on the discrepancy between solo and social performance judgments. One ANOVA examined the effects of Time Horizon (3 Levels: 1 Quiz Before vs. 2 Days Before vs. Directly Before) and Experience (Five Levels: Quizzes 1 through 5) to form a 3x5 factorial design. The ANOVA second examined the effects of a longer Time Horizon (4 Levels: 2 Quizzes Before Target Quiz vs. 1 Quiz Before Target Quiz vs. 2 Days Before Quiz vs. Directly Before Target Quiz vs. 2 Days Before Quiz vs. Directly Before Quiz) and a smaller range of Experience (3 Levels: Quiz 3 vs. Quiz 4 vs. Quiz 5). The third AVOVA examined the role of Temporal Perspective (Pre vs. Post Task) a short Time Horizon (2 Levels: 2 Days Away from the Quiz vs. Directly Away from the Quiz) and Experience (4 Levels: Quizzes 1, 2, 4, and 5).

Figure 8 shows that discrepancies were lower when evaluations are made after the performance event than before. This was supported by a significant effect for Temporal Perspective, <u>F</u> (1,22)=9.12 p<. 01. Figure 8 also shows that the discrepancies between solo and social performance judgments decreased when evaluations were made at times closer to the performance event than those that were farther away from the performance event. This was supported by a significant effect for a short Time Horizon, F (1,22)=7.99 p <. 02. Finally, Figure 8 shows a significant interaction between Temporal Perspective and Time Horizon, F (1, 22)=6.18 p<. 05.

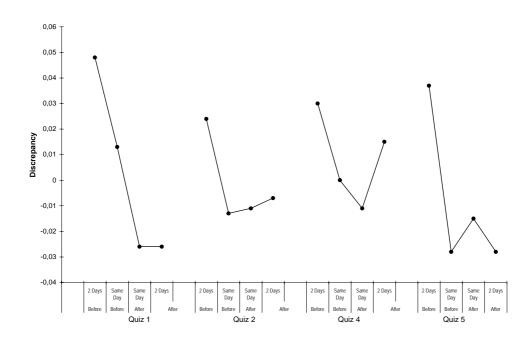


Figure 8. Discrepancies between solo evaluations and social performance judgments as a function of Temporal Perspective and brief Time Horizons.

That is, for the first two quizzes, pre-performance discrepancies between solo evaluations and social performance judgments were much higher than post-performance discrepancies between the two types of estimates. For Quiz 4 however, this pattern disappeared only to reappear for Quiz 5. We did not find similar results when we examined absolute differences between solo and social performance judgments. Only the Experience by Time Horizon interaction was replicated when the MAD between solo and social performance was the dependent variable, F (3, 66)=3.17 p < .05.

Figure 9 shows that there is a significant drop in the discrepancies between solo evaluations and social performance judgments from estimates made long before the performance event to those made right before the performance event. This was supported by a significant effect for short Time Horizon, F (3, 111)=19.55 P<. 001. Figure 9 also shows that with increasing experience, overall discrepancies between solo and social performance judgments are lower. This was supported by a significant effect for Experience, F(2, 74) =3.69, p<.05. Again, we did not find similar results when we examined the MAD between solo and social performance judgments. This can be explained by the mix of positive and negative discrepancy scores; the differences evident in signed scores cancel out with MAD.

Figure 10 shows that when evaluations were made at points that were temporally closer to the performance event, the bias between solo evaluations and social performance judgments reduces. Sometimes, solo evaluations were even lower than social performance judgments (e.g., right before quizzes 2, 4, and 5). This was supported by a significant effect for Time Horizon, <u>F</u> (2, 58)= 22.65 <u>P</u> < .001. Again, we did not find similar results when we examined MAD between solo and social performance judgments.

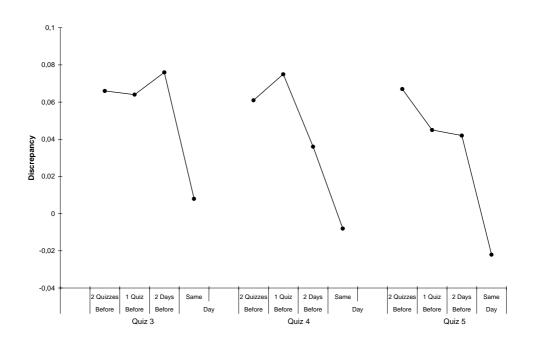


Figure 9. Discrepancies between prospective solo evaluations and social performance judgments over short Time Horizons.

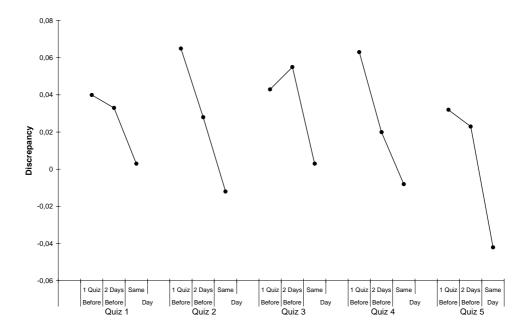


Figure 10. Discrepancies between prospective solo evaluations and social performance judgments over long Time Horizons.

## DISCUSSION

Situations with many people performing essentially the same task are common in the military, education, business, government, and sports. People in these situations can be expected to care about, and frequently evaluate their performance. Their self-evaluations of performance are best examined as social behavior. Whether people perform concurrently or at different times, an implicit or explicit referent for judgments of one's own performance is the performance of others. Consequently, social performance judgments and social comparisons become part of the process of self evaluation of performance. It is of substantial interest to know how people in these domains evaluate their own performance and the performance of others, and compare the two.

This research endeavor uses a dynamic approach to describing three distinct judgments involved in self-evaluation of performance: social comparison, social judgment, and solo evaluation. It is worth pointing out several features of our effort that distinguish it from other work on self-evaluation and social comparison. First, the dimension of interest is performance, and not the ability and trait dimensions commonly found in social comparison research. The nature of the dimension under evaluation in social comparisons has been neglected by researchers (Wood, 1989); yet it may be an important distinction. Second, the self-evaluation made by the participant concerns the performance criterion variable. In our work, it is an estimate of actual performance level on the specified task and not a judgment of competence, ability, or satisfaction. Third, the comparison standard is a group and not an individual, and is offered by the environment rather than selected by the performer. It is a factual standard defined by the average of a group (i.e., a "social category reference point" Higgins, Strauman, and Klein, 1986), where that group provided the social context for performance and evaluation. Fourth, the judgments are evaluated with respect to three measures-bias, MAD, and validity. Whereas most social psychological research attends to questions of accuracy, applied research concentrates on the validity of self-assessments. We used both.

Of course, it is always risky to generalize across natural settings. But settings characterized by coactors, delayed performance feedback, and opportunities for social interaction are comparable to the setting of our study. Our results show what Festinger (1954) hinted and Wood (1989) documented - that people are not unbiased self-evaluators. People in the settings described above are likely to have distortions in their perceptions of their performance relative to that of others. What is more important is that our study shows that the direction and magnitude of the distortion changes with temporal perspective, time horizon, and experience.

Long before the task, most people estimate performance levels for themselves that are far superior to those of the referent group. This illusion fades as the task approaches, and is replaced by a far more pessimistic evaluation following performance. Once experience with the task has provided sufficient information about performance and the accuracy and validity of self-evaluations, realism replaces bias. One interesting Explanation for the pre and post-task change in evaluations is suggested by Feather (1969). In his study, Feather found that unexpected successes were rated as more satisfying than expected ones and although the opposite pattern was found for unexpected failures (i.e., that they were less satisfying than expected failures), the pattern was not significant. These findings suggest that after task

performance, people may lower estimates in order to avoid the negative affect accompanying unexpected failures and to approach positive affect that accompanies unexpected successes.

Sheppard, Ouellette and Fernandez (1996) examined changes in predictions as temporal duration of time to feedback decreased (4 months to .2 weeks before graduating in Study 1 and 1 month before quiz, 3 days, 50 min and 3 seconds before feedback in Study 2). Although they did not distinguish between pre and post-task performance changes in predictions they predicted and found that estimates decreased as time to feedback approached.

It is not typical for participants in social comparison studies to provide explicit judgments about others. This is disappointing given that examination of social judgments of performance helps diagnose the quality of self-evaluations of performance. Bias in social comparisons can be attributed to biased evaluations of others, biased evaluations about oneself, or both. By assessing social performance judgments in addition to solo evaluations and social comparisons we can identify whether inappropriate social comparisons are due to a flawed perception of others' performance or a flawed comparison process. The general answer is that all three judgments, solo evaluations, social performance judgments, and social comparisons, were positively biased. Whether people are more accurate about themselves or others may depend on when the judgments are made. Our data showed relatively less bias about one's own performance compared to that of others after the task. The reverse was true before the task. But with short time horizons there was comparable bias in self and other judgments of performance. Because the bias in direct social comparisons generally paralleled the bias in solo evaluations, there is no reason to suspect biased evaluation processes. That is, direct assessments about one's performance relative to that of others in the form of estimated percentile ranks neither minimized nor magnified the bias in solo evaluations.

We proceed to discuss implications of our results for people in natural settings with the specified characteristics, limitations of this study, and potentially fruitful directions for research on self evaluations of performance.

## Implications for Behaviour

While this picture is interesting in itself, it is more valuable because of its implications for behaviour in these settings. People's evaluations of their performance relative to the each other may exert a broad influence on their behaviour. For example, perceptions of performance superiority enhance feelings of inequity (unless accompanied by perception of superior outcomes). Deleterious effects of inequity such as decreased motivation may be most severe long before the critical performance event. Ironically, because one forecasts superior performance one shifts to a strategy what will reduce performance, and ultimately performs far below the original forecast. If one is seeking to restore equity by conforming in terms of performance, the self-evaluations should decline monotonically up to the time of the event. Of course, the fact that this is the observed pattern in our data does not mean that perceptions of equity were responsible. Nevertheless, this pattern can provide useful insights into the judgments underlying perceptions of equity.

Other consequences of social comparisons for behavior are of considerable interest. For example, they might stimulate competition (Beck and Seta, 1980), produce assorted emotions (Higgins et al., 1986), encourage task persistence, or promote learning and performance. The important point is that for each consequence of social comparison of performance, there are likely

to be changes over the course of the time period leading up to and following the critical performance event. Only with a dynamic view such as we have presented can this source of variance be appreciated.

## Limitations of the Study

Our work follows that of Radhakrishnan et al. (1996) by using manipulations of temporal perspective, time horizon, and experience. It has some of the same limitations, namely restriction to an environment with extrinsic rewards based on a mix of absolute and relative performance. Self-evaluation processes may differ in contexts such as athletic contests and markets where the competitive edge determines all. But the present study's design overcomes several deficiencies in their work because it includes multiple time horizons with the past temporal perspective, and symmetrical time horizons for both temporal perspectives. In addition, this study has more participants, more task trials, and measures than did theirs.

Nevertheless, our research design and execution was not ideal. Methodological problems include the failure to obtain all post-measures for quiz 3, and a lack of control over fluctuations in actual scores over the five quizzes. Although missing data were relatively few in number, there may have been a disproportionate exclusion of data from students at the lower end of the performance continuum. Finally, it is possible that the repeated requests for self-evaluation estimates may have induced unnatural levels of self-monitoring. The picture is considerably brighter when viewed from the perspective of educators: mere self-evaluation could have been beneficial to participants' motivation

(Bandura, 1986) and performance (Sanna and Pusecker, 1994). A between-subjects research design would minimize the problem of artificially increasing self-evaluation frequency or salience (as well as its likely benefits to the performers) by restricting self-evaluation to only one of the time periods of interest. Note that in the present study of two temporal perspectives, three time horizons, and five task repetitions, this approach would require 30 experimental conditions.

# **PROPOSED DIRECTIONS FOR RESEARCH**

While this study has provided a detailed picture of self evaluations of performance over the course of time preceding and following performance, many questions remain. Attempts to understand the causes and consequences of social vs. solo evaluation seem especially worthwhile. Past research activity on the antecedents of social comparison has long focussed on studying the choice of a comparison standard (Wills, 1981). In light of the significance of temporal perspective and time horizon, it would be useful to ask how social comparison processes change over time. More specifically, it is desirable to know when social comparisons occur, and whether preferences for the comparison standard shift with time independently of other factors. According to the "similarity hypothesis" of Festinger's (1954) theory, people seek to compare themselves to others who are similar on the evaluation dimension. The time data from this study suggest the preferred comparison standard will change as one's self-evaluation of performance changes. Another question that can be addressed with the dynamic approach exemplified in this study is whether social performance judgments precede social comparison or are formed only in service of the comparison process. All forms of evaluation may be automatic, but they may also vary in timing and influence each other differently.

A prominent theme in studies of the consequences of social comparison has been peoples' reactions to upward and downward comparisons (Pyszczynski, Greenberg, and LaPrelle, 1985). No doubt that data on this matter are extremely valuable. But many social comparisons are not between oneself and a single individual whom one has selected for comparison purposes. The comparison standard often pertains to an entire group, and is a given for all coactors (Wood, 1989). Such is the case with distributional information, such as mean performance for the group. It is not clear that the reactions people have to comparisons with an individual apply similarly to a group. It is one thing to aggrandize the individual outperformer (Alicke, LoSchiavo, Zerbst, ad Zhang, 1997), and quite another to imagine that an entire group consists of geniuses. In general, judgments about individual persons or events are fundamentally different from judgments about a set of persons (Klar and Giladi, 1997) or events (Sniezek and Buckley, 1991).

Recent research identifies a number of individual difference variables that may lead to differential use of social comparison as well as differential reactance to social comparison information. One variable that seems to hold special promise for explaining variance in self-evaluations of performance is gender. Beyer (1990) found that gender differences in self evaluations existed for certain kinds of tasks. These may be present with time and social comparisons, if as Cross and Madson (1997) argue, that women have more interdependent self orientations and therefore have more elaborate and available information about others. This in turn implies that social comparison information may be more accurate for women in general, and may be more sensitive to informational differences over time. One may also predict that women may consistently give lower estimates of the discrepancy between own vs. others' performance because they have more information about others than men do.

A personality variable that may divide self-evaluations over time into two distinct patterns is uncertainty orientation. Uncertainty-oriented people are motivated to acquire new knowledge about themselves while certainty-oriented people have the desire to avoid ambiguity and thus attempt to maintain existing beliefs. According to Roney and Sorrentino, (1995), self-assessment is more important to the former and self-verification is more important to the latter. The difference between the two personality types becomes most intriguing over time. Certainty-oriented persons are expected to seek information about their performance in the absence of feedback. But once it is obtained, they should be resistant to inconsistent information. In contrast, those who are uncertainty-oriented are predicted to seek information about their performance from multiple sources, and to do so as long as there is something new to discover. They cease self-assessment when information becomes redundant. Thus we speculate that for the certainty-oriented, solo evaluations and social comparisons will be highly correlated, with fewer changes with temporal perspective, time horizon, and experience. The predicted for uncertainty-oriented persons is less correspondence between social comparisons and solo evaluations, and steeper slopes for evaluations as a function of time horizon magnitude, and direction of temporal perspective direction.

A newcomer to the list of individual differences variables involved in self-evaluation is happiness. A study of the hedonic consequences of social comparison by Lyubomirsky and Ross (1997) shows self-rated happy persons to be less sensitive to social comparison information than unhappy persons. They suggest that the relationship is bidirectional, meaning that happiness is a cause of selectivity in making social comparisons as well as a consequence of minimal attention to social comparison information when the comparisons are made. To establish the true nature of this link, it will be necessary to manipulate mood and affect, and to observe social comparison processes. It would be valuable if such research would include multiple temporal perspectives and time horizons, and assess the accuracy and validity of the social comparisons. It may well be that an even more intriguing pattern of differences between happy and unhappy people emerges. The present research shows how much can be learned by examining the quality of self-evaluations, and by tracking them over time.

Finally, Sheppard, Ouellette, and Fernandez (1996) found that self-esteem affected differences in how drastically people changed their estimates as the moment of feedback approached: low self esteem individuals made lower estimates as time for feedback approached. They suggest that people may lower estimates to avoid disappointment (rather than for other reasons such as to regulate, explain pre-feedback anxiety or to escape anxiety).

Although individual differences and personality variables hold promise for explaining variance in self evaluations of performance, it is unlikely that they can be understood adequately without reference to the effects of the temporal factors we have shown to have such consistent effects. The challenge will be to extend theory to understand how individual variables will or will not combine with time horizon and temporal perspective, and experience to alter the patterns observed in the study.

## **ENDNOTES**

- [1] We conducted similar analyses on measures of MAD. We only report the results for mean bias in detail to simplify our presentation and explanation of our results. However, when the results for mean bias differ from those for MAD which were few, we report these and provide possible explanations for such differences.
- [2] For Quiz 1, the estimate termed "1 Quiz Before" was collected during base-line data collection, that is, before the set of all five Quizzes.
- [3] This design was thusly constructed because we had a larger range of pre- feedback estimates for Quizzes 3, 4, and 5. Participants generated estimates on all five Quizzes at each time-point of evaluation. We only analyzed pre-feedback estimates because it was not possible to elicit an equal number of post-quiz and pre-feedback evaluations for these three quizzes.
- [4] Due to experimenter error, "2 Days after Quiz 3" performance assessments were not obtained. Therefore, separate analyses were conducted using the four other quizzes (1,2,4,5) to test for Temporal Perspective effects.
- [5] We conducted similar analyses on measures of MAD. We only report the results for mean bias in detail to simplify our presentation and explanation of our results. However, when the results for mean bias differ from those for MAD--which were few, we report these and provide possible explanations for such differences.

- [6] For Quiz 1, the estimate termed "1 Quiz Before" was collected during base-line data collection, that is, before the set of all five Quizzes.
- [7] Note that comparable long retrospective time horizons were not feasible in this study because they would have required delaying performance feedback to students for several weeks.
- [8] This design was thusly constructed because we had a larger range of prefeedback estimates for Quizzes 3, 4, and 5. Participants generated estimates on all five Quizzes at each time-point of evaluation. We analyzed only pre-feedback estimates because it was not possible to elicit an equal number of post-quiz and pre-feedback evaluations for these three quizzes.
- [9] Due to experimenter error, "2 Days after Quiz 3" performance assessments were not obtained. Therefore, separate analyses were conducted using the four other quizzes (1,2,4,5) to test for Temporal Perspective effects.
- [10] We found that the patterns for MAD were similar to those depicted by measures of bias.
- [11] These were the only 3 pre-feedback estimates we had for all 5 quizzes.
- [12] We did not compute the validity coefficients for social performance judgments - the criteria for these judgments do not change across persons.
- [13] We did not find similar patterns for changes in MAD they were non-significant across variations in Time Horizon, Temporal Perspective and Experience.
- [14] Number of comparisons for raw score estimates and percentile rank estimates do not always match due to ties in one or more comparisons of the correlations.

## REFERENCES

- Alicke, M. D., LoSchiavo, F. M., Zerbst, J. and Zhang, S. (1997). Journal of Personality and Social Psychology, 73 (4), 781-789.
- Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.
- Beck, H. P., and Seta, J. J. (1980). The effects of frequency of feedback o a simple coaction task. *Journal of Personality and Social Psychology*, 38. 75-80.
- Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology*, 59(5), 960-970.
- Biernat, M., Manis, M., Kobrynowicz, D. (1997). Simultaneous assimilation and contrast effects in judgments of self and others. *Journal of Personality and Social Psychology*, 73(2), 254-269.
- Brickman, P., and Berman, J. J. (1971). Effects of performance expectancy and outcome certainty on interest in social comparison. *Journal of Experimental Social Psychology*. 7. 600-609.
- Carver; and Scheier, (1999). Self-Control Theory. In R.S.Wyer (Ed.) Themes and issues in the selfregulation of behavior: *Advances in social cognition*, vol. 12. Mahwah, NJ: Erlbaum.
- Dweck, C. S., and Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.

- Feather, N. T. (1969). Attribution of responsibility and valence of success and failure in relation to initial confidence and task performance. *Journal of Personality and Social Psychology*, 13, 129-144.
- Festinger, L. (1954). A theory of social comparison processes. Human Relations, 7, 117-140.
- Gilovich, T., Kerr, M. and Medvec, V. (1993). Effect of temporal perspective on subjective confidence. *Journal of Personality and Social Psychology*, 64, 552-560.
- Harackiewicz, J. M., Barron, K. E., Carter, S. M., Lehtc, A. T., and Elliot, A. J. (1997). Predictors and consequences of achievement goals in the college classroom: Maintaining interest and making the grade. *Journal of Personality and Social Psychology*, 73(6), 1284-1295.
- Helgeson, V. S., and Mickelson, K. D. (1995). Motives for social comparison. Personality and Social Psychology Bulletin, Vol. 21, No. 11, Nova 1200-1209. Higgins, Strauman, and Klein, 1986. Standards and the process of selfevaluation: Multiple affects from multiple stages. In *Handbook of Motivation and Cognition: Foundations of Social Behavior*, ed. R.M. Sorrentino, E. T. Higgins, 1, 350378. New York: Guilford.
- Jarvis, W. B. G., and Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology*, 70(1), 172-194.
- Klar, Y. and Giladi, E.E. (1997). No one in my group can be below the group's average: A robust positivity bias in favor of anonymous peers. *Journal of Personality and Social Psychology*, 73(5), 885-901.
- Lyubomirsky, S., and Ross, L. (1997). Hedonic consequences of social comparison: A contrast of happy and unhappy people. *Journal of Personality and Social Psychology*, 73(6), 1141-1157.
- Markus, H. and Zajonc, R. B. (1985). The cognitive perspective in social psychology. In G. Lindzey and E. Aronson (Eds.), *Handbook of Social Psychology* (Volt., pp. 137-230). Hillsdale, NJ: Erlbaum.
- Nisan, M. (1972). Dimension of time in relation to choice behavior and achievement orientation. *Journal of Personality and Social Psychology*, 21. 175-182.
- Prentice, D. A. (1990). Familiarity and differences in self- and otherrepresentations. *Journal of Personality and Social Psychology*, 59(3), 369-383.
- Pyszczynski, T., Greenberg, J., and LaPrelle, J. (1985). Social comparison after success and failure: Biased search for information consistent with a self-serving conclusion. *Journal* of Experimental Social Psychology, 21, 195-211.
- Radhakrishnan, P., Arrow, H., and Sniezek, J. A. (1996) Hoping, performing, learning, and predicting: Changes in the accuracy of self-evaluations of performance. *Human Performance*, 9(1), 23-49.
- Roney, C. J. R., and Sorrentino, R. M. (1995) Self-evaluation motives and uncertainty orientation: Asking the "Who" question. *Personality and Social Psychology Bulletin*, Vol. 21 (12). 1319-1329.
- Sanna, L., and Pusecker, P. A. (1994). Self-efficacy, valence of self-evaluation, and performance. *Personality and Social Psychology Bulletin*. Vol. 20 (1), 82-92. Sansone, C. (1986). A question of competence: The effects of competence and task feedback on intrinsic interest. *Journal of Personality and Social Psychology*, 51, 918-931.
- Scheier, M. F. and Carver, C. S. (1983). Self-directed attention and the comparison of self with standards. *Journal of Experimental Social Psychology*, 19, 205-222.

- Sedikides, C. and Strube, M. J. (1997). Self-evaluation: To thine own self be good, to tine own self be sure, to thine own self be true, and to thine own self be better. *Advances in Experimental Social Psychology*, Vol.29, 209-269.
- Seta, J. J. (1982). The impact of comparison processes on coactors' task performance. *Journal of Personality and Social Psychology*, 42(2), 281-291.
- Shepperd, J. A., Ouellette, J. A., and Fernandez, J. K. (1996). Abandoning unrealistic optimism: Performance estimates and the temporal proximity of self-relevant feedback. *Journal of Personality and Social Psychology*. 70. 844-855.
- Sniezek, J. A. and Buckley, T. (1991). Confidence depends on level of aggregation. *Journal of Behavioral Decision Making*. Vol. 4, 263-272.
- Sniezek, J.A. and Reeves, A.P. (1986). Feature cues in probability learning: Database information and judgment. Organizational Behavior and Human Decision Processes, 37, 297-315.
- Suls, J. M., and Tesch, F. (1978). Students' preferences for information about their test performance: A social comparison study, *Journal of Applied Social Psychology*, 8, 189-197.
- Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90, 245-271.
- Wood, J. V. (1989). Theory and research concerning social comparisons and personal attributes. *Psychological Bulletin*. 106, 231-248.

Chapter 48

# **THE EFFECTS OF PERCEPTIONS ON MODE CHOICE**\*

# Maria Vredin Johansson

Department of Economics, Uppsala University, Box 513, SE-751 20 Uppsala, Sweden

# ABSTRACT

While remaining within the traditional micro-economic framework of rational utility maximization, we enrich the standard and random parameters logit choice models with perceptions data. From the estimated models we derive a value of time and we also make a tentative attempt to derive a value of safety. Because we estimate the values simultaneously, we are able to explore whether values estimated in conjunction differ from values estimated in isolation. Survey data is used to measure the individual's perceptions of five modal attributes (time, cost, safety/risk, environmental friendliness and flexibility) and show how these perceptions affect the modal choice for work trips. The respondents' perceptions are elicited by a novel approach in which the names of two modes (car and bus) are used as attribute levels instead of objective levels. A difference between our survey and traditional ones is that we do not attempt to educate the respondents about, for example, the risks of travelling. Instead, we *record* the respondent's perceptions about the risk and the other the modal attributes.

Keywords: Value of Time, Value of Statistical Life, Random Parameters Logit, Stated Preferences

# **1. INTRODUCTION**

In the empirical literature on travel mode choice, most choice models use objective modal attributes and individual characteristics as explanatory variables (cf. Algers et al., 1995).

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Economics, Business and Finance*, edited by Klaus P. Hofmann published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

Nonetheless, few people reject the thought of travel mode choice also depending on harderto-measure, qualitative variables, such as comfort and convenience. It has even been suggested that the choice of travel mode can depend on the perceptions<sup>1</sup> and images of the alternatives.

For instance, Koppelman and Pas (1980) argue that, as models with only demographic and objective attribute levels only provide a limited understanding of the behavioural processes underlying choice, incorporation of perceptions and feelings would improve the understanding of modal choices. Interestingly, Adamowicz et al. (1997) have in a study in the field of environmental economics found that a model based on perceptions slightly outperformed (i.e. had greater explanatory power) a model based on objective attribute measures.

Moreover, as it is recognised that objective information about a travel mode can be quite different from the perception of the information (Lichtenstein et al., 1978; Lee, 1981; Slovic et al.,1981; Mowen, 1990, ch. 2), the idea that objectively measured attributes are the only influential variables in travel choice models is open to question. If people respond to what they perceive, i.e. make choices on basis of their perceptions, the perceptions about the modal attributes are the relevant response variables to which people react, not the objective attribute levels. Thus, reconsideration when it comes to explaining travel choices seems both warranted and welcome.

In this paper, we use a mail survey to measure the individual's perceptions of five modal attributes (time, cost, safety, environmental friendliness and flexibility) and see how these perceptions affect modal choice for work trips. A difference between our survey and traditional ones, is that we do not attempt to inform, or educate, people about, for instance, the costs or risks of travelling. Instead, we *record* the respondents' perceptions about the modal attributes.

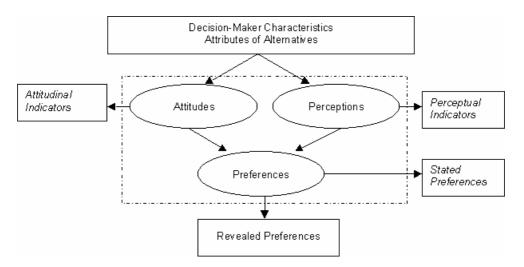


Figure 1. The consumer decision process (adapted from Morikawa et al., 1990, p. 3).

<sup>&</sup>lt;sup>1</sup> In this paper, perception is defined as an individual's subjectively formed opinions and ideas about a stimulus or a piece of information.

The work in this paper is based on the paradigm of individual choice behaviour depicted in Figure 1. In the figure latent variables are depicted by ovals and observable variables by boxes. Decision maker characteristics and objective attributes of alternatives are assumed to affect the *attitudes, perceptions* and *preferences* of the decision-maker.<sup>2</sup> These latent variables - all assumed to affect choice - constitute the "black box" of the consumer. Because perceptions and attitudes are assumed to affect the individual's preferences and, as the preferences are assumed to determine choice, it is possible that incorporation of perceptions in travel choice models will improve the models' explanatory power.

While staying within the traditional microeconomic framework of rational utility maximisers (where preferences are expressed by utility functions), we enrich the standard and random parameters logit choice models by perceptions data. From the estimated models we derive values of time. We also make a tentative attempt to measure the monetary value of safety. Because the Swedish National Road Administration (SNRA) uses both a value of time savings and a value of safety (i.e. value of "statistical life") in their project evaluations, it is desirable to estimate both values in the same model. Presently the SNRA uses values derived from separate studies and, as far as we know, simultaneous estimation has never been performed in Sweden before. Because we estimate the values simultaneously, we are able to explore whether values estimated in conjunction differ from values estimated in isolation.

This paper has three objectives: (*i*) to elicit the individuals' perceptions about the modal attributes time, cost, safety, environmental friendliness and flexibility; (*ii*) to use these perceptions to explain mode choice for work trips, and; (*iii*) to derive monetary values for time and safety in the same model estimation.

The remainder of the paper is organised as follows; the next section describes the data collection and the experimental design, section three gives a theoretical framework for stated preferences discrete choice models, section four presents the results from the empirical estimations and, finally, section five concludes and discusses the results.

## **2. THE SURVEY**

In a mail survey, we focus on modal choice for work trips. There were several reasons for this limitation. First, work trips are easy to define for most people. Second, work trips are repeated journeys for which the nodes does not change unless the respondent has varying places of work. Third, work trips are frequently performed and people are, therefore, often well-informed about the alternative modes of travelling. Consequently the individuals' preferences for different travel modes can be assumed to be fairly well known. Finally, according to Statistics Sweden (SCB, 1998) almost half of the trips in 1997 consisted of trips to and from work/school, a fact that makes work trips imperative to study.

In total, there were 22 questions in the questionnaire. Apart from socio-economic questions and questions regarding the respondent's habitual mode of travelling to and from work (revealed preferences (RP) information), the questionnaire contained a question to elicit the respondent's perceptions of modal attributes of two different travel modes (car and bus) and two stated preferences (SP) choice questions where the respondent was asked to choose

<sup>&</sup>lt;sup>2</sup> Attitudes are defined as the decision-maker's subjective importance of the attributes and *preferences* as the desirability of alternatives.

the preferred of two hypothetical travel modes. Each travel mode was described by four *attributes*.

In December 1997, the questionnaire was mailed to 480 residents of Borlänge and Falu municipalities randomly selected<sup>3</sup> from the Swedish Official Register of Persons and Addresses (SPAR).<sup>4</sup> In the beginning of January 1998, a follow-up reminder was mailed to the non-respondents at that time. The overall response rate from the two mailings was 66.8 percent.

## 2.1. The Attributes

In the SP choice questions the hypothetical alternatives were described by five different modal attributes: time, cost, safety, environmental friendliness, and flexibility (see questions 20-21 in Appendix A).

Time and cost were included for obvious reasons. Safety was included for two reasons: first, there is evidence (Noland, 1995) that decreases in the perceptions of a travel mode's safety reduces the probability of that mode being chosen and, second, safety improvements are the second most important benefit (time savings are the most important) in the costbenefit analyses of road investment projects made by the SNRA and, consequently, of considerable policy interest. The mode's environmental friendliness was added because we wished to analyse if the on-going debate on global warming has affected people's choice of travel mode.

Because we judged it unnecessary, and potentially too burdensome, for the respondents to evaluate five different attributes, we split the total sample in two so that each sub-sample consisted of 240 individuals, and let each sub-sample evaluate only four different attributes. Travel time, travel cost and safety were used in both sub-samples, while environmental friendliness and flexibility were used interchangeably. In this way we obtained an "environmental friendliness" sub-sample (I) and a "flexibility" sub-sample (II). Thus, the distinguishing feature between the two sub-samples is their fourth attribute (see Table 1).

Sub-sample I	Sub-sample II		
Travel time ( <i>t</i> )	Travel time ( <i>t</i> )		
Travel cost $(c)$	Travel cost ( <i>c</i> )		
Safety (s)	Safety (s)		
Environmental friendliness (e)	Flexibility (f)		
n <sub>1</sub> =240	n <sub>11</sub> =240		

## Table 1. Attributes of the two sub-samples

<sup>&</sup>lt;sup>3</sup> Because we wanted respondents who, with a larger probability compared to a completely random sample, were working, the respondents were required to be aged between 36 and 55 years.

<sup>&</sup>lt;sup>4</sup> Borlänge and Falu municipalities are situated approximately 200 kilometres north-east of Stockholm in the Dalecarlia region.

#### 2.2. Attribute Levels

All the attributes vary in two levels, a "high" and a "low" level. For the factorial design, we construct a new, artificial, attribute, *ef*, which symbolises the environmental friendliness (*e*) attribute in sub-sample I and the flexibility (*f*) attribute in sub-sample II. Having four, instead of five variables, reduces the complexity of the factorial design because the number of alternatives is reduced from  $2^5$  to  $2^4$ 

The intrinsic natures of the attributes in this survey differ because the travel times and travel costs are quantitative variables for which objective (i.e. researcher defined) values are common, while safety, environmental friendliness and flexibility are, more or less, qualitative variables where objective values are rather uncommon. For these qualitative variables other measurement scales than objective must be used. For instance, a measure of the travel mode's objective risk level can be defined as the actual risk the individual faces during the specific trip expressed in historical number of annual fatalities. For the mode's environmental friendliness an objective measure would be the average emissions per passenger kilometre (Lenner, 1993). For the mode's flexibility it is harder to conceive of a good objective measure. Possible solutions are to use some kind of categorical measure, such as low, medium and high, or, to endogenise the measurement by defining the prevailing level as the one from which changes are made. If we were to use objective values on all attributes in the SP survey, the values of the qualitative attributes would have to be invented and communicated to the respondents.

We, however, circumvent the problem of constructing objective values for the qualitative variables in this survey by using the respondent's perceptions of these variables. As a matter of fact, in order to simplify and make the questionnaire consistent, we use the respondent's perceptions of *all* variables - even for those with objective values.<sup>5</sup> To elicit the respondents' perceptions of modal attributes, we propose a new approach, implemented by the four-step procedure described below. To the best of our knowledge, this approach has no previous application.

The first step is to use the *names* of two travel modes, car and bus, as levels for *all* attributes. Thus, an alternative could be described as having travel time "like car", travel cost "like bus", safety "like bus" and environmental friendliness or flexibility "like car". Consequently, one of the alternatives describes a travel mode with all attributes "like car", while another has all attributes "like bus". All other alternatives are mixtures of the "like car" and "like bus" attributes and, therefore, hypothetical constructs. To the best of our knowledge, this "like"-approach has only been used once before in an SP survey, by Louviere and Johnson (1991) who, in a marketing study, explored the retail images of different supermarkets.<sup>6</sup>

The second step is to decide whether car (bus) represents the high or the low level of the different attributes. Five assumptions about the attribute levels were used in the experimental design (see Table 2): (i) Going a specific distance by car is normally faster than going the

<sup>&</sup>lt;sup>5</sup> An advantage of using perceptions is that we do not have to concentrate on a specific type of traveller for whom we know the modal attribute levels beforehand (it would be impossible to construct realistic alternatives unless we had information about the respondent's actual travel mode). By using perceptions we are able to survey respondents with a mail questionnaire.

<sup>&</sup>lt;sup>6</sup> Louviere and Johnson (1991) used attribute levels in the "like" form (e.g. "like K-Mart") and measured the perceptions of the different supermarkets through ratings.

same distance by bus, due to the stops the bus makes - consequently, the high level of travel time was designated bus and the low level car; (*ii*) The travel cost of car was considered larger than the equivalent for bus, considering the additional costs of buying and running a car, rendering the car the high and the bus the low level; (*iii*) Based on historical statistics of traffic fatalities (SIKA, 1998; Kommunikationsdepartementet, 1997), bus safety was regarded superior to car safety and, therefore, safety on bus is high, while safety in car is low; (*iv*) Because the discharges per passenger kilometre from buses are lower than those from cars (Lenner, 1993) the environmental friendliness of bus is assumed to be higher than that of car; and (*v*) When assigning a high and a low level to the (admittedly ambiguous) notion of flexibility, we considered three dimensions of mode flexibility: departure times, walking distances and freedom in choice of route. Because cars, in addition to having a greater freedom in route choice, generally, has more flexible "departure times", and can often be parked in proximity of the travel destination, we regarded car as the more flexible travel mode (i.e. high level of flexibility).

The individual's preferences for the attributes are assumed to be monotone (increasing or decreasing). If the respondents perceive the attributes the way we do, lower travel cost and travel time are always preferred, as are higher safety, environmental friendliness and flexibility. Of course, the respondents may perceive the attributes differently, or make different assumptions about the attribute levels.<sup>7</sup> This should not be a weakness of the design, because the respondents make the choices based on their own preferences, not on *our* ideas about their perceptions.

Attribute	High level (+)	Low level (-)		
Travel time ( <i>t</i> )	Bus	Car		
Travel cost ( <i>c</i> )	Car	Bus		
Safety (s)	Bus	Car		
Environmental friend. (e)	Bus	Car		
Flexibility ( <i>f</i> )	Car	Bus		

## Table 2. Attribute levels

The third step is measurement. Most desirable would be to measure modal perceptions on a quotient scale, i.e. an equidistantly graded scale with an absolute zero, so that comparisons could be made both within and between respondents.<sup>8</sup> To obtain quotient scaled data, we apply a method originating in measurement scales for comparative judgement in psychophysics (cf. Björkman and Ekman, 1957). In this method the respondents make pairwise comparisons between a stimulus and a standard. The respondent's perception of the

<sup>&</sup>lt;sup>7</sup> Gårder et al. (1994), for example, show that experts may judge on-street bicycle lanes as less safe (compared with no lanes at all) whereas the cyclists themselves perceive them as being safer.

<sup>&</sup>lt;sup>8</sup> Ranking yields data with individual-specific scale units and origins (ordinal data), whereas rating yields equidistantly measured data with individual-specific origins (interval data). While ordinal data only describe the respondent's preferences (e.g.  $X_1 < X_2 < X_3 < X_4$ ), interval data also describe the magnitude of the difference in preferences (e.g.  $X_4 - X_2 = 2(X_2 - X_1)$ ). Only data measured on a quotient scale can express the *intensity* of preferences (e.g.  $X_2$  is twice as good as  $X_1$ ,  $X_2 = 2 X_1$ ) (Björkman & Ekman, 1957).

stimulus is frequently expressed as *a multiple* or *a fraction* of the standard, so that, for instance, the stimulus is either twice or half of the standard.

In our survey we let the respondents compare bus (the stimulus) to car (the standard)<sup>9</sup> and express any difference perceived in percent. That is, while car is assumed to be "100%" on each attribute, the respondents are asked to state how, compared to the car, they perceive the attributes of the bus (see question 19, Appendix A). For instance, if a respondent perceives the travel time of bus to be three times that of car, she responds "300%" and, if she perceives the cost of bus to be a third of that of car, she responds "33%". Consequently, the perceptions obtained are either multiples or fractions of the car standard.

The fourth step is to infer real values, i.e. values in minutes and Swedish kronor (SEK)<sup>10</sup>, from the percentages, using the RP information provided by the respondent. For example, if the respondent travels to work by car, information about the actual time and cost by car can, in conjunction with the perceptions of the time and cost by bus, be used to give the "like car" and the "like bus" levels real values. A simple example: assume that the respondent travels to work by car and that it takes 10 minutes and costs SEK 20. The same respondent perceives the travel time of bus to be twice that of car and responds "200%" to the relevant part of question 19. Similarly, the travel cost of bus is perceived to be half that of car and, therefore, she responds "50%" to that part of the question. Using the RP information, i.e.  $t_{auto} = 10$  and  $c_{auto} = 20$ , we can calculate the travel time of bus ( $t_{bus}$ ) to 20 minutes and travel time "like bus" equals 20 minutes in the SP choice questions. Analogously, real values can be inferred for the cost attribute.

Equivalent transformations for the time and cost variables can easily be performed for respondents who ride a bus to work and, in fact, for all respondents who has either bus or car in their choice sets. However, both the actual mode's time *and* cost must be known. If the respondent is unaware of an alternative's cost or time, we judge the alternative unlikely for the respondent and do, therefore, not define it as a part of the respondent's choice set.

For the qualitative variables, safety, environmental friendliness and flexibility, no such transformations can be made since we have no RP values for these variables. Consequently, both real values and "pure" attribute perceptions are used to explain the SP choices in the estimation. The main advantage of inferring real values from perceptions is that we are able to calculate a value of time that is easy to interpret and compare with previous research findings.

<sup>11</sup> General formulas for the calculations are:  $t_{bus} = t_{auto} \cdot \tau / 100$  and  $c_{bus} = c_{auto} \cdot \psi / 100$  where  $t_{bus}$  is travel time by bus,  $t_{car}$  travel time by car,  $c_{bus}$  travel cost by bus and  $c_{car}$  travel cost by car.  $\tau$  and  $\psi$  are the perceptions of travel time and travel cost of bus compared to car expressed in percent.

<sup>&</sup>lt;sup>9</sup> The choice of car as the standard is merely one of convenience, but because car is the most frequently used travel mode for work trips (SCB, 1998), we believe that most people find it easier to compare bus to car than the other way around.

<sup>&</sup>lt;sup>10</sup> US\$ 1 was, in 1997, on average equal to SEK 7.64. In 2006, US\$ 1 was on average equal to SEK 7.38. Although there has been large variations in the average exchange rates over the years, the rate in 1997 was on level with the rate in 2006 (www.riksbank.se).

## 2.3. Experimental Design

Because the survey was performed as a postal questionnaire, it was important to keep the questionnaire short and simple to obtain a high response rate. We, therefore, restricted the size of the choice experiment, which is a function of the number of attributes to be varied. With n=16 we could construct as many as n(n-1)/2=120 different paired comparisons (full factorial).<sup>12</sup> As no respondent arguably could be required to evaluate 120 different pairs, we had to reduce the number of comparisons. By selecting the smallest orthogonal main effects plan from the full factorial through "blocking" (see Appendix B), we were able to estimate non-confounded main (i.e. attribute specific) effects. In our survey, the smallest orthogonal main effects design consists of eight pairs. Because we believed that no respondent could be required to perform more than two paired comparisons, i.e. two different choice questions, the eight pairs were divided into four sets of two pairs each. This particular experimental design ensures non-confounded main attribute effects, but confounded (with the blocks) interaction effects. According to Batsell and Louviere (1991), interaction effects are rarely estimated in choice models and, unless the respondents in different blocks differ or behave differently, there will be no block effects.

With two paired comparisons in each questionnaire and a sub-sample size of 240, we got four "sub-sub-samples", A-D and E-H, respectively, consisting of 60 individuals each (see Table 3). Each individual was randomly assigned a treatment.

Because we kept the dominating and the dominated alternatives in the experiment, and as the pairs in the paired comparisons are *fold-over* pairs (which means that they are complementary as the signs of the attribute levels of the first alternative are reversed in the second, see Box et al. (1978) and Appendix B), we have one questionnaire in each sub-sample that contains the choice between the "best" and the "worst" alternatives.<sup>13</sup> In sub-sample I, questionnaire B contains the best-worst pair and in sub-sample II, it is questionnaire F. If these *designed* best-worst questionnaires. However, as can be seen in Table 3, this is obviously not the case since B, in fact, has the lowest response rate in sub-sample I and F has the second lowest response rate in sub-sample II.

	Sub-sample I				Sub-sample II			
	А	В	С	D	Е	F	G	Н
Main sample	60	60	60	60	60	60	60	60
Responses	37	37	44	39	43	38	37	43
Response rate (%)	62.7	61.7	73.3	65.0	71.7	64.4	62.7	72.9

Table 3. Response rates of the different sub-sub-samples

Note: The real sample is the main sample less the questionnaires that were.

<sup>&</sup>lt;sup>12</sup> We do not compare an alternative with itself or consider the sequencing of alternatives.

<sup>&</sup>lt;sup>13</sup> When used as a consistency check, approximately 22 percent in the environmental friendliness sub-sample and about 3 percent in the flexibility sub-sample chose the worst alternative in the best-worst question. However, what is right and what is wrong is based on the individual's perceptions of the attributes. Therefore, we can not conclude that these respondents behaved inconsistently.

# **3. MODELLING CHOICES**

The behavioural assumption underlying the model for travel choice used in this paper is maximisation of individual utility, i.e. the individual who, with certainty, knows the utility associated with every travel alternative, chooses the mode that maximises his/her utility. A prerequisite for this assumption is the existence of stable and well known preferences. Although the utility model has limitations when it comes to describing the human decision making process as it focuses on the final choice (Kahneman and Tversky, 1979), it makes empirical estimation possible. It, therefore, plays a major role in travel choice analysis as an approximation to real decision making.

In traditional microeconomic utility maximisation, the individual's utility is maximised with respect to a bundle of continuous goods, G, subject to a budget constraint. In a 1966 paper, Lancaster recognised that the primary source of utility is the qualitative attributes (e.g. comfort, flavour, softness) of the goods, and not the goods *per se*. Consequently, the individual's choice problem can be framed as a choice between bundles of attributes to achieve maximum utility.

There are several ways of formulating discrete mode choice models (cf. Becker, 1965; DeSerpa, 1971; Train and McFadden, 1978; Jara-Diaz and Videla, 1989). Our model is based on the fairly general disaggregate choice model by Jara-Diaz and Videla (1989).<sup>14</sup> In this model, time is treated as one attribute among others assumed to affect modal choice. There is no explicit time constraint and, in its simplest form (which we will employ), income does not affect choice.

A representative individual n (suppressing indexation) is assumed to choose both a discrete good (a mode) and continuous goods to maximise his/her utility:

$$\max_{\mathbf{G},i} U = U(\mathbf{G}, \mathbf{Q}_i) \tag{1}$$

subject to

$$Y=\mathbf{1}\mathbf{G}+c_{i}i\in M$$
(2)

where 1 is a  $K \times 1$  column vector of ones,  $G=[g_1, g_2, ..., g_K]$  is a  $K \times 1$  column vector of consumed continuous goods,  $c_i$  is the travel cost of mode *i*, *Y* is income,  $Q_i=[q_{i1}, q_{i2}, ..., q_{iR}]$  is

a  $R \times 1$  column vector of "quality" attributes associated with travel mode *i* (excluding cost, but including time) and *M* is equal to the choice set of travel modes. The utility function is assumed to be twice differentiable, quasi-concave and increasing in G. Equation (2) is the budget constraint, in which the travel cost and income are normalised by the price of G. Conditional on the choice of travel mode *i*, we can derive conditional demands for  $g_{ik} = g_{ik}(Y - c_i, Q_i)$  Associated with the conditional demands is a conditional indirect utility function, which defines the maximum attainable utility level, conditional on the choice of travel mode *i*,  $V_i = V_i(Y - c_i, Q_i)$ .

<sup>&</sup>lt;sup>14</sup> See also Jara-Diaz, 1998.

The travel mode chosen will be the alternative that renders the highest conditional indirect utility, i.e. mode *i* is chosen if  $V_i \ge V_j$ ;  $\forall j \ne i$ . From the conditional indirect utility function it follows that the marginal utility of income ( $\mu$ ) is:

$$\mu = \partial V_i / \partial Y = - \partial V_i / \partial c_i.$$

The value of characteristic r is equal to the marginal rate of substitution between that characteristic and the (negative of the) marginal utility of income:

$$VOA_{ir} = \frac{\partial V_i / \partial q_{ir}}{\partial V_i / \partial c_i}.$$

If attribute t for mode i equals time, the value of time  $(VOT_i)$  is given by:

$$VOT_i = \frac{\partial V_i / \partial q_{it}}{\partial V_i / \partial c_i}$$

We need to estimate a generic value of time, because our SP data do not consist of specific modes.

We assume that both the attribute vector associated with travel mode i ( $\mathbf{Q}_i$ ) and the conditional indirect utility function are linear in the arguments. Moreover, we add an error term,  $\varepsilon_i$ , to account for measurement errors. Thus, the conditional indirect utility function equals:

$$V_i = \alpha_i + \mu (Y - c_i) + \gamma' \mathbf{Q}_i + \varepsilon_i.$$
<sup>(3)</sup>

When comparing  $V_i$  and  $V_j$ , only variables that vary between modes affect choice. Therefore, the relevant part of the conditional indirect function is,  $\overline{V_i}$ , a truncated conditional indirect utility function, where income is not included

$$V_i = \alpha_i + \mu c_i + \gamma' \mathbf{Q}_i + \varepsilon_i; \ \forall i \in M.$$

In the estimation  $\gamma = [\gamma_t, \gamma_s, \gamma_{ef}]'$ . If the respondent perceives the attributes the way we do,  $\partial V_i / \partial c_i$  is expected to be negative (so that the negative of this parameter equals the marginal utility of income) because there is disutility from travel cost. By analogous reason, the time parameter,  $\gamma_i$ , is expected to be negative. Consequently, the *VOT* is the ratio of two negative numbers and, therefore, expected to be positive. All other parameters in the vector  $\gamma$  are expected to be positive because utility is assumed to increase from increases in the modal

safety, environmental friendliness and flexibility. The marginal rates of substitution between these attribute parameters and the cost parameter are therefore all expected to be negative.

Thus, the model used in the estimation is a linear, additive in the parameters, model. Linear additive models are *compensatory*, which means that the individual can trade-off a low value on one attribute for a high value on another and achieve the same utility. Empirically, compensatory choice models are found to work well in simple choice tasks such as in the choice between two alternatives (Payne, 1976).

#### 4. ESTIMATION

#### 4.1. Standard Logit

In the SP choice experiment respondent n's choice set (M) consists of two alternatives, A and B. The respondent is asked to indicate which of the two alternatives that (s)he prefers, i.e. which alternative is "chosen".

The probability that the individual (suppressing individual indexation) will choose alternative B over A is given by:

$$\Pr(B) = \Pr(U_B > U_A) = \Pr(V_B + \varepsilon_B > V_A + \varepsilon_A) = \Pr(\Delta \varepsilon > -\Delta V),$$

where  $\Delta \varepsilon = \varepsilon_B - \varepsilon_A$  and  $\Delta V = V_B - V_A$ . The probability that *A* will be chosen over *B* is logically Pr(*A*)=1-Pr(*B*). The dependent variable,  $y_n \in \{0,1\}$  is an indicator variable taking unit value if alternative *B* is chosen and zero if alternative *A* is chosen. Different assumptions about the distribution of the random terms  $\varepsilon_A$  and  $\varepsilon_B$  or about their difference,  $\Delta \varepsilon$ , lead to different choice models. Assuming that  $\Delta \varepsilon$  is normally distributed ( $\varepsilon_A$  and  $\varepsilon_B$  are normally distributed) results in the probit model and assuming that  $\Delta \varepsilon$  is logistically distributed, i.e. that  $\varepsilon_A$  and  $\varepsilon_B$  are independently and identically distributed (IID) extreme value type I distributed, results in the logit model. Unless there are compelling reasons for assuming one distribution or the other, the choice between logit and probit is, in binary cases, only a matter of preference. Here, we assume  $\Delta \varepsilon$  to be logistically distributed. The log-likelihood function is:

$$\ln \ell = \sum_{n=1}^{N} \left[ (1 - y_n) \ln \frac{1}{1 + \exp(\Delta V)} + y_n \ln \frac{1}{1 + \exp(-\Delta V)} \right].$$

We estimate the model:

$$\Delta V = \alpha_B - \alpha_A + \mu (c_B - c_A) + \gamma' (\mathbf{Q}_B - \mathbf{Q}_A) + \Delta \varepsilon.$$

Pooling the environmental friendliness and the flexibility sub-samples, and estimating the combined model with the use of dummy variables for the flexibility sub-sample result in

rejection of the null hypothesis of equal sub-samples.<sup>15</sup> Therefore, the two sub-samples are treated separately.

The results for the environmental friendliness sub-sample are given in Table 4.<sup>16</sup> Table 4 shows that the time and cost variables are significant at the five percent level with the expected signs (i.e. there is disutility from spending time and money on work trips). The safety and the environmental friendliness parameters have the expected signs, but are insignificant. The standard deviation of the environmental friendliness parameter is very low, indicating a good precision of the estimate. Nonetheless, the parameter is insignificant, meaning that considerations about the mode's environmental impacts are unimportant when making modal decisions for work trips. The number of correctly predicted choices in this model is 116.

Attribute	Parameter	St dev	t-value
Constant	0.063	0.214	0.293
TIME <sub>I</sub>	-0.037	0.008	-4.752
COSTI	-0.028	0.010	-2.804
SAFETYI	0.000	0.001	0.289
ENV. FRIENDL	0.000	0.000	1.037
n	77		
Choices made	143		
LRI	0.277		
ln l	-71.50		

#### Table 4. Environmental friendliness sub-sample (I): Results from standard logit estimations of stated choices

Note: The likelihood ratio index  $(LRI) = 1 - (\ln \ell / \ln \ell_0)$ .

# Table 5. Flexibility sub-sample (II): Results fromstandard logit estimations of stated choices

Attribute	Parameter	St dev	t-value
Constant	0.989	0.219	4.506
TIMEII	-0.047	0.010	-4.587
COST II	-0.069	0.018	-3.891
SAFETY II	0.004	0.002	1.899
FLEXIBILITY	0.001	0.001	0.873
п	85		
Choices made	148		
LRI	0.232		
ln l	-71.07		

Note: The likelihood ratio index (*LRI*)=1-( $\ln \ell / \ln \ell_0$ ).

<sup>&</sup>lt;sup>15</sup> LR-test statistic: 11.02 with c<sup>2</sup> critical value: 9.49.

<sup>&</sup>lt;sup>16</sup> The standard logit estimations were conducted in LIMDEP for both sub-samples.

Table 5 shows the results from the flexibility sub-sample. The time and cost parameters are, again, significant at the five percent level with the expected signs. The safety parameter is significant at the six percent level with the expected sign. The flexibility parameter is not significant, but has the expected sign. The relatively large variance of this parameter is likely to be a consequence of the lack of specificity in the definition of this attribute. Thus, the flexibility attribute seems to have been perceived differently by the respondents. The number of correctly predicted choices is in this model 114.

Table 6 gives the values of time for the two sub-samples, calculated as:

$$VOT = \frac{\partial V / \partial TIME}{\partial V / \partial COST} = -\frac{\widehat{\gamma_t}}{\widehat{\mu}},$$

and the value of safety for the flexibility sub-sample, calculated as:

$$VOT = \frac{\partial V / \partial SAFETY_{II}}{\partial V / \partial COST_{II}} = -\frac{\widehat{\gamma_{s,II}}}{\widehat{\mu}}$$

The standard errors are calculated using the Delta method (cf. Greene, ch. 10, 1993).

Table 6. SP values of time and safety (1997 prices), standard deviations (St dev) and 95percent confidence intervals (CI)

	Value	St dev	CI (95%)
VOT <sub>I</sub>	79.50 SEK/hour	29.29	22.09 - 136.91
<i>VOT</i> <sub>II</sub>	40.80 SEK/hour	8.85	23.45 - 58.15
VOA <sub>s, II</sub>	0.07 SEK/percent	0.03	0.01 - 0.13

We refrain from calculating values of environmental friendliness and flexibility because these variables were all insignificant. On the same grounds, we refrain from calculating the value of safety in the environmental friendliness sub-sample.

The estimated values of time and safety are significant. The value of time in the environmental friendliness sub-sample (I) is almost twice the value time in the flexibility sub-sample (II). The difference is not significant, but the  $VOT_{II}$  has better precision than the  $VOT_{I}$ . A possible explanation for the lower value time in the flexibility sub-sample is the significance of the value of safety. That is, when the respondent really trades-off between attributes, the value of time may decrease.

Following Jones-Lee (1990), we express individual *n*'s marginal rate of substitution of wealth for risk by the ratio  $WTP_{ns} = \partial \rho$ , where  $\partial \rho$  is the absolute change in risk level and  $WTP_{ns}$  defines the willingness to pay for a safety increase equal to  $\partial \rho$ . In a group of individuals affected by an increase in safety, the aggregated value of a statistical life equals the mean of these individual marginal rates of substitution. If the individuals affected by a risk reduction are relatively homogenous and the individual risk reductions are small, the value of statistical life is independent of the size of the group and the pattern of individual

risk reductions. Under such circumstances, the total value of statistical life equals  $WTP_s/\partial\rho$  where  $WTP_s$  equals the population mean willingness to pay.

In our survey, the  $VOA_{s, II}$  for every work trip equals SEK 0.07 per percentage reduction in risk. Given that most people have two work trips per day and, on average, 220 workdays per year, we get an annual value of safety, a willingness to pay for safety (*WTP*<sub>s</sub>), equal to SEK 30,80 per percentage reduction in risk.

Because we used the respondents' perceptions of safety we are unable to calculate a *point estimate* for the value of statistical life.<sup>17</sup> Nonetheless, we can give a *function* for the value of statistical life at different levels of perceived risk. Assuming different initial levels of risk for work trips and reducing the risk by exactly 1/100,000 gives rise to the function in Figure 2. For example, in 1997, the *objective* risk for being *killed* or *seriously injured* in a car was approximately 29 per 100,000 of the average population (SIKA, 1998). If the average respondent perceives this objective risk as the initial risk for being killed in traffic, the value of statistical life is equal to SEK 10.6 million, based on the assumptions above.<sup>18</sup>.

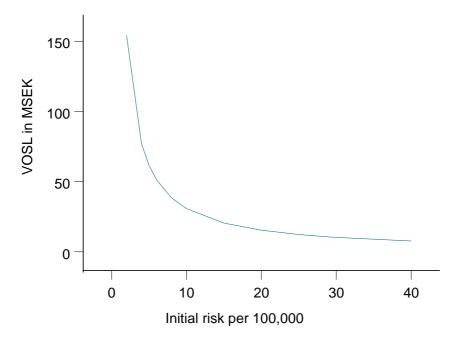


Figure 2. Value of statistical life (VOSL) at different initial risk levels.

<sup>18</sup> Reducing the risk from 29 to 28 in a hundred thousand equals a risk reduction of 3.45 percent. For all work trips in a year we have:  $(-0.07*-3.45*2*220)*(1/100,000)^{-1})$  or, more generally,

<sup>&</sup>lt;sup>17</sup> We assume, everywhere mentioned, that the initial risk is equal to the risk of being *killed* in road traffic. Therefore, the value of statistical life can be calculated from every assumed initial risk level. However, if the respondents perceive the initial risk (safety) differently, so that, for instance, the initial risk is perceived as the risk for being *killed* or *seriously injured* in traffic, an unambiguous value of statistical life can not be calculated.

 $VOSL = VOA_s * \partial \pi * n_{trips} * n_{days} * \partial \rho^{-1}$ , where VOSL is short for value of statistical life and  $\partial \pi$  is the risk reduction expressed in percent.

Similarly, the risk for being *killed* in car was 4 per 100,000 of the average population in 1997 (SIKA, 1998).<sup>19</sup> If the respondent correctly perceives this objective risk level as the initial risk level, the value of statistical life is equal SEK 77 million. Thus, the crucial assumption for the value of statistical life is the assumption about the perceived initial risk. A Swedish contingent valuation survey (Persson et al., 1998) has found that the mean perceived risk for being killed in a traffic accident during a year equal to 72/100,000 (median: 4/100,000). Employing their mean value gives a value of statistical life for work trips.

#### 4.2. Random Parameters Logit

The standard logit model imposes several restrictions on the parameters of the model (McFadden, 1981; Ben-Akiva and Lerman, 1985). First, the parameters are assumed to be the same for all respondents, meaning that two identical respondents (with respect to the observed variables) must have the same parameter values, i.e. tastes. Second, following from the assumption of IID extreme value error term, the logit model suffers from the independence from irrelevant alternatives (IIA) (cf. Ben-Akiva and Lerman, 1985). With IIA, the logit model necessarily predicts that a change in one attribute of an alternative (or the introduction of a new alternative or the elimination of an existing alternative) changes the probability for the other alternatives proportionately, so that the ratios of probabilities remain unchanged (Brownstone and Train, 1999). In many circumstances, this is an unrealistic assumption, resulting in implausible substitution patterns. Third, in cases where there are repeated choices, either over time, as in revealed preferences situations, or over choices, as in stated preferences are independent over time and choices so that there is no correlation in the unobserved utility over time and choices. In many settings, this is also an unrealistic assumption.

To relax the above restrictions, a less restrictive model, the "random parameters" logit model<sup>20</sup>, was estimated. The random parameters logit model is a generalisation of the standard logit model, where the parameters of the attributes are allowed to vary randomly over respondents. The random parameters logit model does not exhibit the IIA property and may, therefore, represent any substitution pattern (Train, 1998). Furthermore, in the random parameters logit model the dependence of unobservable variables over choices is explicitly modelled. For instance, the individual's value of time (the ratio between the time and cost parameters) may depend on the unobserved variable "patience" which may vary among the respondents (Horowitz, 1981).

Random parameters logit models have been applied in various settings, e.g. in recreational demand models for the choice of fishing site (Train, 1998), in consumption demand models for households' choice of appliance efficiency level (Revelt and Train, 1998) and households' choice between vehicles with different fuel types (Brownstone and Train, 1999), and in travel choice models (Algers et al., 1998).

<sup>&</sup>lt;sup>19</sup> In 1997, 348 persons were killed in car accidents (SIKA, 1998). It is important to remember that these are "primary" fatalities (people being killed when travelling by car), whereas there are also "secondary" fatalities (people not travelling by car who are killed by a car).

<sup>&</sup>lt;sup>20</sup> Also known as mixed logit, random parameters logit and error components logit.

In the random parameters logit model, the indirect utility of alternative *i* in choice situation *t* for individual *n*, is specified as  $V_{nit} = \mathbf{\beta}'_n \mathbf{x}_{nit} + \varepsilon_{nit}$ , where  $\mathbf{x}_{nit}$  is a vector of explanatory variables and  $\beta_n$  is a vector of unobserved parameters that varies in the population. The disturbance term,  $\varepsilon_{nit}$ , is still assumed to be IID extreme value distributed. This specification is analogous to the standard logit specification of  $V_i$  in Equation (3), except for the parameter vector  $\beta_n$  which varies over respondents and the sub-indexation *t* which accounts for the data's panel character (the repeated choices). Following Train (1998) the parameter vector  $\beta_n$  is expressed as the sum of the population means, b, and individual deviations,  $\eta_n$ , where  $\eta_n$  represents the respondent's taste relative to the population's. The utility may be re-expressed as:

$$V_{nit} = \mathbf{b}' \mathbf{x}_{nit} + \mathbf{\eta}'_n \mathbf{x}_{nit} + \mathcal{E}_{nit}.$$
(4)

The term  $\mathbf{\eta}'_n \mathbf{x}_{nit} + \varepsilon_{nit}$  is unobserved and, as the researcher can not separate the individual deviation from the independent error term, the existence of  $\eta_n$  induces heteroscedasticity and correlation over choices. That is, because the respondent uses the same "tastes" when evaluating choices and, as the researcher is unable to completely observe these tastes, there is correlation in the unobserved part of the indirect utility function.

Generally, the unobserved parameter vector  $\beta_n$  is assumed to vary in the population with density  $f(\beta_n | \theta)$ , where  $\theta$  is a vector of the true parameters of the distribution (for example, the mean and standard deviation).  $\beta_n$  represents individual *n*'s "tastes", which are assumed to be constant over the repeated choices.<sup>21</sup> Our objective is to estimate  $\theta$ , the true population parameters that describe the distribution of the individual parameters. Here, we assume  $\eta_n$  and  $\beta_n$  to be normally distributed,  $\eta_n \sim N(0, \Sigma)$  and  $\beta_n \sim N(\mathbf{b}, \Sigma)$  and  $\eta_n \sim N(\mathbf{0}, \Sigma)$ .<sup>22</sup> Thus, b and  $\Sigma$  are the vectors of parameters we wish to estimate.

We estimate the these parameters in Equation (4) by maximisation of a simulated loglikelihood function (MSL) with R=1,000 repetitions.<sup>23</sup> All random parameters logit models were estimated with ten different starting values to ensure the finding of a global maximum.<sup>24</sup>

To begin with, the most obvious way to introduce individual heterogeneity would be to estimate models with random intercepts while keeping all other variables fixed. However, compared to the standard logit models, such models do not improve the (simulated) likelihood functions and are rejected by likelihood ratio tests (not reported). We, therefore, fix the intercept terms of both models, and let, instead, the explanatory variables be random.<sup>25</sup>

<sup>&</sup>lt;sup>21</sup> Modification could be done so that the respondent's tastes vary over time (Train, 1998). However, in our case we assume that the time elapsed between the consecutive choices is too small to significantly affect the respondent's tastes.

<sup>&</sup>lt;sup>22</sup>  $\Sigma = diag(\sigma_p)$  where p equals the explanatory variables used in the standard logit model.

<sup>&</sup>lt;sup>23</sup> See Train (1998) and Revelt and Train (1998) on the maximisation of simulated likelihoods. The estimations was conducted in a GAUSS program written by Kenneth Train, David Revelt and Paul Ruud.

<sup>&</sup>lt;sup>24</sup> As starting values we use  $\mathbf{b} = \hat{\lambda} + \zeta$ , where  $\hat{\lambda}$  is a *p*×1 vector of standard logit estimates and  $\zeta \sim N(0, 0.1)$ . For  $\Sigma$ , we use draws from the distribution  $\Sigma \sim Uniform[0, 1]$ .

<sup>&</sup>lt;sup>25</sup> Allowing all parameters to vary results in identification problems (Revelt and Train, 1998).

For the flexibility sub-sample, this model is rejected by a likelihood ratio test ( $\ln \ell = -69.84$ ) but, for the environmental friendliness sub-sample, we can not reject this model ( $\ln \ell = -61.78$ ).

The results from the environmental friendliness estimation are reported in Table 7. The mean of the time variable is significant at the five percent level and the mean and standard deviation of cost are significant at the ten percent level. Thus, there are significant variations in the cost parameter among the respondents in this sub-sample. The point estimate of the time parameter implies that six percent of the respondents have a positive time parameter, while the point estimate of the cost parameter implies that a few respondents may enjoy the time travelling to work, it seems less plausible that some respondents enjoy having more expensive work trips. An explanation for this unexpected value, also found by Algers et al. (1998), may be some respondents' eligibility to tax deductions for work trips.

Comparing the random parameters logit parameter estimates with those from the standard logit, we see that the estimated parameters are larger in absolute magnitude in the random parameters logit. Because the variance of the IID error term is greater in the standard logit than in the random parameters logit and, as the scale of utility is determined by the normalisation of the error term, this is what is to be expected (Brownstone and Train, 1999).

Attribute	Parameter	St dev	t-value	
Constant	0.119	0.426	0.280	
Mean of time	-0.131	0.062	-2.107	
St dev of time	0.086	0.062	1.382	
Mean of cost	-0.190	0.113	-1.680	
St dev of cost	0.191	0.111	1.715	
Mean of safety	0.004	0.005	0.954	
St dev of safety	0.007	0.009	0.773	
Mean of env.friendl.	0.001	0.002	0.793	
St dev of env.friendl.	0.000	0.002	0.176	
N	77			
Choices made	143			
SLRI	0.376			
ln $\ell$ (simulated)	-61.78			

Table 7. Environmental friendliness sub-sample (I): Results from random parameters
logit estimations of stated choices

Note: The simulated likelihood ratio index (*SLRI*)=1-( $\ln \ell_s/\ln \ell_0$ ).

The point estimate of  $VOT_I$  (using  $\hat{b}_{time}/\hat{b}_{cost}$ ) is now SEK 41.37 per hour. Thus, it is lower than the equivalent from the standard logit model. Lower value of time, when employing a random parameters logit model with normal parameters, has also been found by Algers et al. (1998). Because each respondent now has an "own" value of time - which is the ratio between two normally distributed variables - the correct mean value of time would be given by integrating over this distribution. However, this considerably complicates the calculation and also leads to some individuals having implausible values of time as the normal distribution allows both negative and positive individual parameters.

One may assume lognormal distributions for the time and cost parameters and, thereby, make sure that these variables get positive individual parameter values (when estimated on the negative of respective variable). However, for our dataset, this model failed to converge. Convergence failures in estimations with log-normally distributed variables are also recognised by Algers et al. (1998) and Brownstone and Train (1999).

#### **5.** CONCLUSION

Based on a stated choice experiment, we estimate standard and random parameters logit models of commuters' choice of travel mode. The choices are explained by the individual's perceptions of five modal attributes; time, cost, safety, environmental friendliness and flexibility. Time, cost and safety were evaluated by all respondents while environmental friendliness and flexibility were used interchangeably. Depending on the attribute evaluated by the respondent, the sample was divided into an environmental friendliness and a flexibility sub-sample. The respondents' preferences were elicited by a novel method in which the names of two modes (car and bus) were used as attribute levels instead of objective levels. Based on the respondents' revealed preferences, we infer "perceptions-based" real values for the time and cost attributes. In the analysis, the inferred values are used together with the perceptions of the other attributes. From the estimated model we derive a value of time for work trips and a monetary value for safety improvements (value of statistical life). Time savings and saved statistical lives are the major benefits in project evaluations of the Swedish National Road Administration (SNRA). Whereas the SNRA uses values derived from separate studies, we estimate them in the same model.

The survey produced several interesting findings. Overall, time and cost are found to be the most important attributes for work trip modal choice, while safety is non-negligible in one of the sub-samples. Whereas a model specification with normally distributed parameters (the random parameters logit) improved model fit for the environmental friendliness sub-sample, the standard logit model with fixed parameters could not be rejected for the flexibility sub-sample. For both sub-samples, the models with the best fit (highest log-likelihoods) resulted in values of time of about SEK 40 per hour. In a national survey from 1994 (Algers et al., 1995), the value of time for work trips was found to be SEK 34 - 54 per hour for bus, car and train. Adjusted for price changes, these figures are SEK 35 - 56 (1997 prices), which encompasses our estimate, even though a direct comparison is erroneous because the different values of time are estimated on different travel modes. Thus, in a "test-retest" sense, the point estimates of the value of time in our survey are reliable. In 2005, the value of time recommended to the Swedish National Road Administration (SNRA) for cost-benefit calculations is SEK 42 per hour for regional (<50 kilometres) trips (SIKA, 2005).

In this survey the work trip value of statistical life in road traffic depends on the perceived baseline, or initial, risk. We find a lower bound for the value of statistical life equal to SEK 4.3 million (1997 prices), based on a perceived initial risk of 72/100,000. As is recognised in psychological research, people tend to underestimate the risks of well known, voluntary, low probability events (Lichtenstein et al., 1978). Because most people probably

define driving to work as a low probability event, a much higher value of statistical life is implied. A Swedish contingent valuation survey found the value of statistical life to be SEK 19.4 million (Persson et al., 1998). In 2005, the value of statistical life recommended to the SNRA for cost-benefit calculations was SEK 14.2 million (SIKA, 2005).

Though many things could have been done differently, the use of perceptions data is a novelty to SP analysis.<sup>26</sup> And although our perceptions data do not dramatically change, or improve, the value of time estimate, the estimations show that perceptions can be useful for explaining travel mode choices. Nonetheless, the findings from this survey are encouraging and constitute a point of departure for future research. In order to improve model fit and predictive power of travel choice models, future research will focus on identifying latent variables that affect mode choice. For instance, purposefully collected survey data, modelled in linear structural equation models, where the latent variables are related to indicator variables (cf. Vredin Johansson et al., 2006; Morikawa et al., 1990; Bollen, 1989), will, by estimation in the LISREL program (Jöreskog and Sörbom, 1993), help to identify latent variables. Such estimations could provide fitted values that may be used as explanatory variables in discrete travel choice models.

#### **ACKNOWLEDGMENTS**

The author would like to thank Per Johansson for generous and competent guidance throughout the work on this essay. The essay also benefited much from the comments and suggestions of Thomas Aronsson, Jörgen Hellström, Lars Hultkrantz and Karl-Gustaf Löfgren. Financial support from The Swedish Road Administration is gratefully acknowledged. The usual disclaimer applies.

## APPENDIX A: THE PERCEPTIONS AND CHOICE QUESTIONS (QUESTIONNAIRE A, QUESTIONS 19-21)

In the question below we would like you to *compare* two different travel modes for <u>a</u> *journey to your workplace;* car and bus. We would like you to do the comparisons *even* if you do not have a car and *even* if it is impossible to go by bus to your workplace. Try to imagine what it would be like if both alternatives existed.

19. As you know different travel modes are associated with different attributes, for example different travel time, cost, safety and environmental friendliness. How do you *perceive* the different travel modes with respect to respective attribute?

<sup>&</sup>lt;sup>26</sup> For example, information gathering on exactly how the respondents perceive the attributes should be improved. Moreover, the survey should account better for situational constraints (such as bad weather and/or seasonal differences) and, in order to maximise validity, be related to specific and recent choice situations.

Start out from the car's attributes and make comparisons. The attributes of the car are always 100%. *Example:* If you perceive the travel time with bus *equal to* the travel time with car, you answer 100%. If you perceive the travel time with bus *three times* as long as the travel time with car, you answer 300%. If you perceive the travel time with bus *half of* the travel time with car, you answer 50%. You may, of course, use any percentages - the important thing is that you, as accurately as possible, describe *your own* perceptions.

a)	Compared with <i>car</i> , I perceive the TRAVEL TIME with <i>bus</i> to be:	_%
b)	Compared with <i>car</i> , I perceive the COST with <i>bus</i> to be:	_%
c)	Compared with <i>car</i> , I perceive the SAFETY with <i>bus</i> to be:	%
Ca	moved with east I perceive the ENVIDONMENTAL EDIEND I INESS	

Compared with *car*, I perceive the ENVIRONMENTAL FRIEND-LINESS with *bus* to be: \_\_\_\_\_%

We ask the following two questions to find out how people value different attributes in a journey to work.

Suppose you are confronted with the choice between two hypothetical travel modes (A and B) for a journey to your workplace. These travel modes are characterised by four different attributes (travel time, cost, safety and environmental friendliness). There are no "real" travel modes that are described, but travel modes characterised by combinations of different attributes. Now we ask you to make a choice between A and B on the basis of the attributes that characterise travel modes A and B. How you choose is up to you, but we ask you to think thoroughly through your opinions before answering question 20 and 21.

*Remember:* There is no right or wrong answer! You have the right to think and choose as you like - it is *your* opinion that are of interest to us.

Attributes	А	В	
Travel time like:	Car	Bus	
Cost like:	Bus	Car	
Safety like:	Car	Bus	
Environmental friendliness like:	Car	Bus	

#### 20. Which alternative (A eller B) do you prefer?

Answer: I prefer (mark with one cross):(	)	Travel mode A
(	)	Travel mode B
(	)	The alternatives are equal

21. Which alternative (A eller B) do you prefer?

Attributes	А		В		
Travel time like:	Bus		Car		
Travel cost like:	Bus		Car		
Safety like:	Car		Bus		
Environmental friendliness like:	Car		Bus		
Answer: I prefer (mark with one cross):(		) )	Travel mode A Travel mode B		
			The alternatives are equal		

## **APPENDIX B:** THE EXPERIMENTAL DESIGN

Following the notation of Box et al. (1978) the three variables and the combination variable are denoted as follows; 1=t (travel time), 2=c (travel cost), 3=s (safety), 4=ef(Environmental friendliness/Flexibility). All variables have two levels, a high level and a low level. As previously, we employ a plus sign (+) to indicate the high level and a minus sign (-) to indicate the low level. Furthermore, if the elements of any column are multiplied by themselves, we obtain a column of plus signs, I, i.e. I=11=22=33=44.

Alt.		Va	riable	es					Inter	actior	n effe	cts			
	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
1	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+
2	+	-	-	-	-	-	-	+	+	+	+	+	+	-	-
3	-	+	-	-	-	+	+	-	-	+	+	+	-	+	-
4	+	+	-	-	+	-	-	-	-	+	-	-	+	+	+
5	-	-	+	-	+	-	+	-	+	-	+	-	+	+	-
6	+	-	+	-	-	+	-	-	+	-	-	+	-	+	+
7	-	+	+	-	-	-	+	+	-	-	-	+	+	-	+
8	+	+	+	-	+	+	-	+	-	-	+	-	-	-	-
9	-	-	-	+	+	+	-	+	-	-	-	+	+	+	-
10	+	-	-	+	-	-	+	+	-	-	+	-	-	+	+
11	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+
12	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
13	-	-	+	+	+	-	-	-	-	+	+	+	-	-	+
14	+	-	+	+	-	+	+	-	-	+	-	-	+	-	-
15	-	+	+	+	-	-	-	+	+	+	-	-	-	+	-
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

## Table B1. Signs for a 2<sup>4</sup> factorial design

In the following description of the experimental design, it is important to distinguish between the concepts of alternatives, blocks, pairs and sub-sub-samples (denoted sss in Table B2). When we have four variables varying at two different levels and desire two pair-wise comparisons in each sub-sub-sample, we need three "block" variables to do the job. Therefore, let us introduce the block variables, B<sub>1</sub>, B<sub>2</sub> and B<sub>3</sub>, where B<sub>1</sub>=12, B<sub>2</sub>=13 and B<sub>3</sub>=34 (see Table B2). Hence, B<sub>1</sub> is equal to the two-way interaction between *t* and *c*, B<sub>2</sub> is equal to the two-way interaction between *t* and *s* and B<sub>3</sub> is equal to the two-way interaction between *s* and *ef*.

Alternative	<i>B</i> <sub>1</sub>	<i>B</i> <sub>2</sub>	<i>B</i> <sub>3</sub>	Pair no	SSS	Pair no	Alternatives	Questionnaire
1	+	+	+	8	1	8	1, 16	A and E
2	-	-	+	7				
3	-	+	+	6		7	2, 15	A and E
4	+	-	+	5				
5	+	-	-	4	2	6	3, 14	B and F
6	-	+	-	3				
7	-	-	-	2		5	4, 13	B and F
8	+	+	-	1				
9	+	+	-	1	3	4	5, 12	C and G
10	-	-	-	2				
11	-	+	-	3		3	6,11	C and G
12	+	-	-	4				
13	+	-	+	5	4	2	7, 10	D and H
14	-	+	+	6				
15	-	-	+	7		1	8,9	D and H
16	+	+	+	8				

 Table B2. Block variables and experimental design

Depending on the signs of the block variables, we assign "pair" numbers to the different alternatives. Hence, the two alternatives with all block variables at the high level (+ + +) constitute a pair (pair no 8) while alternatives with block variables minus, minus, plus (- +) constitute another (pair no 7), and so on. The result of using the block variables is eight different pairs. These pairs are divided into four sub-sub-samples consisting of two pairs each. For example, sub-sub-sample one consists of two pair-wise comparisons, pairs number eight (8) and seven (7), where pair eight equals the choice between alternative one and 16 and pair seven equals the choice between alternative two and 15. For the other three sub-sub-samples the alternatives are given in Table B2. The two alternatives in each pair are complementary in signs, as the signs of the attribute levels of the first alternative are reversed in the second. Therefore, each pair is said to be a *fold-over* pair (Box et al., 1978). For example, in pair no 8 there is a choice between alternative one and 16. Regarding Table B1, we see that while alternative one is low on all variables, alternative 16 is high on all variables.

This blocking arrangement is copied from Box et al. (1978, p. 347). Clearly, there are some, but not so many, other possible combinations of block variables. However, caution is called for, because the use of wrong block variables may lead to main attribute effects being confounded with block effects. Using any other three *possible* (without confounding the block effects with the main attribute effects) combinations of block variables leads to the same

result as the ones we use; all two-way interactions and the four-way interaction are confounded with the block effects and the pairs consist of the same alternatives.

If the we use the notation and the multiplicative properties from above, we have;  $B_1B_2=1213=23$ ,  $B_1B_3=1234$ ,  $B_2B_3=1334=14$  and  $B_1B_2B_3=121334=24$ , which means that interaction effects 12 (time-cost), 13 (time-safety), 14 (time-environmental friendliness/flexibility), 23 (cost-safety), 24 (cost-environmental friendliness/flexibility), 34 (safety-environmental friendliness/flexibility) and 1234 (time-cost-safety-environmental friendliness/flexibility) are confounded with the block effects and, therefore, not possible to estimate separately. However, none of the main attribute effects are confounded.

#### REFERENCES

- Adamowicz, W., Swait, J., Boxall, P., Louviere J., and Williams, M. (1997). Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation. *Journal of Environmental Economics* and Management, 32, 65-84.
- Algers, S., Bergström, P., Dahlberg, M., and Lindqvist Dillén, J. (1998). *Mixed Logit Estimation of the Value of Travel Time*. Working Paper 15, Department of Economics, Uppsala University.
- Algers, S., Hugosson, B., and Lindqvist Dillén, J. (1995). 1994 års tidsvärdesstudie. Slutrapport. Del 1 Resultat. Transek, mimeo. (In Swedish with a summary in English).
- Batsell, R. R., and Louviere, J. J. (1991). Experimental Analysis of Choice. *Marketing Letters*, 2(3), 199-214.
- Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal*, 75, 493-517.
- Ben-Akiva, M., and Lerman, S. R. (1985). Discrete Choice Analysis: Theory and Application to Travel Demand. Cambridge, MA, MIT Press.
- Björkman, M., and Ekman, G. (1957). *Experimentalpsykologiska metoder*. Stockholm, Almqvist and Wiksell.
- Bollen, K. A. (1989). Structural Equations with Latent Variables. New York, Wiley.
- Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978). Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building. New York, John Wiley and Sons.
- Brownstone, D., and Train, K. (1999). Forecasting New Product Penetrations with Flexible Substitution Patterns. *Journal of Econometrics*, 89, 109-129.
- DeSerpa, A. C. (1971). A Theory of the Economics of Time. *The Economic Journal*, 81, 828-846.
- Greene, W. H. (1993). Econometric Analysis (2nd edition). New York, Macmillan.
- Gårder, P., Leden, L., and Thedéen, T. (1994). Safety Implications of Bicycle Paths at Signalized Intersections. *Accident Analysis and Prevention*, *26(4)*, 429-439.
- Horowitz, J. L. (1981). Identification and Diagnosis of Specification Errors in the Multinominal Logit Model. *Transportation Research*, 15B, 345-360.

- Jara-Diaz, S. R. (1998). Time and Income in Travel Choice. Towards a Microeconomic Activity-Based Theoretical Framework. In T. Gärling, T. Laitila and K. Westin (Eds.), *Theoretical Foundations of Travel Choice Modeling* (51-74). Oxford, Elsevier Pergamon.
- Jara-Diaz, S. R., and Videla, J. (1989). Detection of Income Effect in Mode Choice: Theory and Application. *Transportation Research*, 23B, 393-400.
- Jones-Lee, M. W. (1990). The Value of Transport Safety. Oxford Review of Economic Policy, 6(2), 39-60.
- Jöreskog, K. G., and Sörbom, D. (1993). *LISREL 8: Structural Equation Modelling with the SIMPLIS Command Language*. Chicago, SSI.
- Kahneman, D., and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263-291.
- Kommunikationsdepartementet (1997): På väg mot det trafiksäkra samhället. Ds 1997:13, Stockholm.
- Koppelman, F. S., and Pas, E. I. (1980). Travel-Choice Behavior: Models of Perceptions, Feelings, Preference and Choice. *Transportation Research Record*, 765, 26-33.
- Lancaster, K. J. (1966). A New Approach to Consumer Theory. Journal of Political Economy, 14, 132-157.
- Lee, T. R. (1981). Perceptions of Risk. The Public's Perception of Risk and the Question of Irrationality. *Proceedings of the Royal Society of London, A376*, 5-16.
- Lenner, M. (1993). Energy Consumption and Exhaust Emissions Regarding Different Means and Modes of Transportation. VTI Meddelande nr 718. (In Swedish with a summary in English).
- Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., and Combs, B. (1978). Judged Frequency of Lethal Events. *Journal of Experimental Psychology: Human Learning and Memory*, 4(6), 551-578.
- Louviere, J. J., and Johnson, R. D. (1991). Using Conjoint Analysis to Measure Retail Image. In A. Ghosh, and C. A. Ingene (Eds), *Research in Marketing. Spatial Analysis in Marketing: Theory, Methods and Applications*. Supplement 5, 137-156, Jai Press Inc.
- McFadden, D. (1981). Econometric Models of Probabilistic Choice. In C. Manski, and D. McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications* (198-272), Cambridge, The MIT Press.
- Morikawa, T., Ben-Akiva, M., and McFadden, D. (1990). Incorporating Psychometric Data in Econometric Choice Models. Draft version, July 1990.
- Mowen, J. C. (1990). Consumer Behaviour (2nd edition). New York, McMillan.
- Noland, R. B. (1995). Perceived Risk and Modal Choice: Risk Compensation in Transportation Systems. Accident Analysis and Prevention, 27(4), 503-521.
- Payne, J. W. (1976). Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis. Organizational Behavior and Human Performance, 16, 366-387.
- Persson, U., Nilsson, K., Hjalte, K., and Norinder, A. (1998). Beräkning av Vägverkets riskvärden. En kombination av contingent valuation-skattningar och uppmätta hälsoförluster hos vägtrafikskadade personer behandlade vid fyra sjukhus. The Swedish Institute of Health Economics (IHE), mimeo. (In Swedish).
- Revelt, D., and Train, K. (1998). Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. *Review of Economics and Statistics*, 80(4), 647-657.

- SCB (1998). Svenskarnas resor. Riks-RVU resultatrapport för 1997. Statistics Sweden. Stockholm. (In Swedish with a summary in English).
- SIKA (1998). Traffic Injuries 1997. Official Statistics of Sweden, Örebro, SCB.
- SIKA (2005). Kalkylvärden och kalkylmetoder (ASEK). En sammanfattning av Verksgruppens rekommendationer 2005. SIKA PM 2005:16. (In Swedish).
- Slovic, P., Fischoff, B., and Lichtenstein, S. (1981). Perceived Risk: Psychological Factors and Social Implications. *Proceedings of the Royal Society of London*, A376, 17-34.
- Train, K. and McFadden, D. (1978). The Goods/Leisure Tradeoff and Disaggregate Work Trip Mode Choice Models. *Transportation Research*, 12(5), 349-353.
- Train, K. E. (1998). Recreation Demand Models with Taste Differences over People. Land Economics, 74(2), 230-239.
- Vredin Johansson, M., Heldt, T., and Johansson, P. (2006). The effects of attitudes and personality traits on mode choice. *Transportation Research Part A*, 40, 507-525.

Chapter 49

## **DO OUTSIDER-DOMINATED BOARDS AND LARGE BOARD SIZE CURTAIL FIRM RISK TAKING?**\*

## Michael Graham<sup>1</sup> and Alfred Yawson<sup>2</sup>

<sup>1</sup>School of Economics, Finance and Marketing, Royal Melbourne Institute of Technology (RMIT) University, Melbourne, Victoria 3001, Australia
<sup>2</sup>School of Banking and Finance, The University of New South Wales, Sydney, NSW 2052 Australia

#### ABSTRACT

This chapter examines the impact of board composition and board size on three firm risk taking variables; strategy risk, stock returns risk, and income instability risk. There is recognition in the literature that corporate governance processes need to encompass mechanisms for motivating board and managerial behavior towards enhancing firm risk taking. Agency theory and regulatory recommendations advocate for an increasingly greater roles for outsiders on the board of directors. The evidence documented here indicates a positive relationship between majority independent board composition and firm risk taking. The literature also suggests board size affects firm activities independent of other board attributes and that there are biases against risk taking as board size grows. We do not, however, find any evidence to support the proposition that large board size influences firm risk taking. Thus, whereas the recommendation of increasing independent members on board of directors by regulatory bodies as well as the Cadbury (1992) and Hampel (1998) facilitate wealth creation by firms, the total representation on the board of directors does not exert any measurable influence on firms' wealth creation.

**Keywords**: agency theory, board of directors, board size, corporate governance, risk taking, firm enterprise.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Risk Taking and Legal Contexts*, edited by Rachel N. Kelian, published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

#### **1. INTRODUCTION**

Short et al (1999) provide a framework of corporate governance and stress the need for governance processes to encompass mechanisms for motivating managerial behavior towards increasing the wealth of the business. That is, to enhance economic enterprise activities or risk-taking<sup>68</sup>. Firm risk-taking is generally a high risk-high return strategy that is attractive to stockholders given that they expect a positive effect on performance. Furthermore, stockholders can reduce their inherent risk by diversifying their investment portfolios.

Agency theory assigns the task of checking self-serving and value-decreasing behaviors by managers to the board, and in particular outside members of the board of directors. Generally, there is the expectation that board of directors composed primarily of outside directors should be superior to boards composed of insider directors in minimizing agency cost and contributing to managerial effectiveness and organizational performance. This view is supported by a number of empirical studies (see e.g. Baysinger and Butler 1985 for outside directors and firm performance, Weisbach 1988 for CEO turnover, Byrd and Hickman 1992 for tender offer bids, and Brickley et al. 1994 for poison pill adoptions and control auctions). Along this line, Rosenstein and Wyatt (1990) document positive investor reactions to appointment of outside directors. An important development on this front relates to recent regulatory activity in many countries, for example Finland, UK, and Australia that recommends that majority of the board of listed firms should be outside directors.<sup>69</sup> The root of this recommendation can be traced to the Cadbury (1992) and the Hampel (1998) reports.

A critical agency issue involving the recommendation of greater representation by independent directors on the board of listed firms as a means of minimizing agency cost that has not been given much attention in the literature is their impact on firm risk taking. We believe this is an important area that needs evidencing and this chapter examines this issue. Furthermore, this chapter investigates the impact of board size on firm risk taking. The number of directors on a board has been theorized to affect corporate governance independent of other board attributes. Jensen (1993) theorizes that boards of directors that are too large may not be able to operate effectively because the co-ordination and information processing problems outweigh the advantages of having a large number of people to draw on. Lipton and Lorsch (1992) point to the dysfunctional nature of board are amplified as board size gets larger. Although Lipton and Lorsch's (1992) recommendation did not specify the optimal board size, the proposal amounts to a supposition that even if the board's capacity for monitoring increases with its size, the benefits are outweighed by such costs as biases against risk-taking.

The Cadbury (1992) and Hampel (1998) reports have shaped the various changes in corporate governance principles and practices around the world. However, to the best knowledge of the authors, the academic literature has not captured the relationships studied in this chapter prior to the last of the two reports. For this reason, this chapter limits the data used in the empirical analysis to the year 1998 to capture important aspects of the

<sup>&</sup>lt;sup>68</sup> Enterprise activities and risk-taking are used interchangeably to mean the same thing in this chapter.

<sup>&</sup>lt;sup>69</sup> See, for example, corporate governance recommendation of The Nordic Exchange and the Australian Stock Exchange.

relationships that led to the recommendation of having majority outsiders on firms' board of directors.

Miller and Bromiley (1990) suggest that firm risk taking measures used in the literature can be grouped into three categories: income instability risk, strategic risk, and stock return risk. Following that, this chapter studies the effect of board composition and board size on these three measures of firm enterprise. This chapter examines the issues raised above using data from Finland.

We find a positive relationship between outside dominated board of directors and firm risk taking when the relevant measures of risk taking are the proxies for stock returns (measured by beta) and strategic (measured by capital intensity). Board composition relates differently to different risk taking measures and the relationship is strongest when the relevant risk measure is capital intensity. There is no evidence found to support the assertion that board size influences firm risk taking. This chapter contributes to the body of empirical literature that examines whether boards of directors, affect firm enterprise activities. Mayer (1997 p.152) remarked that "corporate governance has become a subject on which opinion has drowned fact". To play a part in enhancing factual representation in the subject, the findings in this chapter add to the existing stock of empirical knowledge on how some elements of the governance process may influence the economic performance of companies. The contributions of this chapter to the existing literature enhance our understanding of the governance process and business prosperity, an important aspect of corporate performance.<sup>70</sup>

The remainder of the chapter is structured as follows: We examine agency theory and firm control and follow that up by looking into the impact of outsider-dominated boards on firm enterprise. We then investigate board size and firm enterprise. The sample construction, methodology and results from our empirical investigations follow. We then conclude the chapter.

#### **2.** Agency Theory and Firm Control

Agency theory views the board as a potentially effective element of corporate governance and, hence internal control (Fama and Jensen 1983). The issue of control is very important in agency theory, where the traditional interest lies in contracting. Agency theory considers the optimal contract form for that ubiquitous control relationship in which the principal delegates work to the agent (the principal-agent relationship is mediated by the board of directors). Eisenhardt (1985) formalize the agency problem that determines the optimal contract for the agent's service. The theory can be stated in two cases. Firstly, when the behavior of the agent is observed, a behavior-based contract is optimal because the agent's behavior is the purchased commodity. In this case, there is complete information. The two parties, principal and the agent, know what the agent has done. The second is a situation of incomplete information. Agents are aware of their behavior but the principal is not. A quandary arises because the principal cannot determine if the agent has acted appropriately. If the agent is to be rewarded based upon agreed job behaviors, but without confirmation of those behaviors by

<sup>&</sup>lt;sup>70</sup> For example, the literature suggests that earlier studies involving board size have centred on its effect on CEO compensation (see e.g., Holthausen and Larcker 1993), and firms performance (see for instance, Yermack 1996, Eisenberg et al. 1998, and Conyon and Peck 1998).

the principal, the agent may shirk. Given this, the agent cannot be relied on to perform as agreed. In this case of incomplete information, the principal has two alternatives to put things right. First, information about the agent's behaviors could be purchased and rewards directed at those behaviors. This requires employing surveillance mechanisms such as cost accounting measures, budgeting systems, or additional layers of management. On the other hand, the principal can reward the agent based on outcomes (e.g., profitability). Such outcomes are surrogate measures for behaviors. In this alternative, however, the agent is rewarded or punished for outcomes partially outside his or her control. That is, good results could be attained despite poor efforts and poor results can occur despite good efforts.

Agency theory, thus, suggests two underlying strategies of control. These are behaviorbased and outcome-based strategies of control. Both of these strategies rely on performance evaluation. Performance-based control strategies emphasize monitoring, evaluating, and rewarding an agent's performance and thus focus on the information aspect of control. Agents' performance can be evaluated on inferences concerning the quality of their decisions or on the measurable outcomes of decision-making process. Senior level corporate managers make decisions under uncertainty that are not programmable as other non-managerial tasks. Hence, following Baysinger and Hoskisson (1990), the terms strategic controls and financial controls, respectively, are used in this paper to convey behavioral and outcome control concepts in the context of controlling corporate managers. Gupta (1987) describes what is referred to here as strategic control in terms of openness in the relationship between corporate-level managers and division managers and the willingness of management to rely on subjective information when evaluating division managers' performances. That is, under a system of strategic controls, division managers are evaluated on the basis of how strategically desirable their decisions were before implementation and on the basis of performance of the firm after the decisions were implemented. On the other hand, in a system employing financial controls, managers are evaluated solely on the basis of their success in meeting performance criteria. Given the detailed nature of strategic controls, putting them into practice requires the controller to have more information than is required for implementing financial controls.

### **3. OUTSIDER-DOMINATED BOARDS OF DIRECTORS AND FIRM RISK TAKING**

The effect independent members of the board of directors have on firm enterprise activities flow from the control strategies they adopt in monitoring firm management, including the direction the firm takes in its risk taking activities. The composition of the board of directors determines the control strategies employed. Baysinger and Hoskisson (1990) suggest that outside directors usually have different types of information from inside directors. Hence, they are likely to differ in the control strategies they employ to advance firm events. Inside directors, in their position as participants in the decision processes, have access to information relevant to assessing strategic desirability of enterprise activities, regardless of their long or short run performance outcomes. Given that insiders and the CEO frequently interact in ways that are relevant to assessing the quality of the decision-making process, relations between the two parties are more likely to be open and subjective. Hence, in terms of information processing, it could be expected that the inside party of any board would evaluate firm enterprise proposals on the basis of strategic controls.

The outside board members usually have few of these informational advantages and this reduces the likelihood of using strategic controls. Outside directors, by definition, have limited contact with the day-to-day decision process of the firm. Their evaluation of firm processes and enterprise activities is limited to board interaction, at which point the strategic plans may be at their final stage, needing only ratification by the board. Thus, in terms of information processing, Baysinger and Hoskisson (1990) suggest that outside directors lack the type of subjective information needed for evaluating managerial decisions. Jensen (1993) also points to the severe information problems that limit the effectiveness of outside board members. For example, the top management of firms headed by the CEO almost always determines the information given to the board. This limitation on information severely hinders the ability of even highly talented outside board members to significantly contribute to the evaluation of company's risk taking strategy.

It must be noted that outside directors may be able to reach detached conclusion about quality of firm's wealth creation strategies through years of experience. However, Mintzberg et al. (1976) suggest that the very nature of strategic risk taking decisions makes them unique and unstructured. Consequently, for outside directors to fully appreciate the quality of decisions, they may need supplemental experience with that firm's process. This knowledge is what is not available to the truly independent outside director. On the other hand, outside directors who are former executives of the firm will be well informed. However, such directors are not truly independent and advocates for outside directors in the composition of the board of directors have independent outsiders in mind. Hence, former company employees do not fit into the category of outsiders on the board of directors. Consequently, whereas relations between inside directors and top management may be open and subjective, relations between outside directors and top management in discussing firm enterprise activities may be more objective and formulaic. In other words, it is expected that the predominance of outside directors on boards will be associated with financial controls in board-management relations. Along these lines, Hoskisson et al. (1995) argue that outside directors may favor expansion, a risk taking action, via external means, such as acquisitions to enter new markets as these are better suited to evaluation using financial criteria.

Given the type of information available to outside directors and their likely emphasis on financial controls, outsider-dominated boards may skew the direction of managerial effort away from optimally risky strategies that many shareholders prefer and the long-term orientation that is required for competitive advantage. From the perspective of top management, financial controls correlate managerial rewards directly with short-term variations in the market value of the firm (see e.g. Demski 1987, Gupta 1987). Functionally, this is achieved either by disciplinary practices that are highly sensitive to short-term cash flow, net profit, growth, or market share results or by compensation schemes that tie much of the pecuniary rewards to quarterly earnings (see Baysinger and Hoskisson 1990). Along the same line of reasoning, Fischel and Bradley (1986) discuss behavioral implications of alterations in corporate law, which increases management's liability for short-term performance declines. They argue that if managers are penalized whenever risk taking decisions that were optimal ex ante turn out poor ex post, they will tend to avoid risky projects.

Increasing managers' liability for financial performance in the short run may increase their diligence in maximizing short-term profits. It may, however, lessen their incentive to take risks and enhance firm enterprise activities. A greater outsider representation on the board of directors may have the same effect because the insiders' influence on the decisioncontrol process is reduced. This chapter examines this relationship.

#### 4. BOARD SIZE AND FIRM RISK TAKING

This chapter also examines the relationship between board size and firm risk taking. The board of directors, at the apex of the internal control system, has the responsibility for the firm's risk taking activities. There is the recognition that an active board can improve decision-making and, thus, firm enterprise activities. John and Senbet (1998) argue that the effectiveness of board is determined, among other things, by its size. Lipton and Lorsch (1992) and Jensen (1993) contend that board size affects corporate governance and firm activities independent of other board attributes. Their arguments focus on the productivity losses that arise when work groups grow large, an insight borrowed from organizational behavior research such as Hackman (1990). Jensen (1993) states that "... as groups increase in size they become less effective because the coordination and process problems overwhelm the advantage from having people to draw on..." Lipton and Lorsch (1992) also state that "... . the norms of behavior in most boardrooms are dysfunctional", because directors rarely criticize the policies of top managers or hold candid discussions about corporate performance, including the direction of firm enterprise activities. Jensen (1993) further suggests that larger boards lead to less candid discussion of firm enterprise direction. This situation arises because, according to Jensen (1993), there is "great emphasis on politeness and courtesy at the expense of truth and frankness in boardrooms". This implies that firms with large boards can reduce the board's ability to resist CEO control of firm agenda. Yermack (1996) suggests that "... CEO performance incentives provided by the board through compensation and the threat of dismissal operate less strongly as board size increases ... "

Indeed empirical results presented by Yermack (1996), Eisenberg et al. (1998), and Conyon and Peck (1998) suggest that board size affects firm profitability in both large and small firms. To reduce the problems of communication and co-ordination and the decreased ability of the board to set firm agenda, including risk taking, Lipton and Lorsch (1992) recommend limiting the membership of boards of directors. Implicitly this suggests costs, such as biases against risk taking, to large board sizes.

#### 5. SAMPLE CONSTRUCTION AND METHODOLOGY

Data is collected over a nine-year period, from 1990 to 1998, with focus on two sample periods, 1994 and 1998, for the empirical analyses. The estimation of the variables used in this chapter includes all the years between 1990 and 1998. As noted above, the second of the two reports, namely the Cadbury Report (1992) and the Hampel Report (1998), that has significantly influenced global corporate governance principles and practices was published in 1998. Therefore, this chapter limits the data used in the empirical analysis to the year 1998

to capture an important aspect of the relationships that led to the recommendation of having majority outsiders on firms' board of directors.

Generally, firms are selected from publicly traded companies in Finland satisfying two basic data requirements. First, the firm is required to have ownership data available for each sample year. Data regarding board members, institutional investors and block owners are obtained from the respective firms' annual reports. A further requirement is that firms included in the dataset should have five consecutive fiscal years of stock market and financial statement data, including the focus year, for each sample. The final sample consists of 48 firms in the 1994 sample and 68 firms in the 1998 sample.

Following previous research (for example, Gilson 1990 and Zahra 1996), this chapter considers outside directors as those who are not former employees (officers) of a firm or its subsidiaries (divisions), or do not possess contractual relationship with it. The only formal association between the outside directors and the firm are their duties as directors. The outsider-dominated boards are those boards that have higher numbers of independent outside directors than inside board members. In other words, it is measured by dividing the number of independent outside directors by the total number of directors on a board. The samples for years 1994 and 1998 are both dominated by outsider-dominated boards. There are 43 outsider-controlled boards in the 1994 sample, representing 89.6% of the firms in consideration, and 58 outsider-controlled boards in the 1998 sample, representing 85% of the firms under consideration. Data on the board of directors are accessed from the respective firm publication (annual reports). Table 1 presents summary statistics on the board of directors. From Table 1, it could be seen that there were, in total, 215 outside directors in the 1994 sample and 358 outside directors in 1998.

The board size is the total number of persons on the board of directors. This number is also accessed from company publications (annual reports) for the respective years. Table 1 also presents a summary statistics of the variable board size (under all directors). The average number of directors on the board was about the same for the two sample periods, 6.5 and 6.6 for 1994 and 1998, respectively. There were, in total, 310 directors in 1994 and 453 directors in 1998.

Variable	Mean	Median	STDEV	Mode	Min	Max	No. of directors	No. of firms
Panel A: 1994 Sample								
Outside directors	4.5	4	2.78	2	0	10	215	48
Inside directors	2	1	2.05	1	0	8	95	48
All directors	6.45	6	1.76	5	4	10	310	48
Panel B: 1998 Sample								
Outside directors	5.3	5	2.2	6	0	10	358	68
Inside directors	1.4	1	1.58	1	0	8	95	68
All directors	6.6	7	1.37	7	5	10	453	68

Table 1. Summary statistics of the board of directors: 1994 and 1998

Outside directors are independent directors who have no affiliations with the firm. Inside directors are managers in the firm.

Following Miller and Bromiley (1990) this chapter uses 3 measures of risk-taking to examine the hypothesized relationships. These are the proxy for income stream risk (the standard deviation of return on equity (ROE)), the proxy for industry or strategic risk (capital intensity), and risk based on stock returns (beta). Analogous to the methodology employed by Bowman (1980 and 1982), and Miller and Bromiley (1990), the standard deviation of return on equity over a five-year period for each sample firm is used in calculating the proxy for income stream risk. Capital intensity is calculated as the ratio of total assets to sales. This variable is calculated as the mean value over a five-year period. Capital intensity increases risk in two ways (see for instance Brealey and Myers 2006, Shapiro and Titman 1986). If capital inputs are less variable than labor inputs in the short run, a company choosing to produce a given output with large amounts of capital and low amounts of labor increases its fixed costs and lowers its variable cost. The firm consequently will experience larger variations in profits if demand fluctuates (see Lev 1974 for a detailed derivation of this point). In addition, a firm using large amounts of capital runs a high risk of capital obsolescence-the possibility that technological change will make its capital investment worth little or nothing. Beta, the risk-taking measure for stock returns data, is estimated from the conventional market model regression equation (see Sharpe 1964) over a three-year period using weekly returns. In the capital asset pricing model (CAPM), systematic risk reflects the sensitivity of return on a firm's stock to general market movement. Table 2 presents a summary statistics for the variables used in this chapter.

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
Panel A: 1994 Sample					
ROE (STDV)	14.37	9.61	18.47	1.1	98.22
Capital intensity	1.99	1.16	2.84	0.52	14.96
Beta	0.81	0.84	0.3	0.17	1.5
Firm size	1070	423	1422	26	5840
CEO and board ownership	3.1	0.1	9.3	0.00	56.00
Panel B: 1998 Sample					
ROE (STDV)	8.03	4.07	17.15	0.903	43.18
Capital intensity	1.81	0.95	3.37	0.002	19.17
Beta	0.68	0.69	0.33	0.04	1.37
Firm size	1326	345	2840	21	15414
CEO and board ownership	7.4	0.37	14.3	0	61.2

Table 2. Summary statistics of risk taking measures and control variables

The 1994 and 1998 samples included 48 firms and 68 firms, respectively. ROE (STDV) is the standard deviation of return on equity measured over a five year period; capital intensity is measured by the ratio of total assets to sales, averaged over a five year period. Beta is estimated from the conventional market model regression equation (see Sharpe 1964) over a three-year period using weekly returns Firm size is measured by total assets (millions of Euros). CEO and board ownership is the combined equity stakes held by the CEO and the board of directors (in percentage).

The methodology employed to examine the impact of outsider-dominated board of directors and board size on firm risk taking is a cross-sectional regression analysis in which firm enterprise activities is regressed against board composition and board size. The control variables introduced in the analysis pertain to equity ownership by the board and the CEO, firm size and industry effects. Board and CEO ownership is assessed from company annual reports. Total assets capture firm size effect. For the data period, firms were grouped into fourteen industrial classifications.<sup>71</sup> Due to the small sample size, we combined industries using industry relatedness as criteria and reduced the industry classification to three.

Following Wright et al. (1996), we investigate the relationship between board composition and board size on firm enterprise activities, by estimating equations (1) and (2), respectively:

$$Risk_{i,t}^{z} = \alpha_{0}^{z} + \alpha_{1}^{z}OU_{i,t} + \alpha_{2}^{z}FS_{i,t} + \alpha_{3}^{z}BCEO_{i,t} + \alpha_{4}^{z}d_{i,t}^{1} + \alpha_{5}^{z}d_{i,t}^{2} + \varepsilon_{i,t}$$
(1)

$$Risk_{i,t}^{z} = \alpha_{0}^{z} + \alpha_{1}^{z}BS_{i,t} + \alpha_{2}^{z}FS_{i,t} + \alpha_{3}^{z}BCEO_{i,t} + \alpha_{4}^{z}d_{i,t}^{1} + \alpha_{5}^{z}d_{i,t}^{2} + \varepsilon_{i,t}$$
(2)

where  $Risk_{i,t}^{z}$  is the risk-taking measure (or enterprise activities) for firm *i* at time *t* where z = 1 for standard deviation of return on equity, z = 2 for capital intensity, and z = 3 for beta;  $BS_{i,t}$  is board size variable (logarithm of board size);  $OU_{i,t}$  is the fraction of independent board of members for firm *i* at time *t*;  $FS_{i,t}$  is firm size (logarithm of total assets) for firm *i* at time *t*;  $BCEO_{i,t}$  is the percentage of equity held by the board and the CEO for firm *i* at time *t*;  $d_{i,t}^{1}$  is firm *i* classified in industry 1 at time *t*;  $d_{i,t}^{2}$  is firm *i* classified in industry 2 at time *t*;  $\varepsilon_{i,t}$  is the error term.

The regression models in equations (1) and (2) are estimated using ordinary least squares regressions (OLS). We estimate the models separately for 1994 and 1998. The standard errors are corrected for heteroscedasticity (White, 1980).

#### 6. EMPIRICAL RESULTS AND DISCUSSION

Table 3 reports the results of the estimation model examining the relationship between board composition and firm risk taking. For the 1994 sample, we find a statistically positive relationship between the two variables both for the full model (model 3) and the model (model 2) where we estimate the relationship without any control variables. This findings only relate to risk taking measured by the proxy for stock returns risk. When firm enterprise is

<sup>&</sup>lt;sup>71</sup> The classifications were chemical, construction, energy, food, forest, investment, media, metal and engineering, multi-business, trade, transport, telecom and electronics, other industries, and other services. We combined industries into three based on relatedness of the industries, taking into consideration the GICS.

	Standard deviation ROE		E	Beta		Capital intensity	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	
Intercept							
1998	8.37***	9.36***	0.96***	0.25	0.51	8.22**	
	(0.01)	(0.07)	(0.00)	(0.51)	(0.40)	(0.04)	
1994	13.64***	30.88	0.60***	-0.18	1.81***	10.71**	
	(0.01)	(0.21)	(0.00)	(0.45)	(0.01)	(0.03)	
OU							
1998	-2.93	-1.21	-0.19	-0.23	1.69*	0.63	
	(0.40)	(0.56)	(0.18)	(0.39)	(0.08)	(0.54)	
1994	1.69	3.26	0.30**	0.26**	0.19	-0.37	
	(0.82)	(0.68)	(0.03)	(0.04)	(0.83)	(0.73)	
BCEO							
1998		0.14		-0.08		-0.07**	
		(0.27)		(0.26)		(0.03)	
1994		-0.26		0.01		-0.06*	
		(0.28)		(0.11)		(0.07)	
FS							
1998		-0.53		-0.81***		-0.06*	
		(0.31)		(0.01)		(0.06)	
1994		-2.39		0.10***		-0.89*	
		(0.37)		(0.00)		(0.06)	
$d^1$							
1998		2.68		0.03		-1.20**	
		(0.27)		(0.76)		(0.04)	
1994		1.12		0.02		-1.66**	
		(0.84)		(0.80)		(0.05)	
$d^2$							
1998		-0.12		0.44		-3.02***	
		(0.94)		(0.32)		(0.01)	
1994		1.66		-0.05		-2.75**	
		(0.83)		(0.66)		(0.02)	
Adjusted R <sup>2</sup>							
1998	0.01	0.12	0.02	0.02	0.03	0.16	
1994	0.02	0.07	0.08	0.23	0.02	0.22	
Correlation							
1998		-0.12		-0.80		0.13	
1994		-0.26		0.15		-0.16	

#### Table 3. Effect of board composition on firm risk taking measures

Regression analyses, model:

 $Risk_{i,t}^{z} = \alpha_{0}^{z} + \alpha_{1}^{z}BS_{i,t} + \alpha_{2}^{z}FS_{i,t} + \alpha_{4}^{z}BCEO_{i,t} + \alpha_{4}^{z}d_{i,t}^{1} + \alpha_{5}^{z}d_{i,t}^{2} + \varepsilon_{i,t}$ , where  $Risk_{i,t}^{z}$  is risk taking measure for firm *i* at time *t* where z=1 for standard deviation of return on equity, z=2 for capital intensity, and z=3 for beta,  $BS_{i,t}$  is board size for firm *i* at time *t*,  $FS_{i,t}$  is firm size (logarithm of total assets) for firm *i* at time *t*,  $BCEO_{i,t}$  is the percentage of equity held by the board and the CEO for firm *i* at time *t*,  $d_{i,t}^{1}$  is firm *i* classified in industry 1 at time *t*,  $d_{i,t}^{2}$  is firm *i* classified in industry 3 at time *t*,  $\varepsilon_{i,t}$  error term for firm *i* at time *t*. Probability values are in parentheses: \*\*\* significance at 1%; \*\* significance at 5%; \*significance at 10%. In model 1, we estimate equation (1) without control variables. We estimate the full model given in equation (1) in model (2). The correlation results reported related to the correlations between board composition and the 3 firm risk taking variables.

	Standard deviation ROE		Beta		Capital intensity	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Intercept						
1998	0.23	8.69	-0.86	-1.32	0.22	5.24*
	0.89	0.26	0.43	0.35	0.89	0.09
1994	47.35**	47.65*	0.48	0.27	4.3	10.34**
	0.04	0.08	0.14	0.32	0.13	0.03
BS						
1998	0.85	-0.46	0.89	1.17	0.84	0.4
	0.36	0.91	0.18	0.33	0.35	0.76
1994	-17.85	-13.63	0.17	-0.2	-1.27	0.05
	0.11	0.12	0.33	0.41	0.36	0.96
BCEO	-	-		-		
1998		0.15		-0.003		-0.07**
		0.25		0.73		0.02
1994		-0.28		0.007		-0.06**
1777		0.25		0.25		0.05
FS		0.20		0.20		0.02
1998		-0.48		0.05		-0.67**
1770		0.38		0.05		0.05
1994		0.97		0.11***		-0.88**
1777		0.67		0.01		0.05
d <sup>1</sup>		0.07		0.01		0.05
1998		2.86		-0.55		0.97**
1998		0.16		0.35		0.97
1994		-0.04		0.33		-1.64**
1994		-0.04 0.99		0.03		0.05
$d^2$		0.99		0.78		0.05
a 1998		0.16		-0.61		2.9***
1770		0.16		0.3		
1004						0.01
1994		0.25		-0.08		-2.7
<u> 11. 1 D C</u>		0.97		0.5		0.02
Adjusted R-Square	0.01				0.01	0.14
1998	0.01	0.12	0.02	0.02	0.01	0.16
1994	0.05	0.04	0.01	0.16	0.004	0.22
Correlation						
1998	-0.12		-0.04			0.05
1994	0.03		0.34			0.02

Table 4. Effect of board size on fir	<b>III I ISK LAKINY IIIEASULES</b>

Regression analyses, model:

 $Risk_{i,t}^{z} = \alpha_{0}^{z} + \alpha_{1}^{z}BS_{i,t} + \alpha_{2}^{z}FS_{i,t} + \alpha_{4}^{z}BCEO_{i,t} + \alpha_{4}^{z}d_{i,t}^{1} + \alpha_{5}^{z}d_{i,t}^{2} + \varepsilon_{i,t}$ , where  $Risk_{i,t}^{z}$  is risk taking measure for firm *i* at time *t* where *z*=1 for standard deviation of return on equity, *z*=2 for capital intensity, and *z*=3 for beta,  $BS_{i,t}$  is board size for firm *i* at time *t*,  $FS_{i,t}$  is firm size (logarithm of total assets) for firm *i* at time *t*,  $BCEO_{i,t}$  is the percentage of equity held by the board and the CEO for firm *i* at time *t*,  $d_{i,t}^{1}$  is firm *i* classified in industry 1 at time *t*,  $d_{i,t}^{2}$  is firm *i* classified in industry 3 at time *t*,  $\varepsilon_{i,t}$  error term for firm *i* at time *t*. Probability values are in parentheses: \*\*\* significance at 1%; \*\* significance at 5%; \*significance at 10%. In model 1, we estimate equation (1) without control variables. We estimate the full model given in equation (1) in model (2). The correlation results reported related to the correlations between board size and the 3 firm risk taking variables

measured by proxies for capital intensity and income instability enterprise variables, the results suggest that board composition do not exert any significant influence on firm risk taking. The results for the 1998 sample, on the other hand only finds a statistically positive relationship when the relevant risk taking measure is capital intensity. We also conduct a non parametric test to examine the relationship between the board of directors and firm enterprise. As reported in Table 3, the levels of correlation between the firm risk taking and board composition are relatively low, with the exception of the correlation between beta and board composition for the 1998 sample, with inconsistent signs. Table 3 also reports results for the control variables. The results show firm size is negatively related to firm enterprise. The sign and statistical significance of the industry effects are shown to be inconsistent. The results also show a statistically negative relationship between board and CEO equity ownership and enterprise activities.

Jensen's (1993) proposition lends some support to the weak empirical findings reported in this section. Jensen (1993) suggests that the factors that motivate board of directors to take actions that create value are, in general, inadequate. Two focal points of Jensen's (1993) proposition could be discerned. The first is the threat of legal liabilities, like class action suits initiated by shareholders (lawsuits which are often activated by unexpected declines in stock price), often faced by board members could make them act to cover their interest. These legal liabilities are more often consistent with minimizing downside risk rather than maximizing value. Anecdotal evidence of Jensen's proposition is the lawsuit filed against the nonexecutive directors of Equity Life Insurers, United Kingdom. The outside directors were sued for 3.3 billion euros under the claim that the directors failed to protect policyholders. On 17 October 2003, an English High Court ruled that the case could go ahead. Furthermore, threats of adverse publicity from the media or political and regulatory authorities also serve to inhibit actions by the board to enhance value through risk taking, although desirable.

Table 4 reports the results of the model examining the relationship between board size and firm risk taking. It could be seen from Table 4 that board size exerts no measurable influence on firm enterprise. This result holds for both the parametric and non-parametric analysis. Similar to the results reported for equation (1), we find that board and CEO equity ownership and firm size relate negatively to firm risk taking.

In related studies, Yermack (1996), Conyon et al. (1998), and Eisenberg et al. (1998) all find a negative relationship between board size and firm performance.

#### CONCLUSION

The separation of ownership and control in publicly owned firms induces potential conflicts between the interests of professional managers and stockholders. The board of directors, at the apex of the decision control system in organizations is assigned the task of checking self-serving and value-decreasing behaviors by managers. Many have recommended a greater representation by outside directors, on the board of directors of listed firms, as a means of minimizing agency cost. A critical issue involved in this recommendation is the impact of board composition on increasing the wealth of the firm or, simply, enterprise activities. This issue is empirically examined in this chapter. Also investigated in this chapter is the impact of board size on firm risk taking. The relationship between board size and firm

The evidence produced in this chapter provides some support to the assertion that board composition has an impact on firm risk taking. The impact, however, depends on the risk-taking measure. Outsider-dominated boards have positive and measurable effect on strategic risk or stock return risk but not income variability, albeit weak. These effects are seen across different time periods and suggest a changing nature of firm risk taking. This suggests the need for shareholders to reconsider their approach to improving firm performance with respect to risk-taking. Rather than simply increasing the representation of independent members on a board to aid and check management, shareholders need to explore ways to enhance the motivation of outside board members to foster their commitment to wealth creation or enterprise activities of firms. The empirical examination does not lend any support to the hypothesis that there is a relationship between board size and firm enterprise activities. Lipton and Lorch's (1992) proposition does not hold in our analysis.

The positive relationship documented here confirms the recommendations of the Cadbury (1992) and Hampel (1998) reports of increasing the representations of independent board members. That is, a board dominated by independent members delivers value for shareholders.

#### REFERENCES

- Baysinger, B. D. and H. N. Butler (1985). Corporate governance and the board of directors: Performance effects of changes in board composition, *Journal of Law, Economics, and Organisation* 1, pp. 101-124.
- Baysinger, B. D. and R. E. Hoskisson (1990). The composition of boards of directors and strategic control: effects on corporate strategy, *Academy of Management Review* 15, vol. 1, pp. 72-87.
- Bowman, E. H. (1980). A risk/return paradox for strategic management. *Sloan Management Review* (Spring), 17-31.
- Bowman, E. H. (1982). Risk seeking by troubled firms. *Sloan Management Review*, 23 (4), 33-42.
- Brealey, R., S. Myers (2006). Principles of corporate finance (3d ed.). New York: McGraw-Hill.
- Brickley, J. A, J. Coles, and R. Terry (1994). Outside directors and the adoption of poison pills, *Journal of Financial Economics* 35, pp. 317-332.
- Byrd, J. and K. Hickman (1992) Do outside directors monitor management? Evidence from tender offer bids, *Journal of Financial Economics* 32, pp. 195-222.
- Cadbury, A. (1992). Report of the committee on the financial aspects of corporate governance. London: Gee Publishing
- Conyon, M. J. and S. I. Peck (1998). Board size and corporate performance: evidence from European countries, *The European Journal of Finance* 4, pp. 291-304.
- Demski, J. S. (1987). Managerial accounting, *The executive course*, ed. G.E Germane, Addison-Wesley, pp. 101-122.
- Eisenberg, T. (1995). Konkurs eller reconstruction. SNS Förlag, Stockholm.

- Eisenberg, T., S. Sundgren, and M. T. Wells (1998) Large board size and decreasing firm value in small firms, *Journal of Financial Economics* 48, pp. 35-54.
- Eisenhardt, K. (1985). Control: Organisational and economic approaches, *Management Science* 31, pp. 134-149.
- Fama, E. F. and M. Jensen (1983). Separation of ownership and control, *Journal of Law and Economics* 26, pp. 301-325.
- Fischel, D., and M. Bradley (1986). The role of liability rules and derivative suit in corporate law: A theoretical and empirical analysis, *Cornell Law Review* 71, pp. 261-283.
- Gilson, R. and M. Roe (1993). Understanding the Japanese keiretsu: overlaps between corporate governance and industrial organisation. *Yale Law Journal* 102, 871-906.
- Gilson, S. C. (1990). Bankruptcy, boards, banks, and blockholders, *Journal of Financial Economics* 27, pp. 355-387.
- Gupta, A. K. (1987). SBU strategies, corporate-SBU relations, and SBU effectiveness in strategy implementation, *Academy of Management Journal* 30, pp. 477-500.
- Hackman, J. R. (1990). Groups that work Jossey-Bass, San Francisco. CA.
- Hampel, R. (1998). Committee on Corporate Governance: Final Report. London: Gee Publishing
- Holthausen, R. and D. F. Larcker (1993). Boards of directors, ownership structure and CEO compensation, Unpublished manuscript, Wharton School, University of Pennsylvania, Philadelphia, PA.
- Hoskisson, R. M. Hitt, R. Johnson, and W. Grossman (1995). The effects of internal governance and ownership control on corporate entrepreneurship. University of Texas AandM, mimeo.
- Jensen, M. C. (1993). The modern industrial revolution, exit, and the failure of internal control system', *Journal of Finance* 3, pp. 831-877.
- John, K. and L. W. Senbet (1998). Corporate governance and board effectiveness, *Journal of Banking and Finance* 22, pp. 371-403.
- Lev, B. (1974). On the association between operating leverage and risk. *Journal of Financial and Quantitative Analysis* 9, 627-642.
- Lipton M. and J. W. Lorsch (1992). A modest proposal for improved corporate governance, *Business Lawyer* 48, vol.1, pp. 59-77.
- March, J. G. and Z. Shapira (1987). Managerial perspectives on risk and risk taking, *Management Science* 33 (11), 1404-1414.
- Mayer, C. (1997). Corporate governance, competition and performance, *Journal of Law and Society* 24, pp. 152-176.
- Miller, K. D. and P. Bromiley (1990). Strategic risk and corporate performance: An analysis of alternative risk measures, *Academy of Management Journal* 33, pp. 756-779.
- Mintzberg, H., D. Raisinghani, and A. Theoret (1976). The structure of "unstructured" decision processes, *Administrative Science Quarterly* 21, pp. 246-275.
- Roe, M. (1994). Strong mergers, weak owners: the political roots of American corporate *finance*. Princeton University Press, Princeton, NJ.
- Rosenstein, S. and J. G. Wyatt (1990). Outside directors, board independence and shareholder wealth, *Journal of Financial Economics* 26, pp. 175-191.
- Shapiro, A. C. and S. Titman (1986). An integrated approach to corporate risk measurement. In J. M. Stern and D.H Chew (Eds.), *The revolution in corporate finance*, Oxford: Blackwell, 215-229.

- Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19 (3), 425-442.
- Short, H., K. Keasey, M. Wright, and A. Hull (1999). Corporate governance: from accountability to enterprise, *Accounting and Business Research* 29 (4), pp. 337-352.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, pp. 817-838.
- Wright, P., S. P. Ferris, A. Sarin and V. Awasthi (1996). Impact of corporate insider, blockholder, and institutional ownership on firm risk taking, *Academy of Management Journal* 39 (2), pp. 441-463.
- Yermack, D. (1996). Higher market valuation of companies with a small board of directors, *Journal of Financial Economics* 40, pp. 185-211.
- Zahra, S. (1996). Governance, ownership and corporate entrepreneurship: the moderating impact of industry technological opportunities, *Academy of Management Journal* 39 (6), pp. 1713-1735.

Chapter 50

## **EXPLORING GENDER DIFFERENCES IN DECISION-MAKING USING THE IOWA GAMBLING TASK<sup>\*</sup>**

## Ruud van den Bos<sup>†</sup>, Esther den Heijer, Suzanne Vlaar and Bart Houx

Ethology & Welfare, Faculty of Veterinary Sciences, Utrecht University, Yalelaan 2, NL-3584 CM Utrecht, the Netherlands

#### ABSTRACT

Human and non-human animals essentially face the same problem: how to find the best long-term option within an environment that contains uncertainty for relevant items. Against this background the Iowa Gambling Task is a biologically relevant task to study such decision-making processes. We have recently developed an animal analogue of this task in rodents. An interesting cross-species finding in this task is that performance differences exist between males and females: while males tend to focus exclusively on long-term goals, females tend to balance short- and long-term interests, in other words males shift from exploration to exploitation, while females remain exploratory. In this chapter we try to answer the question what may underlie these differences between males and females, focussing thereby on humans. First, we discuss a neurobehavioural model for the Iowa Gambling Task. Subsequently, we look at the contribution of the menstrual cycle using both data from the literature and an experiment that we conducted. We conclude that the menstrual cycle is not a decisive factor for these differences to occur. Finally, we discuss the possibility that differences in choice behaviour may be due to differences in the general dynamics of neurotransmitter systems. Based on recent experiments we conclude that differences in brain serotonergic and dopaminergic activity may contribute to the observed behavioural differences.

<sup>\*</sup> A version of this chapter was also published in *Psychology of Decision Making in Education, Behavior and High Risk Situations*, edited by Jeanine A. Elsworth published by Nova Science Publishers, Inc. It was submitted for appropriate modifications in an effort to encourage wider dissemination of research.

<sup>&</sup>lt;sup>†</sup> Correspondence to: Ethology & Welfare, Faculty of Veterinary Sciences, Utrecht University, Yalelaan 2, NL-3584 CM Utrecht, the NetherlandsTel: ++31-30-2534373; fax: ++31-30-2539227; e-mail: r.vandenbos@vet.uu.nl

#### INTRODUCTION

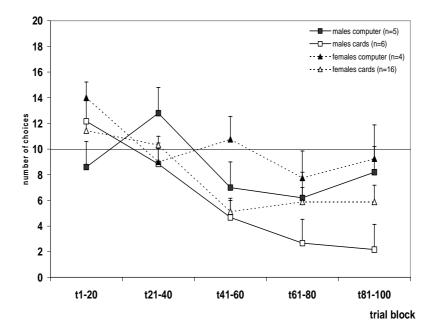
The Iowa Gambling Task (IGT) was originally developed as tool to diagnose patients suffering from lesions in the ventromedial prefrontal cortex (VMF; Bechara et al., 1994). Patients with such lesions are characterised in their everyday life by the absence of establishing profitable long-term behavioural strategies in both social and non-social affairs (Damasio, 1994). The Iowa Gambling Task was meant to model the development of everyday life long-term profitable strategies. Test subjects have to develop a long-term profitable monetary scenario in a situation of uncertainty and a conflict between the chance of encountering an immediate large reward (100\$) in two long-term losing decks (A and B; -250\$ per 10 cards) and the chance of encountering an immediate small reward (50\$) in two long-term winning decks (C and D; +250\$ per 10 cards; Bechara et al., 1994). Normal subjects develop a preference for the so called 'good' or advantageous decks C and D over the course of the experiment, while VMF patients do not (Bechara et al., 1994).

While much effort has been devoted to substantiate and characterize the effects observed in VMF patients using physiological measurements, such as skin conductance responses, and by comparing the test results of VMF patients with those of patients suffering from other kinds of brain lesions (Bechara et al., 1997, 1998, 1999, 2000; Clark et al., 2003; Manes et al., 2002; Tranel et al., 2000, 2002), little effort has been devoted to put the test in a biological context. Yet, this test may be positioned in the context of animal research that studies similar phenomena, such as optimal foraging and risk sensitive foraging: how to find the best longterm strategy in an environment in which costs and benefits vary at different sites and at different times (Van den Bos, 2004).

Inglis and colleagues (1997, 2001) have described a model of animal exploration and foraging behaviour in an environment containing uncertainty, i.e. the information primacy model, which in essence may capture the problem underlying the Iowa Gambling Task. In short, uncertainty in environments arises because of three factors: fluctuating patch quality (broadly taken as overall cost-benefit relationships), location specificity (broadly taken as spatio-temporal stimuli characterising a location), and recency of information (broadly taken as 'when was I here last time'). To reduce uncertainty to zero – Has a site changed in quality? Have stimuli characterising a site changed? Have I been here before today? - an individual should be everywhere anytime. However, as this is physically impossible, and runs right against fulfilling physiological needs, such as obtaining enough food, daytime ongoing behaviour is a compromise between 'information gathering' and 'satisfying physiological needs': when hungry go to the best sites, i.e. the easiest to find sites with the lowest costbenefit ratios, when not, or less, hungry, explore - update information - and eat while on the way. All information gathered on the way might be useful at a later stage (De Valois, 1954). So, in order to be successful the organism has to build a cognitive and affective map of its surroundings which contains the 'when, where, what' and 'what's the value' features of environmental stimuli respectively.

The Iowa Gambling task may be seen as an environment that contains uncertainty and through exploration the individual has to decipher what this environment looks like. The order of win and loss cards per deck differs per block of 10 cards and no information is present which indicates that decks will remain the same throughout the game. The individual has to pay attention to the amount of money he or she has (cf. a physiological need), and create and update an 'affective' and a 'cognitive' map.

In this chapter we explore differences between men and women in IGT performance. For, such differences have been reported in the literature (Bechara and Martin, 2004; Bolla et al., 2004; Reavis and Overman, 2001; Overman, 2004; Overman et al., 2004, 2006). Figure 1 shows the results of our own studies (data from Van den Bos, 2004; Van den Bos et al., 2006a) which are in line with those from the literature: regardless of the specific experimental set up, it turns out that female subjects take more cards from the disadvantageous decks than male subjects, especially in the second half of the task. One could therefore argue that male subjects shift their behaviour from exploration to exploitation while female subjects remain exploratory. In an animal analogue of this task we observed similar gender differences in mice (Van den Bos et al., 2006b) and rats (unpublished data; see later sections of this chapter).



Choices from disadvantageous decks

Figure 1. IGT test performance in men and women expressed as the number of cards taken from decks A and B (AB) per block of 20 choices (mean+SEM). The data are taken from two experiments (Van den Bos, 2004; Van den Bos et al., 2006a). Differences existed between the two experiments. In the cards-experiment (Van den Bos, 2004) subjects were subjected to the original experiment using cards as described by Bechara et al. (1994). In the computer-experiment (Van den Bos et al., 2006a) subjects were subjected to a computerised version. Subjects in the first experiment were older (average age±SEM:  $36.3\pm2.4$  yr) than subjects in the second experiment (average age±SEM:  $27.2\pm3.0$  yr). Statistical analysis revealed that regardless of gender or experiment subjects chose fewer and fewer AB cards as the IGT progressed (three-way ANOVA with gender and experiment as independent factors and trial-block as repeated measurement: F(4,108)=9.978,  $p\leq0.001$ ). Regardless of gender subjects chose more AB cards than men (F(1,27)=4.622,  $p\leq0.041$ ), whereas regardless of gender subjects chose more AB cards in the computer than in the cards version (F(1,27)=10.670,  $p\leq0.003$ ). The latter may be age- rather than version-related (see Crone & Van der Molen, 2004).

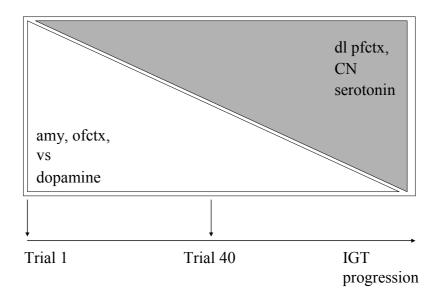
In the following sections we shall explore what may underlie these differences. First, we will present a neurobehavioural model of the performance in the IGT. Subsequently we will discuss as well as present data that the menstrual cycle may not contribute to the observed differences. We will then turn to the question what exactly may characterize the performance differences between males and females using a recently developed formal mathematical model to analyse IGT behaviour. We will conclude that men and women may engender different strategies to approach problems such as presented in the IGT and that differences in neurotransmitter activity may explain differences in performance between men and women.

#### THE IOWA GAMBLING TASK: A NEUROBEHAVIOURAL MODEL

Thus far studies have shown that the amygdala and ventromedial prefrontal cortex are important structures in developing a choice for the long-term winning decks C and D (Bechara et al., 1997, 1998, 1999). It has recently been shown that reducing dopamine levels in healthy volunteers by a mixture containing the branched-chain amino acids (BCAA) valine, isoleucine and leucine led to more choices for the disadvantageous decks (Sevy et al., 2006). Earlier findings suggested that the dopaminergic system may be important in the early stages of the test (Bechara et al., 2001), i.e. when representations may be formed of the different options (Sevy et al., 2006). As discussed by others this dopaminergic activity may also be related to exploring the different decks and responding to rewards (Fiorillo et al., 2003). The ventral striatum may also be critically involved herein (Knutson et al., 2001).

A serotonergic component appears to be present in the later stages of the test (Bechara et al., 2001). This serotonergic activity may be related to regulating the extent to which individuals continue or maintain their choice behaviour for the best long-term option, as opposed to being tempted to visit the long-term losing decks that contain immediate large rewards, i.e. self-control, as serotonin has been implicated in controlling levels of impulsive behaviour (Higley et al., 1996a,b; Mehlman et al., 1994, 1995) or controlling the extent to which mesolimbic dopaminergic activity gains control over behaviour (Katz, 1999). Recent studies in female rats in which serotonin levels have been elevated in the brain by genetically affecting serotonin-reuptake are in line with these findings (Homberg et al., 2007; see later sections of this chapter). The dorsolateral prefrontal cortex may be critically involved in this later stage of the test as well (Ernst et al., 2002; McClure et al., 2004; Ridderinkhof et al., 2004).

The differential involvement of these neurotransmitter systems suggests a transition from one set of neuronal structures to another as the test progresses, i.e. from the reward system, necessary for learning the best long-term option by assessing and integrating trial-by-trial wins and losses, to a cognitive control system, aiding in maintaining to choose once chosen options for which pay-off lies ahead in the future (Figure 2; Bechara & Damasio, 2002; Ernst et al., 2002; McClure et al., 2004; Ridderinkhof et al., 2004; Tanaka et al., 2004; Tranel et al., 2000; Shizgal & Arvanitogiannis, 2003). McClure and colleagues (2004) have demonstrated that the activity of this cognitive control system (delta system in their nomenclature) is stronger than the activity of the reward system (beta system) when decisions are made for options where pay-off lies ahead in the future, while the beta system has a slightly higher activity than the delta system when decisions are made for which pay-off is immediate. The IGT contains a conflict between immediate and long-term pay-off. It has been shown that individual differences exist in the extent to which subjects are able to withhold responding to the immediate pay-offs of the disadvantageous decks (Bechara & Damasio, 2002; Van den Bos, 2004; Crone et al., 2004).



# Neurobehavioural Model of the IGT

Figure 2. Neurobehavioural model of the IGT. The horizontal axis represents the progression of the task. The upper and lower triangle represent the relative contribution of the different brain systems which may be involved in the different stages of the test, learning the relevant task-features and choosing cards from the long-term winning decks (see text for further explanation). Abbreviations: amy: amygdala; ofctx: orbitofrontal cortex; vs: ventral striatum; dl pfctx: dorsolateral prefrontal cortex; CN: caudate nucleus.

We have not as yet validated all aspects of the model (Van den Bos et al., 2006a; Van den Bos & De Ridder, 2006). However, we do have data supporting features of the model.

Firstly, we conducted a recent study on the relationship between dieting and self-control (Kuijer et al., 2007). This study revealed that self-control moderated IGT choice behaviour of women scoring high or low on food restraint in the second half of the test (trials 41-100) but not the first half (trials 1-40). Women scoring high on restraint with low self-reported self-control took more cards from the disadvantageous decks in the second half of the task than women scoring high on restraint who also scored high on self-reported self-control (Kuijer et al., 2007). These data are in line with the model, as they show that self-control plays an important role in the second but not the first half of the task.

Secondly, data from a recent student-practical experiment revealed that stimulating elements of the reward system in the second half of the task disrupted task performance (Van den Bos & Houx, unpublished data). In this experiment we presented young male subjects (average age  $\pm$ SEM: 22.4 $\pm$ 0.3 yr, n=90 subjects) at trials 41, 61 and 81, i.e. during the second half of the task, blocks of three pictures each consisting of beautiful ('hot') or average ('not') women. As control pictures, pictures of flowers were used. Participants were awarded 3 Euro

upon completion of the task. In addition they received 1 Euro for every 1000 Euro they earned in the task. We used the computerised form of the standard IGT (for details see Bechara et al., 1994; Van den Bos et al., 2006a). It has been reported that pictures of beautiful women enhance activity in the ventral striatum of men, which is related to their willingness to exert effort for viewing them (Aharon et al., 2001) and may be related to an activated mesolimbic ventral striatal dopamine system, which subserves willingness to exert effort, even up to the point of risk-taking behaviour or tolerating negative cost-benefit outcomes (Cardinal et al., 2001; Fiorillo et al., 2003; Matthews et al., 2004; Richards et al., 1999; Salamone & Correa, 2002; Spruijt et al., 2001). So, we predicted that pictures of beautiful women may compromise men's monetary decision-making, as these pictures weaken cognitive control, needed to maintain a long-term perspective, by activating the ventral striatum and increase the temptation to engage in less than optimal options. In a recent study such an effect was observed using a temporal discounting paradigm (Wilson & Daly, 2004). Men settled for a smaller immediate amount of money after viewing pictures of beautiful women than after viewing pictures of average women. It would appear that they tolerate a less than optimal cost-benefit balance. The results of our study are shown in Figure 3. The men clearly rated the pictures of beautiful women higher than those of average women, which is in line with literature findings (Figure 3, panel A; cf. Wilson & Daly, 2004). No differences were observed with respect to viewing times for the different pictures, suggesting that IGT effects were not related to attentional differences for the different classes of pictures or mnemonic differences for the different classes of pictures due to the fact that the task was interrupted (Figure 3, panel B). The data from the IGT showed that viewing pictures of beautiful women slightly increased risk-taking behaviour compared to viewing pictures of average women or flowers (Figure 3, panel C). It is clear that the effect was small, which is partly due to the fact that the range over which disturbances may be found in this age group is small, given that the performance after 100 trials is still poor in males as also shown in Figure 1 (cf. Crone & Van der Molen, 2004). Thus any factor interfering with their performance may mask the results. During the student-practicals many disturbing factors were present. Indeed, a replication of this study in another student-practical may have failed for this reason. Nevertheless, the present data are in line with the prediction of the model.

## THE MENSTRUAL CYCLE AS A FACTOR IN THE IOWA GAMBLING TASK

So, if the model holds, one obvious explanation for the difference between men and women therefore would be the way these brain systems operate in men and women. As dopaminergic and serotonergic activity fluctuate across the menstrual/oestrus cycle in humans and animals (Becker, 1999; Dluzen & Ramirez, 1985; Fernández-Ruiz et al., 1991; Ho et al., 2001; Justice & De Wit, 1999; Nördstrom et al., 1998; Rubinow et al., 1998; Shimizu & Bray, 1993; Thompson & Moss, 1997) we hypothesised that the performance of women may depend on the phase of their menstrual cycle.



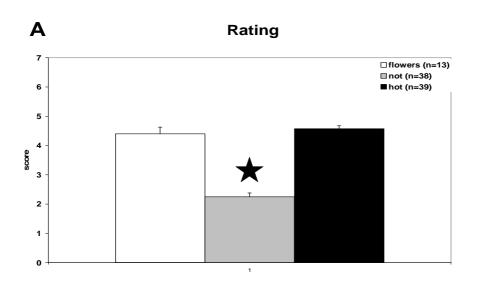


Figure 3, panel A: Mean (+SEM) rating scores of pictures offered during the second half of the IGT (scale 1-7; see text). Pictures were taken from the 'hot/not' internet site (cf. Daly & Wilson, 2004). Beautiful women ('hot') were ranked >9 at this site, average ('not') women <5 (scale:1-10). One-way ANOVA: F(2,87)=106.077,  $p\leq0.001$ ; \*:  $p\leq0.05$ , significantly different from the other two groups (Student-Newman Keuls test).

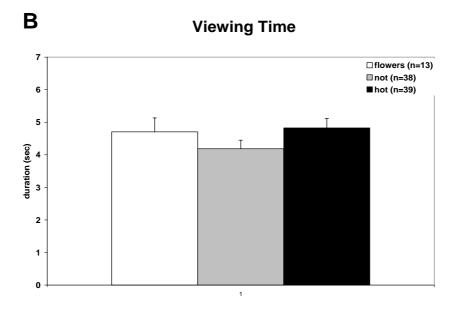


Figure 3, panel B: Mean (+SEM) picture viewing times of pictures offered during the second half of the IGT (see text). One-way ANOVA: F(2,87)=1.453, NS.



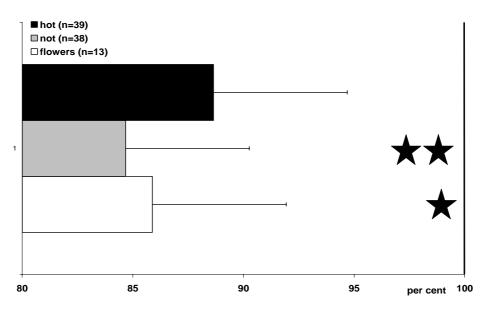


Figure 3, panel C: Mean (+SEM) per cent of choices from the disadvantageous decks during trials that were intermitted by pictures of hot- and not-women or flowers (trial 41-100) compared to baseline (trial 1-40). Statistics were paired t-tests (hot: t=1.879, df=38, p≤0.07; not: t=2.742, df=37, p≤0.009; flowers: t=2.335, df=12, p≤0.038). Initially data were expressed as the number of choices from the disadvantageous decks per block of 20 trials. We noted that the hot-group did not improve its performance from blocks 1 and 2 to blocks 3, 4 and 5. The not-group did improve its performance but more weakly than expected which may have been due to the presentation of the pictures. This masked the difference between the hot- and not-group. Therefore we decided to analyse the data using a within-subject analysis. We compared the average scores of blocks 1 and 2 with the average scores of block 3, 4 and 5. This is shown in the figure.

Data of one recent study suggested that the menstrual cycle has no effect on choosing cards from disadvantageous decks in the IGT (Reavis & Overman, 2001). However, the absence of an effect may depend on the magnitude of reward used. We showed that by increasing the differences between the immediate rewards of decks A and B versus decks C and D, risk-taking behaviour may be enhanced which could be due to a prolonged or increased dopaminergic activity (Van den Bos et al., 2006a). Female subjects are more sensitive to stimulation of dopaminergic activity in the follicular than the luteal phase of the menstrual cycle (Justice & De Wit, 1999; White et al., 2002). We therefore tested female subjects in the follicular and luteal phase in a within design using the IGT with increased reward magnitude differences (Van den Bos et al., 2006a). We hypothesised that subjects would make more choices for disadvantageous options in the follicular than the luteal phase of the menstrual cycle. We also tested subjects in a newly developed 'speed IGT' in order to explore the effect of a different type of reward. We reasoned that increased or decreased computer speed as reward or punishment would affect the subjects more directly than the more indirect wins and losses of abstract money. We expected that this 'speed IGT' would be more sensitive to potential effects of the menstrual cycle. We hypothesised that subjects would make more choices for disadvantageous options in the follicular than the luteal phase of the menstrual cycle.

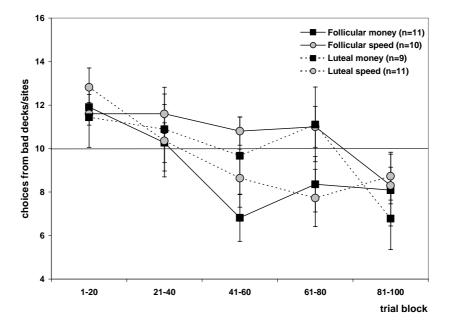
Forty-one female subjects (average  $age\pm SEM$ :  $21.5\pm0.5$  yr) fulfilling strict criteria with respect to cycle irregularities, mood disorders and drug use, were tested in the follicular phase (until day 9 after menses) and luteal phase (day 3-8 before menses). We used two versions of the Iowa Gambling Task, i.e. the standard 'money IGT' and the newly developed 'speed IGT'.

The basic instructions of the 'speed IGT' were similar as to those of the 'money IGT'. The overall goal in the 'speed IGT' was to keep the speed of a black square (puppet) at a high level rather than earning as much money as possible as in the 'money IGT'. On a computer screen subjects saw four coloured boxes ('food patches') that were connected by a central area. At the start of each trial the puppet was present in the middle of the central area. When subjects clicked on one of the four boxes using the mouse the puppet would move towards that box. Subjects could then obtain 'good' or 'bad' food by their choice instead of winning or losing money. While 'good' food increased the puppet's speed, 'bad' food slowed it down. Upon returning to the central area the puppet would adopt its new speed as well as on its journey to the next box in the next trial. The monetary values of the original IGT were used to calculate increases and decreases in speed. The feedback that subjects obtained was the same as for the 'money IGT' in terms of absolute values, so they could see whether they had wins or losses. The puppet's speed was indicated graphically by a speedometer as well as numerically by a value to the left of the boxes. The four boxes represented the four decks of cards of the 'money IGT': upper left: A (green), upper right: B (blue), lower left: C (red) and lower right: D (orange-brown). A bar and numbers indicated how many choices remained for each box. Just as in the 'money IGT' 40 choices per box existed.

The experimental design was such that subjects either started with the 'money IGT' or the 'speed IGT' in either the follicular or the luteal phase according to a complete cross-over design. Thus they visited our institute twice. They were awarded course credits or money (6 Euro) upon completing the task.

A 4-way ANOVA [menstrual cycle phase, IGT-version, trial-block (repeated), visit-order (repeated)] showed that: (i) regardless of phase, IGT-version or visit-order, all subjects gradually decreased the number of choices from the disadvantageous options as the IGTs progressed (trial-block: F(4,40)=8.559,  $p\leq0.001$ ) and (ii) regardless of IGT-version or phase, subjects improved performance on their second IGT (visit-order: F(1,35)=5.586,  $p\leq0.024$ ). As the latter may mask phase differences, we analysed the subjects' performance separately for each visit.

Only for the first visit a menstrual cycle effect was found. The data are shown in Figure 4. It would appear that women performed better on the 'money IGT' in the follicular phase than the luteal phase while the opposite was true for the 'speed IGT', which was supported by a near significant (F(4,148)=2.284, p $\leq$ 0.063) three-way interaction term, trial-block \* phase \* IGT-version. Subsequent two-way ANOVAs showed that only for blocks 3-5 of the 'speed IGT' significant differences existed between the follicular and luteal phase (phase: F(1,19)=5.384, p $\leq$ 0.032). Furthermore, a significant trial-block effect (F(4,76)=4.095, p $\leq$ 0.005) as well as IGT-version effect (F(1,19)=6.986, p $\leq$ 0.016) was found for the follicular phase.



#### Choices of bad options

Figure 4: Mean (±SEM) number of choices from 'bad' decks/sites during the follicular and luteal phase. See text for further explanation.

In none of the analyses did the performance on the 'money IGT' differ from the performance on the 'speed IGT' per se, suggesting that the two Gambling Tasks are equivalent.

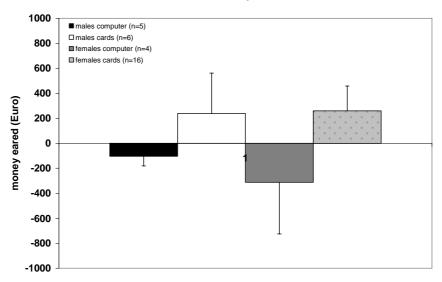
In line with earlier findings (Reavis & Overman, 2001) we did not observe clear menstrual cycle phase differences in performance on the 'money IGT', even when we increased the conflict between 'good' and 'bad' decks (Van den Bos et al., 2006a). If anything, we observed that women tended to perform more poorly on the 'speed IGT' in the follicular phase than the luteal phase. Whether this is related to differences in metabolic rate in these phases (Buffenstein et al., 1995) or differences in perceptual motor skills in these phases (Kimberly Epting & Overman, 1998) remains to be determined.

The fact that the menstrual cycle has little effect on IGT performance is supported by recent studies that suggest that gender differences in IGT performance already occur early in life, i.e. from the age of 7 years onwards (Crone et al., 2005; Overman, 2004; Overman et al., 2004), that is, well before the onset of puberty and the period in which the menstrual cycle starts.

### Differences between Males and Females Revisited: Neurobiology

Having discarded the menstrual cycle as a possible confound in the monetary IGT, we now again turn to the question what may underlie performance differences between men and women. To this end we first describe the choice behaviour of men and women in slightly more detail. The test instruction in the IGT is to win as much money as possible. Figure 5 shows the amount of money won by men and women in the experiment that was shown in Figure 1. It is clear that men and women have earned the same amount of money at the end of the game. Also when the amount of money was analysed for the different trial blocks no differences were observed (data not shown). Thus in terms of complying with the task instructions no major differences existed between men and women. The only difference is that women take more cards from the disadvantageous decks than men, and thereby take the risk of losing money, yet in return gain information on whether the disadvantageous decks may have changed or not. In fact, a few cards from the disadvantageous decks, although the choices nevertheless are risky, may lead to high immediate pay-off. As we have argued earlier (Van den Bos, 2004) the performance in the IGT may be a balance between exploration and exploitation.

To assess whether men and women are differently sensitive to wins and losses, we analysed the data in Figure 1 using the Expectancy-Valence model (EV model; Stallen, 2006), which has been shown to be useful for analysing choice behaviour in the IGT (Busemeyer & Stout, 2002; Sevy et al., 2006; Yechiam et al., 2005).



### Amount of money earned after 100 trials

Figure 5: The mean (+SEM) amount of money earned after 100 trials. A two-way ANOVA [experiment, gender] revealed no significant differences: experiment: F(1,27)=2.263, NS; gender: F(1,27)=0.093, NS; experiment\*gender: F(1,27)=0.142, NS.

In general the model seems to describe the behaviour of men better than that of women, as the percentage of subjects in which the model does not do better than a random model is higher in women than in men (women versus men: 35 % (n=20) versus 9.9% (n=11)), and the test-statistic (G<sup>2</sup> statistic; Busemeyer & Stout, 2002, page 258) that describes its usefulness is higher (but not significantly so) in men than in women for those individuals in which the EV model did better than a random model (mean ±SEM: men versus women:  $43.7\pm12.1$ 

(n=10) versus 26.8±6.0 (n=13)). This model contains three different parameters, each describing a different feature of task performance.

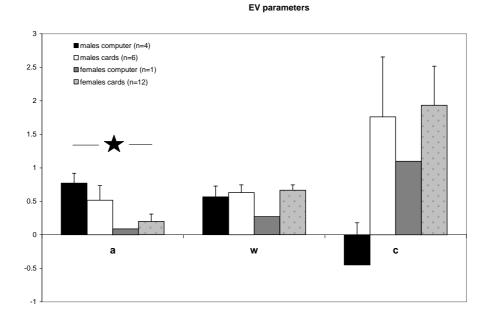


Figure 6: Mean (+SEM) values of different parameters from the Expectancy-Valence model. See text for an explanation on parameters a, w and c. Only for the learning rate parameter (a) a significant gender-difference was observed (two-way ANOVA [experiment, gender]: gender effect: F(1,19)= 3.879, p≤0.064; since in the computer version only one woman was present we also analysed the data using a one-way ANOVA with gender as factor: F(1,21)=6.316, p≤0.020); gender effect: \*p≤0.05).

Firstly, the attention to losses/wins parameter w, which measures the relative weight given to losses and wins and ranges from 0 (attention to wins only) to 1 (attention to losses only). As Figure 6 shows no differences exist between men and women on this parameter. This suggests that women and men are equally sensitive to rewards or punishments. Indeed, when the original IGT was changed such that the immediate reward conflict between decks A and B versus C and D was eliminated by having only the chance to win 50 Euro maximally in all decks, while keeping the net gains and losses per 10 cards the same, i.e. -250 Euro (AB) and +250 Euro (CD), no differences were observed between men and women in choice behaviour (re-analysis of data of Van den Bos et al., 2006a), suggesting that women and men do not differ in at least their sensitivity to losses.

Secondly, the recency parameter a, which represents the updating rate of recent information. The value of this parameter ranges from 0 to 1. A value of 0 means long associative memories, that is, information of the valence of decks is kept over long lags of choices, and thus little updating of information. A value of 1 means short associative memories, rapid forgetting and a strong urge to update information. As Figure 6 shows, men show higher values on this parameter than women. This would suggest that women retain information of the valences of different decks across longer lags than men, i.e. women have longer associative memories than men.

Thirdly, the choice consistency parameter c, which measures how consistent subjects are in their choices across trials. This parameter is related to response mechanisms. Its value ranges from -5 to +5. Negative values mean that the subject's choices are random and not led by expectancies. Positive values mean that the subject behaves according to deck expectancies. No differences were observed between men and women on this parameter.

Thus is seems that only on the recency parameter men and women differ. As already mentioned above, it was recently shown that this parameter is sensitive to manipulation with the dopaminergic system (Sevy et al., 2006). Lowering dopaminergic activity led to higher values of the parameter and a weaker performance in the IGT, which was interpreted as being due to destabilizations of representations in the orbitofrontal cortex, and thereby to more attention to recent events (Sevy et al., 2006). Combined, the present data suggest that dopaminergic activity in women may be higher than in men while their performance appears to be worse than that of men as judged by the number of cards from disadvantageous decks. One way to reconcile this apparent contradiction is to suggest that females 'emotionally' know the differences between decks, but remain, despite this knowledge, exploratory or do not yet form a routine of choosing only cards from decks C and D. That is, they do not make the switch to the other system as readily as men do. This suggests that a difference with respect to serotonergic activity or dorsolateral prefrontal cortical activity may contribute to these differences in choice behaviour. Indeed, evidence for this may be found in the literature and our own experiments.

In the rodent version of the IGT, we observed that female rats chose the 'bad' arms more often than male rats as the rat-IGT progressed, as was also observed earlier in mice (Van den Bos et al., 2006b). However, female rats lacking the serotonin reuptake transporter gene completely (homozygous rats) or partially (heterozygous rats) showed a performance similar to that of male rats, that is, they gradually started to choose the 'good' arm more often as the rat-IGT progressed (Homberg et al., 2007). Reuptake of serotonin is an important mechanism for regulating serotonergic tone (Murphy et al., 2004). Both homozygote and heterozygote rats have higher levels of serotonin release in brain structures than wild-type rats (Homberg et al., unpublished data). These data suggest that an increase of brain serotonin levels in female rats leads to a suppression of the continuing tendency to explore.

Combined, the findings in rats and humans suggest gender differences in dopaminergic and serotonergic regulation, that is, in males the balance between the two systems might be in favour of serotonin, while in females this balance might be in favour of dopamine.

From the foregoing it may be predicted that differences between men and women disappear when the dopaminergic system is strongly activated in both sexes. In an earlier published study we increased reward magnitude differences between advantageous and disadvantageous decks (AB:CD=200:50 Euro or 300:50 Euro) while keeping the overall gains and losses per 10 cards of the different decks the same (AB:-250 Euro; CD: +250 Euro) to test how sensitive the emotional system is to manipulation (Van den Bos et al., 2006a). We hypothesised that the increased preference for the disadvantageous decks and concomitant increased loss of money was due to a strong activation of the dopaminergic system. When the data were reanalysed with women and men as a separate factor it turned out that the differences between men and women that were observed in the original 100-50 IGT condition disappeared in the 200-50 and 300-50 IGT condition (re-analysis of data of Van den Bos et al., 2006a). This is in line with the hypothesis formulated above.

A study using PET-scans while subjects performed the IGT observed differences in several brain structures between male and female subjects (Bolla et al., 2004). A withinsubject analysis revealed that men activated extensive regions of the right and left lateral orbitofrontal cortex and right dorsolateral prefrontal cortex, whilst women activated only the left medial orbitofrontal cortex. Starting from these within-subject differences Overman and colleagues (2006) predicted that personal moral dilemmas, which activate the lateral dorsal prefrontal cortex, would reduce IGT performance differences between men and women. Indeed their study showed that such dilemmas either contemplated during or before the IGT led women to choose the same number of advantageous cards as men as the test progressed, i.e. during the second half of the task (Overman et al., 2006). They attributed this to an increased cognitive control due to activation of the dorsolateral prefrontal cortex. These data also support our neurobehavioural model of the IGT.

Overall, these neurobiological data suggest that women do not activate the cognitive control system to the same extent as men. From this it may be predicted that when tasks are made more difficult, i.e. when subjects have to pay more attention to the test features, leading to activation of the dorsolateral prefrontal cortex (e.g. MacDonald et al., 2000), gender differences may disappear. Indeed we observed that when task-conditions were made more difficult, differences between men and women disappeared (Vlaar, 2007), suggesting that differences between men and women actually may depend on task-conditions and thus in this sense are not fixed.

An intriguing finding that we have not yet discussed is the observation that differences also occur between men and women at the level of individual decks. These differences will be discussed in the next paragraph.

### Male-Female Differences at the Level of Individual Decks

Overman and colleagues (2006; Overman, 2004; cf. Hooper et al., 2004) noted that men and women differed with respect to the decks they choose. Within the disadvantageous decks A and B both men and women chose more cards from deck B than from deck A. However, women also chose more cards from deck B than men, which accounts for the overall performance difference. Within the advantageous decks C and D men had no preference for cards from deck C or D, while women preferred cards from deck D over cards from deck C. We also observed such differences (Figure 7), regardless of whether reward magnitude differences were increased (re-analysis of data of Van den Bos et al., 2006a). When we made the long-term differences between decks AB versus CD smaller (-100 Euro versus +100 Euro) the differences between men and women with respect to decks C and D remained (Vlaar, 2007).

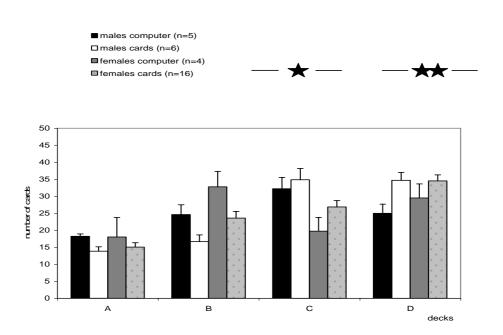


Figure 7: Mean (+SEM) number of choices from the different decks for women and men in different test-conditions. A three way ANOVA [experiment, gender, deck as repeated measure] showed a significant deck effect (F(3,81)=15.160, p≤0.001), a significant deck\*experiment effect (F(3,81)=5.052, p≤0.003) and a significant deck\*gender effect (F(3,81)=5.098, p≤0.003). Gender effects existed for deck B (two way ANOVA [gender, experiment]: F(1,27)=6.084, p≤0.02) and deck C (F(1,27)=10.668, p≤0.003). Experiment effects existed for deck B (F(1,27)=7.878, p≤0.009) and deck D (F(1,27)=6.669, p≤0.016). Gender effects: \*: p≤0.05; \*\*: p≤0.01.

The differences between decks lie in the frequency with which wins and losses occur. While deck A and C stand out by the fact that wins and losses follow one another on a regular basis (frequent loss decks), decks B and D stand out by the fact that losses only occur after long series of wins (infrequent loss decks; see scheme in Bechara et al., 1994). Thus, these gender differences seem to suggest a differential sensitivity to the way information is presented and decisions are made, which in turn may affect the overall task performance. It may thus be concluded that women have a tendency to select decks with infrequent losses and thereby may thus not perform optimally in this task. However, data of recent exepriments done in our laboratory using several variants of the IGT suggest that frequency alone is not a decisive factor for the differences between men and women (Vlaar, 2007; cf. Garon & Moore, 2007). Alternatively, therefore, the data may imply that women are more sensitive to punishment and thus avoid regular losses (Hooper et al., 2004). However, the EV model did not support such a notion as shown in Figure 6. Thus it remains to be determined why women and men are differentially sensitive to frequent or infrequent losses under some test-conditions.

Choices from decks A-D after 100 trials

### CONCLUSION

Although evidence is accumulating that differences exist between men and women in the performance on the Iowa Gambling Task, little is known about why such differences exist and what they mean. In this chapter we have explored these differences and suggested that neurobehavioural differences may explain why men and women differ in their performance. Women tend to be more flexible in the way they organise their responses: they pay attention to both long-term and short-term aspects. In contrast men appear to be more focussed on long-term goals (Overman, 2004; Overman et al., 2006). While men tend to focus exclusively on exploitation, women combine exploration and exploitation. Thus women gain and update information in return for losing some resources, while men focus solely on the latter. Whether this may be framed within an evolutionary framework remains speculative, but if so, it may be related to the differences in the roles men and women may have had in ancient times. Men may have been more goal-directed, related to for instance their hunting trips, while women were more flexible, since they were sedentary. Furthermore, men may have needed to pay more attention to energy resources than women, for instance to secure their competitive and hunting potential, and as such do not take too many risks in such decision making processes. However, such explanations remain just-so stories if they cannot be put to test. For instance therefore, it may be hypothesised that women become more goal-directed, when they are pregnant, i.e. when energy resources are more important. It is clear that more experiments are needed in human as well as non-human animals to assess how these differences emerge, what they mean and why they have evolved.

### REFERENCES

- Aharon, I., Etcoff, N., Ariely, D., Chabris, C.F., O'Connor, E. & Breiter, H.C. (2001) Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*, 32, 537-551.
- Bechara, A., Damasio, A.R., Damasio, H. & Anderson, S.W. (1994) Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7-15.
- Bechara, A., Damasio, H., Tranel, D. & Damasio, A.R. (1997) Deciding advantageously before knowing the advantageous strategy. *Science*, 275, 1293-1295.
- Bechara, A., Damasio, H., Tranel, D. & Anderson, S.W. (1998) Dissociation of working memory from decision making within the human prefrontal cortex. *The Journal* of *Neuroscience*, 18, 428-437.
- Bechara, A., Damasio, H, Damasio, A.R. & Lee, G.P. (1999) Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *The Journal of Neuroscience*, 19, 5473-5481.
- Bechara, A., Tranel, D. & Damasio, H. (2000) Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, 123, 2189-2202.
- Bechara, A., Damasio, H. & Damasio, A.R. (2001) Manipulation of dopamine and serotonin causes different effects on covert and overt decision making. Society of Neuroscience Abstracts 465.5.

- Bechara, A. & Damasio, H. (2002) Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia*, 40, 1675-1689.
- Bechara, A. & Martin, E.M. (2004) Impaired decision making related to working memory deficits in individuals with substance addictions. *Neuropsychology*, *18*, 152-162.
- Becker, J.B. (1999) Gender differences in dopaminergic function in striatum and nucleus accumbens. *Pharmacology, Biochemistry & Behavior, 64*, 803-812.
- Bolla, K.I., Eldreth, D.A., Matochik, J.A. & Cadet, J.L. (2004) Sex-related differences in a gambling task and its neurological correlates. *Cerebral Cortex*, 14, 1226-1232.
- Bos, van den R. (2004) Emotions and Cognition. In: M. Bekoff (Ed.) *Encyclopedia of Animal Behavior* (pp.554-557) Westport (CT): Greenwood Press.
- Bos, van den R., Houx, B.B. & Spruijt, B.M. (2006a) The effect of reward magnitude differences on choosing disadvantageous decks in the Iowa Gambling Task. *Biological Psychology*, 71, 155-161.
- Bos, van den R., Lasthuis, W, Heijer, den E., Harst, van der J. & Spruijt, B. (2006b) Towards a rodent model of the Iowa Gambling Task. *Behavior Research Methods*, *38*, 470-478.
- Bos, van den R. & Ridder, de D. (2006). Evolved to satisfy our immediate needs: Self-control and the rewarding properties of food. *Appetite*, 47, 24-29.
- Buffenstein, R., Poppitt, S.D., McDevitt, R.M. & Prentice, A.M. (1995) Food intake and the menstrual cycle: a retrospective analysis, with implications for appetite research. *Physiology & Behavior, 58*, 1067-1077.
- Busemeyer, J.R. & Stout, J.C. (2002) A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara Gambling Task. *Psychological Assessment*, 14, 253-262.
- Cardinal, R.N., Pennicot, D.R., Sugathapala, C.L., Robbins, T.W. & Everitt, B.J. (2001) Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science*, 292, 2499-2501.
- Clark, L., Manes, F., Antoun, N., Sahakian, B.J. & Robbins, T.W. (2003) The contributions of lesion laterality and lesion volume to decision-making impairment following frontal lobe damage. *Neuropsychologia*, 41, 1474-1483.
- Crone, E.A. & Molen, van der M.W. (2004) Developmental changes in real life decision making: performance on a gambling task previously shown to depend on the ventromedial prefrontal cortex. *Developmental Neuropsychology*, 25(3), 251-279.
- Crone, E.A., Somsen, R.J.M., Beek, van B. & Molen, van der M.W. (2004) Heart rate and skin conductance analysis of antecedents and consequences of decision making. *Psychophysiology*, 41, 531-540.
- Crone, E.A., Bunge, S.A., Latenstein, H. & Molen, van der M.W. (2005) Characterization of children's decision making: sensitivity to punishment frequency, not task complexity. *Child Neuropsychology*, 11, 245-263.
- Damasio, A.R. (1994) *Descartes' error. Emotion, reason and the human brain.* New York: Avon Books.
- Dluzen, D.E. & Ramirez, V.D. (1985) In vitro dopamine release from the rat striatum: diurnal rhythm and its modification by the estrous cycle. *Neuroendocrinology*, *41*, 97-100.
- Ernst, M., Bolla, K., Mouratidis, M., Contoreggi, C., Matochik, J.A., Kurian, V., Cadet, J-L., Kimes, A.S. & London, E.D. (2002) Decision-making in a risk-taking task : a PET study. *Neuropsychopharmacology*, 26, 682-691.

- Fernández-Ruiz, J.J., Hernandez, M.L., Demiquel, R. & Ramos J.A. (1991) Nigrostriatal and mesolimbic dopaminergic activities were modified throughout the ovarian cycle of female rats. *Journal of Neural Transmission - General section*, 85 (3), 223-229.
- Fiorillo, C.D., Tobler, P.N. & Schultz, W. (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299, 1898-1902.
- Garon, N. & Moore, C. (2007) Developmental and gender differences in future-oriented decision-making during pre-school period. *Child Neuropsychology*, 13, 46-63.
- Higley, J.D., Mehlman, P.T., Poland, R.E., Taub, D.M., Vickers, J., Suomi, S.J. & Linnoila, M. (1996a) CSF testosterone and 5-HIAA correlate with different types of aggressive behaviors. *Biological Psychiatry*, 40, 1067-1082.
- Higley, J.D., Mehlman, P.T., Higley, S.B., Fernald, B., Vickers, J., Lindell, S.O., Taub, D.M., Suomi, S.J. & Linnoila, M. (1996b) Excessive mortality in young free-ranging male nonhuman primates with low cerebrospinal fluid 5-hydroxyindoleacetic concentrations. *Archives of General Psychiatry*, 53, 537-543.
- Ho, H-P., Olsson, M., Westberg, L., Melke, J. & Eriksson, E. (2001) The serotonin reuptake inhibitor fluoxetine reduces sex steroid-related aggression in female rats: an animal model of premenstrual irritability? *Neuropsychopharmacology*, 24, 502-510.
- Homberg, J.R., Heijer, den E., Suer, R., Spruijt, B., Bos, van den R. & Cuppen, E. (2007) Modulation of decision making in the Iowa Gambling Task by partial and complete genetic ablation of the serotonin transporter in the rat and the human serotonin transporter polymorphism. Endo-Neuro-Psycho meeting 5<sup>th</sup>-8<sup>th</sup> June 2007, Doorwerth, the Netherlands.
- Hooper, C.J., Luciana, M., Conklin, H.M. & Yarger, R.S. (2004) Adolescents' performance on the Iowa Gambling Task: Implications for the development of decision making and ventromedial prefrontal cortex. *Developmental Psychology*, 40, 1148-1158.
- Inglis, I.R., Forkman, B. & Lazarus, J. (1997) Free food or earned food? A review and fuzzy model of contrafreeloading. *Animal Behaviour*, 53, 1171-1191.
- Inglis, I.R., Langton, S., Forkman, B. & Lazarus, J. (2001) An information primacy model of exploratory and foraging behaviour. *Animal Behaviour*, 62, 543-557.
- Justice, A.J.H. & Wit, de H. (1999) Acute effects of d-amphetamine during the follicular and luteal phases of the menstrual cycle in women. *Psychopharmacology*, 145, 67-75.
- Katz, L.D. (1999) Dopamine and serotonin: Integrating current affective engagement with longer-term goals. *Behavioral and Brain Sciences*, 22, 527.
- Kimberly Epting, L. & Overman, W.H. (1998) Sex-sensitive tasks in men and women: A search for performance fluctuations across the menstrual cycle. *Behavioral Neuroscience*, *112*, 1304-1317.
- Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L. & Hommer, D. (2001) Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, 12, 3683-3687
- Kuijer, R., Ridder, de D., Ouwehand, C. Houx, B. & Bos, van den R. (2007) Dieting and behavioural decision making. *Appetite* (submitted).
- MacDonald, A.W III, Cohen, J.D., Stenger, V.A. & Carter, C.S. (2000) Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*, 1835-1838.
- Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N, Aitken, M.. & Robbins, T.W. (2002) Decision-making processes following damage to the prefrontal cortex. *Brain*, 125, 624-639.

- Matthews, S.C., Simmons, A.N., Lane, S.D. & Paulus, M.P. (2004) Selective activation of the nucleus accumbens during risk-taking decision making. *Neuroreport*, 15, 2123-2127.
- McClure, S.M., Laibson, D.I., Loewenstein, G. & Cohen, J.D. (2004) Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*, 503-507.
- Mehlman, P.T., Higley, J.D., Faucher, I., Lilly, A.A., Taub, D.M., Vickers, J., Suomi, S.J. & Linnoila, M. (1994) Low CSF 5-HIAA concentrations and severe aggression and impaired impulse control in nonhuman primates. *American Journal of Psychiatry*, 151, 1485-1491.
- Mehlman, P.T., Higley, J.D., Faucher, I., Lilly, A.A., Taub, D.M., Vickers, J., Suomi, S.J. & Linnoila, M. (1995) Correlation of CSF 5-HIAA concentration with sociality and timing of emigration in free-ranging primates. *American Journal of Psychiatry*, 152, 907-913.
- Murphy, D.L, Lerner, A., Rudnick, G. & Lesch, K-P. (2004) Serotonin transporter: gene, genetic disorders, and pharmacogenetics. *Molecular Interventions*, *4*, 109-123.
- Nördstrom A., Olsson, H. & Halldin, C. (1998) A pet study of D<sub>2</sub> dopamine receptor density at different phases of the menstrual cycle. *Psychiatry Research: Neuro-imaging section*, 83, 1-6.
- Overman, W.H. (2004) Sex differences in early childhood, adolescence, and adulthood on cognitive tasks that rely on orbital prefrontal cortex. *Brain and Cognition*, *55*, 134-147.
- Overman, W.H., Frassrand, K., Ansel, S., Trawalter, S., Bies, B. & Redmond, A. (2004) Performance on the IOWA card task by adolescents and adults. *Neuropsychologia*, 42, 1838-1851.
- Overman, W., Graham, L., Redmond, A., Eubank, R., Boettcher, L., Samplawski, O. & Walsh, K. (2006) Contemplation of moral dilemmas eliminates sex differences on the Iowa gambling task. *Behavioral Neuroscience*, 120, 817-825.
- Reavis, R. & Overman, W.H. (2001) Adult sex differences on a decision-making task previously shown to depend on the orbital prefrontal cortex. *Behavioral Neuroscience*, 115, 196-206.
- Ridderinkhof, K.R., Wildenberg, van den W.P.M., Segalowitz, S.J. & Carter, C.S. (2004) Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain* and Cognition, 56, 129-140.
- Rogers, R.D., Everitt, B.J., Baldacchino, A., Blackshaw, A.J., Swainson, R., Wynne, K., Baker, N.B., Hunter, J., Carthy, T., Booker, E., London, M., Deakin, J.F.W., Sahakian, B.J. & Robbins, T.W. (1999) Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. *Neuropsychopharmacology*, 20, 322-339.
- Rubinow, D.R., Schmidt P.J. & Roca C.A. (1998) Estrogen-serotonin interactions: implications for affective regulation. *Biological Psychiatry*, 44, 839-850.
- Salamone, J.D. & Correa, M. (2002) Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural Brain Research*, 137, 3-25.
- Sevy, S., Hassoun, Y., Bechara, A., Yechiam, E., Napolitano, B., Burdick, K., Delman, H & Malhotra, A. (2006) Emotion-based decision-making in healthy subjects: short-term effects of reducing dopamine levels. *Psychopharmacology*, 188, 228-235.

- Shimizu, H. & Bray, G.A. (1993) Effects of castration, estrogen replacement and estrus cycle on monoamine metabolism in the nucleus accumbens, measured by microdialysis. *Brain Research*, 621, 200-206.
- Shizgal, P. & Arvanitogiannis, A. (2003) Gambling on dopamine. Science, 299, 1856-1858.
- Spruijt, B.M., Bos, van den R. & Pijlman, F.T.A. (2001) A concept of welfare based on reward evaluating mechanisms in the brain: anticipatory behaviour as an indicator for the state of reward systems. *Applied Animal Behaviour Science*, *72*, 145-171.
- Stallen, M. (2006). *Excel application of the Expectancy-Valence model*. Technical Report, University of Amsterdam.
- Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamaoto, Y., & Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature Neuroscience*, 7, 887-893.
- Thompson, T.L. & Moss, R.L. (1997) Modulation of mesolimbic dopaminergic activity over the rat estrous cycle. *Neuroscience Letters*, 229, 145-148.
- Tranel, D., Bechara, A. & Damasio, A.R (2000). Decision making and the somatic marker hypothesis. In : M.S. Gazzaniga (Ed). *The new cognitive neurosciences* (pp.1047-1061). Cambridge (MA): A Bradford book.
- Tranel, D., Bechara, A & Denburg, N.L. (2002) Asymmetric functional roles of right and left ventromedial prefrontal cortices in social conduct, decision-making, and emotional processing. *Cortex*, 38, 589-612.
- Valois, De R.L. (1954) The relation of different levels and kinds of motivations to variability of behaviour. *Journal of Experimental Psychology*, 47, 392-398.
- Vlaar, S. (2007) *Three new variants of the Iowa Gambling Task: deck structure effects and gender differences.* Master Thesis, Utrecht University, Utrecht, the Netherlands.
- White, T.L., Justice, A.J.H. & Wit, de H. (2002) Differential subjective effects of damphetamine by gender, hormone levels and menstrual cycle phase. *Pharmacology*, *Biochemistry & Behavior*, 73, 729-741.
- Wilson, M., & Daly, M. (2004). Do pretty women inspire men to discount the future? The Royal Society. Biology Letters 271 nr. S4, 177-179.
- Yechiam, E., Busemeyer, J.R., Stout, J.C. & Bechara, A. (2005) Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological Sciences*, 16, 973-978.

## INDEX

accessibility, xxvii, 169, 473, 483, 890, 903, 904, # 905, 907, 908, 909, 980 accidents, 201, 378, 388, 458, 827, 842, 844, 863, 3D, 761 864, 871, 876, 1087 accommodation, 266, 447, 480, 497 accountability, 54, 55, 1113 Α accounting, 87, 369, 482, 697, 714, 717, 879, 936, 1102, 1111 AARP, 942, 947, 949 accreditation, 22, 58 ABC, 497, 922, 930 acculturation, 303, 559, 568 abdomen, 588 accuracy, xxix, 29, 76, 77, 78, 84, 86, 87, 110, 191, abdominal, 480, 523 195, 284, 295, 318, 344, 351, 468, 562, 626, 904, abiotic, 1021 907, 914, 916, 917, 947, 991, 995, 1008, 1013, abnormalities, 343, 383, 388 1047, 1048, 1050, 1052, 1053, 1064, 1068, 1069, abortion, xix, xx, xxii, xxiii, 295, 351, 364, 365, 391, 392, 393, 394, 395, 396, 398, 399, 401, 402, 403, 1070 acetylcholine, 256, 258 404, 405, 406, 410, 414, 415, 416, 422, 423, 424, ACF, 633, 634 425, 426, 427, 428, 429, 430, 432, 435, 436, 437, achievement, 52, 214, 368, 458, 484, 544, 558, 560, 438, 439, 440, 441, 442, 443, 446, 448, 450, 451, 453, 503, 505, 506, 515, 571, 573, 583, 585, 592, 568, 620, 982, 983, 985, 986, 998, 1005, 1010, 1047, 1070 619, 620 achievement orientation, 1070 absolute zero, 1078 acid, 258 absorption, 94, 882 acoustic, 229, 666 abstinence, xxiii, 576, 584, 619, 620, 621, 622, 623, acquisition of knowledge, 989 624, 625, 626, 628, 629, 630, 631, 632, 633, 634, acquisitions, 1103 635 action potential, 240, 241, 245, 246, 249, 258 abuse, 267 action research. 180 abusive, 198, 372, 476, 478, 479, 482, 483, 485, 486, activation, 3, 205, 206, 207, 208, 209, 210, 211, 213, 489, 494, 497, 519, 588 214, 215, 217, 218, 220, 222, 223, 224, 226, 227, academic, 484, 1100 232, 233, 251, 255, 256, 257, 258, 904, 908, 982, academics, 24 1127, 1128, 1131, 1133 ACC, 207, 210, 211, 212, 213, 214, 215, 219, 220, activities, 264, 265, 266, 272 221, 222, 223, 224, 225, 226 acts of aggression, 82 access, xv, 120, 122, 142, 145, 158, 159, 162, 166, actual risk, xxi, 532, 537 168, 169, 173, 175, 176, 216, 282, 284, 345, 346, actuarial, 76, 79, 80, 83, 90, 91, 700 353, 354, 363, 372, 375, 392, 394, 395, 398, 405, acupuncture, 137 425, 450, 453, 460, 461, 465, 470, 480, 547, 576, acute, xv, 29, 31, 38, 42, 54, 60, 78, 89, 99, 105, 111, 578, 579, 581, 582, 585, 589, 592, 667, 668, 781, 183, 184, 187, 265, 266, 267, 474, 486, 492, 493 836, 872, 904, 922, 923, 927, 1004, 1010, 1102

- acute myelogenous leukemia, 111
- Adams, 111, 213, 231, 255, 562, 566, 1009, 1132
- adaptation, xvi, 8, 164, 187, 237, 681, 927, 955, 960, 963, 978, 1014, 1015, 1017, 1028
- addiction, xvi, 225, 227, 229, 237, 249, 250, 252, 255, 258, 259, 265, 611, 655, 1131
- adenine, 313
- adequate housing, 83
- adjustment, xviii, 154, 180, 202, 218, 311, 315, 329, 471, 754, 756
- administration, 7, 59, 241, 253, 255, 257, 470, 630, 776, 943, 947, 997, 1007, 1038
- administrative, 56, 87, 175, 449, 1002, 1051
- administrators, 85, 86, 87, 90
- adolescence, xx, 406, 407, 458, 459, 469, 471, 473, 476, 478, 482, 503, 504, 505, 515, 526, 532, 539, 568, 589, 861, 979, 1042, 1133
- adolescent problem behavior, 514
- adolescents, xx, xxi, xxiii, 18, 253, 392, 393, 394, 395, 396, 444, 453, 458, 463, 466, 470, 471, 472, 473, 475, 479, 486, 488, 500, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 537, 539, 557, 558, 559, 560, 566, 567, 569, 581, 619, 620, 621, 623, 624, 626, 627, 629, 633, 634, 635, 859, 861, 862, 1036, 1133
- adult, vi, xiii, 93, 94, 95, 100, 101, 102, 103, 104, 105, 107, 108, 149, 160, 161, 164, 273, 289, 298, 303, 320, 324, 327, 338, 368, 458, 463, 466, 468, 469, 471, 480, 481, 485, 486, 490, 492, 505, 506, 518, 526, 532, 537, 634, 666, 675, 844, 960, 1036, 1039
- adult education, 149, 463
- adulthood, 254, 345, 351, 394, 396, 402, 458, 459, 461, 462, 463, 464, 466, 468, 469, 470, 471, 473, 486, 488, 505, 860, 1042, 1133
- adults, xx, xxiii, xxix, 18, 38, 43, 91, 114, 151, 154, 195, 230, 242, 253, 298, 314, 393, 395, 457, 458, 460, 461, 462, 463, 466, 467, 468, 469, 472, 477, 485, 486, 498, 499, 500, 506, 518, 525, 526, 528, 529, 558, 559, 616, 619, 622, 623, 624, 627, 628, 660, 673, 737, 936, 937, 960, 1033, 1034, 1035, 1036, 1039, 1040, 1041, 1133
- adverse event, 38, 47, 49, 55, 60, 868, 874
- advertisements, 901, 905
- advertising, 979
- advocacy, xx, 55, 167, 459, 475, 632
- aerobic, 151
- aerobic exercise, 151
- aetiology, 169
- affect, 477, 484, 485, 495, 506
- affective dimension, 166
- affective disorder, 484
- affective reactions, 204, 212, 235

- affective states, 18, 202, 203
- Africa, 719, 722
- African American, xviii, 119, 189, 194, 196, 297, 298, 300, 301, 302, 304, 307, 308, 309, 310, 462, 463, 464, 470, 533, 534, 536, 538
- African American women, 196
- African Americans, 300, 301, 302, 307, 308
- afternoon, 1023
- ageing, 118, 135, 434
- agent, 140, 586, 597, 601, 610, 612, 641, 645, 654, 655, 680, 719, 753, 912, 914, 919, 926, 1101, 1102
- agents, xxiv, 43, 328, 576, 597, 641, 707, 708, 709, 710, 807, 894, 902, 912, 922
- aggregates, 736, 738
- aggregation, 753, 805, 1071
- aggression, xiii, xxi, 75, 79, 82, 87, 313, 479, 483, 485, 486, 513, 517, 520, 521, 526, 665, 844, 862, 1132, 1133
- aggressive behavior, 78, 82, 479, 484, 486, 525, 1132
- aggressive therapy, 301
- aggressiveness, 861, 863
- aging, 140, 178, 678, 682, 937
- aging population, 678, 682
- aging society, 678
- agriculture, 982
- aid, viii, xii, xv, xxi, xxv, xxvi, 9, 10, 11, 78, 81, 84, 130, 144, 145, 184, 185, 186, 188, 189, 191, 192, 193, 194, 195, 272, 317, 381, 440, 557, 735, 767, 768, 774, 775, 776, 777, 778, 779, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 793, 796, 797, 845, 870, 875, 936, 947, 954, 999, 1005,
- 1007, 1010, 1028, 1111
- aiding, 338, 381, 1118
- AIDS, 393, 472, 505, 506, 511, 535, 538, 622, 628, 629, 634, 636
- air, 714, 870
- Alaska, 298, 871
- Alberta, 157, 885
- alcohol, xx, xxi, 87, 228, 249, 254, 255, 412, 437, 452, 459, 462, 465, 466, 468, 469, 471, 473, 479, 504, 507, 510, 511, 512, 513, 514, 531, 533, 535, 526, 507, 507, 510, 511, 512, 513, 514, 531, 533, 535,
  - 536, 580, 586, 587, 621, 843, 863, 900
- alcohol abuse, 471
- alcohol consumption, 452, 511, 512, 900
- alcohol dependence, 469
- alcohol use, xx, 459, 469, 473, 504, 511, 512, 513
- alcoholics, 250, 460
- alertness, 208, 987
- alfalfa, 1018
- algorithm, 40
- alienation, 460

allele, 96, 313 alleles, 95, 96, 110, 313 allogeneic, 100, 106, 110, 111 alpha, xxi, 241, 531, 534, 535, 561, 562, 613, 657, 776, 777, 791 alternative hypothesis, 605, 648 alternatives, xxix, 10, 108, 199, 200, 204, 216, 264, 266, 289, 314, 350, 357, 367, 449, 461, 482, 535, 572, 573, 578, 588, 598, 642, 681, 730, 770, 774, 786, 832, 892, 893, 901, 902, 919, 920, 922, 973, 976, 977, 978, 979, 1013, 1014, 1015, 1016, 1018, 1028, 1034, 1037, 1051, 1074, 1075, 1076, 1077, 1080, 1083, 1087, 1091, 1092, 1093, 1094, 1095, 1102 alters, xxv, 735, 1019 altruism, xxiv, 223, 707, 713 aluminium, 875, 876 Alzheimer's disease, 137, 187, 188, 196, 380 amalgam, 915 ambiguity, 215, 264, 548, 613, 622, 624, 633, 657, 731, 804, 809, 901, 903, 907, 910, 1067 ambivalence, 369, 805, 886 ambivalent, 348, 581 American Academy of Pediatrics, 500 American Civil Liberties Union, 620 American Indian, 298 American Psychological Association, 91, 472, 500 amino, 1118 amino acid, 1118 amino acids, 1118 AML, 111 amnesia, 255 amniocentesis, xv, 183 AMPA, 259 amphetamine, 212, 233, 251, 436, 613, 614, 657, 1132, 1133, 1134 amphetamines, 412, 413, 436 Amsterdam, 385, 593, 637, 764, 1009, 1030, 1134 amygdala, xvi, 197, 205, 207, 208, 209, 210, 211, 212, 214, 215, 216, 218, 219, 220, 223, 224, 225, 226, 227, 228, 229, 230, 233, 239, 242, 245, 247, 251, 253, 255, 617, 661, 1118, 1119, 1130 anaesthetists, 171 analgesic, 178 analog, 752, 755, 756, 1032 analysis of variance, 563 analytical framework, 870 analytical thinking, 315 anatomy, 229, 241, 257, 394, 497 ANC, 444, 445 anemia, 94

anger, xii, 16, 18, 45, 57, 60, 167, 180, 202, 208, 210, 216, 284, 459, 479, 483, 485, 486, 844, 845, 859, 862, 863 anger management, 57 Anglo-Saxon, 263, 267 animal cognition, 1026 animal learning, 234, 1024 anorexia, 482 anorexia nervosa, 482 ANOVA, 563, 669, 694, 695, 1054, 1057, 1061, 1117, 1121, 1123, 1125, 1126, 1129 antagonism, 60, 793 antecedents, xxii, 229, 469, 498, 557, 566, 612, 616, 655, 659, 733, 980, 1066, 1131 anterior cingulate cortex, xvi, 197, 207, 228, 232, 1132 anthropological, 355, 572 anthropology, 158, 355, 572 anticancer, 32, 43 antidepressant, 474 antigen, xv, 183 antipsychotic, 79 antisocial, 264, 268, 481 antisocial acts, 81 antisocial behavior, 481 antisocial personality, 87, 264, 268 antisocial personality disorder, 87, 264, 268 antithesis, 477 anus, 626 anxiety, 47, 111, 115, 167, 179, 181, 186, 201, 209, 214, 229, 233, 267, 283, 284, 316, 324, 328, 348, 352, 371, 379, 409, 418, 485, 491, 519, 524, 527, 558, 566, 775, 903, 1068 anxiety disorder, 485 apathy, 483 aplasia, 100 appetite, 99, 101, 103, 258, 1131 apples, 594, 638 application, xv, 7, 13, 14, 15, 17, 28, 34, 41, 42, 65, 73, 87, 184, 263, 265, 269, 315, 382, 500, 545, 711, 715, 758, 769, 870, 901, 1077, 1134 applied research, 518, 1064 appraisals, 1046, 1053 Argentina, 756, 757, 758, 759, 762, 763, 764 argument, 149, 187, 263, 270, 292, 546, 547, 577, 608, 651, 710, 770, 793, 856, 1003 Aristotle, 987 arithmetic, 544 Army, 140 arousal, 16, 211, 217, 229, 258, 480 arrest, 83, 286, 316 arthritis, 71, 119, 124, 125, 134, 136, 137, 138, 139, 140, 149, 150, 152, 154, 155, 179

arthroplasty, 149, 150, 152, 153, 156 arthroscopy, 137 artificial, 488 Asia, 455, 591 Asian, xviii, xxi, 194, 297, 298, 299, 300, 302, 303, 306, 308, 454, 455, 462, 463, 531, 533, 534, 536, 538, 559, 567, 568, 569, 894 Asian American, 302, 462, 463, 559, 568, 569 Asian Americans, 302, 559, 568, 569 Asian countries, 303 asphalt, 574, 579 aspiration, xxi, 49, 314, 532, 533, 537, 559, 565, 698, 895, 1008 assault, 83, 85, 478, 492, 494, 499, 500 assaults, 520 assertiveness, 58 assessment procedures, 84, 908 assessment tools, 91, 169, 844, 860 assets, xxviii, 594, 639, 682, 683, 685, 687, 688, 702, 752, 754, 829, 936, 941, 942, 943, 946, 947, 948, 1106, 1107, 1108, 1109 assignment, 218 assimilation, xxviii, 309, 916, 917, 951, 985, 997, 1006, 1007, 1008, 1069 association, 266, 478, 483, 486, 499, 505, 532 associations, 267 assumptions, xxix, 23, 56, 77, 122, 140, 147, 158, 176, 269, 306, 314, 320, 368, 568, 572, 589, 596, 600, 640, 644, 679, 685, 687, 694, 709, 710, 711, 834, 894, 912, 914, 916, 923, 928, 951, 1004, 1026, 1077, 1078, 1083, 1086 asthma, 478 asymmetric information, 768 asymmetry, 217, 962, 963 Athens, 541 Atlantic, 630, 812, 940, 944 atmosphere, 401, 406, 445, 988, 992 atrocities, 163 attachment, 461, 481, 483, 506, 559, 566 attacks, 38, 201, 525 attention, xii, xiv, xvi, xxi, xxv, xxvii, 2, 15, 17, 22, 30, 70, 105, 124, 157, 159, 160, 164, 165, 169, 179, 193, 194, 197, 204, 230, 244, 246, 247, 249, 264, 269, 272, 298, 313, 320, 329, 397, 442, 484, 487, 488, 537, 544, 546, 547, 557, 558, 579, 632, 673, 685, 687, 735, 792, 848, 864, 882, 885, 911, 988, 996, 1003, 1008, 1068, 1070, 1100, 1117, 1126, 1127, 1128, 1130 attractiveness, 204, 682, 697, 699, 736, 873, 897, 902 attribution, 345, 733, 863, 1010 attribution theory, 345 atypical, 527

audio, 521

Australasia, 769

- Australia, xi, 5, 264, 265, 266, 275, 387, 476, 489, 767, 768, 771, 773, 776, 781, 795, 797, 841, 842,
  - 845, 846, 850, 861, 863, 1099, 1100

Austria, 1009

- authoritarianism, 309
- authority, 52, 107, 304, 305, 411, 493, 559, 565, 566, 585, 697, 990, 1001

autolysis, 528

automatic processes, 203, 207

- autonomic, 485
- autonomic activity, 206
- autonomy, v, xiv, xvii, xx, 5, 6, 94, 96, 107, 108, 163, 182, 184, 263, 280, 284, 285, 287, 288, 292, 295, 297, 299, 303, 306, 307, 308, 317, 328, 333, 461, 465, 469, 503, 560, 565, 578, 704, 769, 796
- autoradiography, 257
- autosomal dominant, 312, 313, 380, 383
- availability, 6, 10, 94, 116, 176, 212, 215, 249, 264, 281, 302, 312, 316, 365, 369, 436, 460, 461, 584, 589, 621, 772, 789, 903, 904, 937, 1014, 1016, 1048
- averaging, xxix, 762, 954, 960, 968, 976, 1013, 1015, 1017, 1018, 1019, 1020, 1021, 1024, 1027, 1030
- aversion, xiii, xxiv, xxv, 16, 63, 65, 66, 70, 72, 74, 199, 200, 201, 203, 214, 216, 227, 595, 597, 600, 605, 611, 639, 641, 644, 648, 654, 677, 686, 689, 694, 695, 696, 704, 707, 712, 714, 716, 717, 718, 730, 736, 737, 746, 751, 752, 753, 754, 755, 758, 759, 762, 763, 764, 800, 804, 805, 830, 834, 871, 884, 890, 896, 898, 905, 908, 910, 918, 919, 921, 932, 962, 963, 966, 976
- aviation, 57
- avoidance, xvi, 16, 165, 201, 204, 233, 261, 262, 264, 265, 266, 268, 269, 270, 272, 274, 281, 282, 373, 492, 605, 648, 880

avoidant, 479

awareness, xii, 5, 6, 9, 41, 54, 57, 60, 96, 108, 109, 161, 164, 165, 166, 168, 169, 177, 180, 286, 289, 299, 300, 321, 335, 338, 349, 356, 357, 359, 360, 361, 362, 363, 364, 365, 370, 372, 376, 377, 394, 452, 461, 477, 479, 543, 545, 546, 549, 558, 708, 943, 990

### В

babies, 346, 365, 421, 425, 426, 428, 430, 438, 443, 483, 584

back pain, 67, 73, 133, 160, 162, 179, 180, 181, 182 background information, xv, 158, 159, 356, 562 backlash, 107, 479

backwardness, 581 bail, 874 Bali, 454, 838 banks, 1112 bargaining, 19 barrier, xv, 52, 58, 161, 168, 184, 189, 369, 395, 629, 667, 789, 795, 800 barriers, 52, 161, 163, 168, 169, 185, 187, 191, 270, 284, 340, 360, 461, 466, 467, 470, 558, 559, 568, 667, 668, 788 basal forebrain, 252, 256, 257 basal ganglia, 225, 226, 239, 240, 241, 249, 1134 basal nuclei, 208 base rate, 77, 80, 82, 85, 88, 1040 base rate information, 1040 base rate neglect, 1040 base rates, 77, 80, 82, 88 basic needs, 164, 477, 488 basic research, 980 Bayesian, v, xii, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 66, 72, 675, 739, 802, 930 Bayesian analysis, xii, 25, 26, 30, 35, 39, 41, 42 Bayesian estimation, 31, 43 Bayesian methods, 27, 29, 30, 31, 32, 36, 41, 43 beating, 422, 425, 428, 434, 448, 478 beer, 250, 435, 535 behavioral disorders, 252 behavioral manifestations, 495 behavioral problems, xx, 475, 480, 1037 behavioral sciences, 986 behavioral theory, 613, 656 Beijing, 15, 561, 569, 725 Belgium, 292, 1033 belief systems, 170, 177 beliefs, xii, 25, 26, 28, 29, 52, 53, 60, 69, 73, 122, 123, 135, 136, 139, 147, 158, 167, 168, 169, 170, 172, 177, 181, 267, 286, 298, 300, 303, 307, 308, 339, 344, 353, 382, 397, 402, 448, 452, 458, 459, 469, 548, 550, 554, 555, 558, 560, 563, 566, 624, 633, 708, 773, 785, 792, 803, 807, 809, 891, 892, 926, 1034, 1035, 1067 bell, 727 beneficial effect, 107 benefits, xiv, xv, xxiv, xxvi, 16, 45, 46, 64, 65, 68, 70, 71, 94, 107, 108, 109, 117, 160, 183, 184, 186, 187, 189, 198, 199, 221, 222, 223, 224, 264, 280, 288, 289, 315, 332, 339, 343, 488, 594, 597, 638, 640, 664, 675, 678, 718, 725, 726, 730, 738, 768, 769, 770, 771, 772, 773, 774, 775, 776, 786, 787, 789, 792, 793, 799, 800, 801, 802, 803, 804, 805, 806, 807, 809, 811, 873, 874, 880, 884, 941, 946, 1066, 1090, 1100, 1116

benign, 38 bereavement, 8 Bernanke, Ben, 719 Best Practice, 151 beta, 1101, 1106, 1107, 1108, 1109, 1110 betrayal, 480 bias, xxx, 11, 69, 79, 85, 169, 191, 200, 202, 204, 209, 230, 250, 308, 344, 369, 483, 490, 495, 518, 519, 520, 666, 698, 732, 738, 739, 773, 785, 833, 844, 856, 862, 884, 886, 908, 918, 931, 952, 953, 954, 962, 968, 975, 1034, 1035, 1043, 1045, 1047, 1052, 1054, 1055, 1056, 1057, 1058, 1062, 1064, 1065, 1068, 1069, 1070 Bible, 308, 588 binding, 244, 255, 257, 290 biodiversity, 10, 13, 14 bioethics, 299, 308 biological, 478, 485, 506, 509, 512, 513 biological markers, 485 biological parents, 509, 512, 513 biological processes, 28 biological systems, 485 biology, 632 biomedical model, 144, 312, 584 **Biomedical Perspective**, 144 biopsies, 110 biopsy, 96, 110 biotechnology, 386, 389, 800, 801, 810 birds, xxiv, 599, 600, 613, 643, 657, 663, 666, 667, 668, 669, 671, 673, 674, 675, 1015, 1023 birth, viii, xxiii, xxv, xxvi, 312, 331, 334, 343, 345, 346, 350, 356, 379, 381, 392, 395, 398, 402, 403, 483, 506, 532, 535, 573, 575, 576, 578, 580, 583, 584, 585, 587, 589, 619, 620, 632, 767, 768, 771, 772, 773, 774, 775, 776, 783, 784, 785, 786, 787, 788, 789, 790, 791, 793, 794, 795, 797, 936 birth control, 346, 356, 395, 535, 573, 575, 576, 578, 583, 584, 587, 589, 632 birth rate, xxiii, 619, 620 birth rates, xxiii, 619, 620 birth weight, 483 birthing women, 797 births, 350, 353, 392, 577, 585, 620, 771, 796 black, 267, 272 black hole, 49 blame, 180, 342, 345, 359, 361, 407, 437, 444, 480, 489, 726 bleeding, 49, 414, 425, 427, 428, 429, 437, 438, 441, 480 blindness, 267 blocks, 17, 163, 228, 253, 1080, 1093, 1119, 1122, 1123, 1125

blood, 94, 105, 255, 312, 322, 333, 363, 394, 428, 437, 442, 443, 485, 492, 584, 882, 984, 996, 1006 blood flow, 255 blood pressure, 322, 485 blood supply, 394 blood transfusion, 94, 105 blood transfusions, 94, 105 blurring, 383 BMA, 286 board members, 1103, 1105, 1110, 1111 boats, 543, 548, 549 body, 480, 491, 492 body dissatisfaction, 979 body image, 161, 523, 524, 527 body language, 57 body size, 599, 609, 643, 653, 966, 971 boils, 610, 654 BOLD, 257 bonding, 429, 461 bonds, 687, 717 bone marrow, vi, xiii, 93, 95, 96, 97, 110, 111 bone marrow transplant, vi, xiii, 93, 95, 110, 111 borderline, 87 borderline personality disorder, 87 boredom, 843 borrowing, 700 Boston, 61, 110, 112, 180, 183, 254, 383, 472, 590, 592, 636, 863, 885, 931, 1008, 1009 bottom-up, 12 bounds, 838 boutons, 240 bowel, 167, 169, 179, 180, 352, 380, 499 boys, 393, 439, 508, 509, 510, 511, 512, 513, 518, 520, 521, 549, 587, 588, 634, 859 brain, xi, xvi, xxix, xxxi, 1, 2, 3, 4, 17, 28, 38, 137, 159, 161, 197, 198, 206, 207, 208, 209, 211, 212, 213, 214, 215, 216, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 233, 235, 238, 239, 240, 242, 243, 245, 248, 249, 250, 251, 252, 254, 256, 257, 301, 313, 471, 484, 499, 609, 613, 614, 615, 616, 617, 653, 657, 658, 659, 660, 870, 912, 913, 1033, 1042, 1043, 1115, 1116, 1118, 1119, 1120, 1127, 1128, 1131, 1134 brain activity, xi, 1, 2, 228, 231, 252 brain chemistry, 870 brain damage, 207, 301, 484 brain injury, xvi, 3, 197 brain size, 609, 653 brain stem, 208, 209, 213, 216, 220, 225 brain structure, 3, 207, 239, 240, 242, 248, 1127, 1128 Brazil, vii, xxii, 571, 572, 574, 575, 577, 578, 579, 581, 582, 583, 584, 590, 591, 592

Brazilian, xxii, 571, 572, 574, 575, 576, 577, 579, 590, 591 BRCA, 312, 315 breakdown, 164, 171, 249, 367, 581, 582 breakfast, 15 breast cancer, 69, 73, 312, 387, 472 breathing, 184 breeding, 671 Britain, xvi, 29, 181, 261, 264, 265, 266, 276, 350, 635 British, 261, 264, 265, 267, 269, 271, 275, 276 British Columbia, 332, 729 broadband, 983 Bronfenbrenner, 461, 469 brothers, 321, 330, 395, 427, 435, 586 brutality, 519, 520 budding, 251 buffer, 514 building blocks, 17 bullies, 486 bullying, 51, 510 burn, 266 burning, 478 burnout, 272 buses, 1078 business, 535, 752, 762, 765, 812, 1100, 1101, 1107 business cycle, 752, 765 buttons, 727 bypass, 363, 883

### С

cabbage, 133, 146

- cache, 1020, 1022, 1029
- caesarean section, xxv, xxvi, 767, 768, 771, 772,
  - 776, 785, 786, 787, 788, 793, 794, 795, 796, 797
- California, 622, 722, 733, 828, 838
- campaigns, 487, 800, 860
- Canada, 61, 72, 115, 150, 157, 265, 266, 311, 313, 317, 331, 333, 368, 382, 464, 538, 769, 772, 795, 885, 1045
- cancer, xviii, 5, 8, 28, 29, 69, 73, 99, 105, 110, 111, 164, 181, 185, 189, 284, 286, 287, 292, 293, 294, 295, 297, 300, 301, 302, 304, 305, 307, 309, 312, 315, 316, 322, 326, 331, 332, 333, 334, 335, 373, 380, 387, 471, 472, 473, 485, 726
- cancer treatment, 293
- candidates, 95, 316, 319, 327, 328, 329, 920, 978
- capacity, 479, 489, 1100
- capital, 1101, 1106, 1107, 1108, 1109, 1110
- Capital Asset Pricing Model, 752
- capital input, 1106
- capital markets, 678, 694, 699

capitalist, 573 car accidents, 1087 cardiopulmonary, 185, 187, 188, 301 cardiopulmonary resuscitation (CPR), 185, 187, 188, 301 cardiovascular disease, 72 career counseling, 565 career development, 566, 567, 568, 569 career success, 566 caregiver, 163, 283, 285, 289, 314, 319, 324, 325, 327, 329, 477 caregivers, xvii, 6, 8, 49, 192, 193, 279, 280, 284, 285, 287, 292, 294, 309, 319, 324, 325, 330, 472, 483, 790 caregiving, 312, 317, 323, 329, 330 caretaker, 429, 434, 478, 479, 489, 498, 579 carrier, 312, 329, 379, 380, 384, 386, 387 cartilage, 137 case law, 290 case study, 383, 399, 870, 877 cash flow, 1103 cast, 713 castration, 1134 catatonia, 79 catatonic, 79 catatonic schizophrenia, 79 catchments, 1037 categorization, 319, 705, 800, 909, 932 category a, 300 Catholic, 432, 433, 434, 455, 575 Catholic Church, 432, 433, 434, 575 cats, 258 cattle, 1029 Caucasian, 119 Caucasians, 34 causal attribution, 139, 863 causal model, 307, 538, 864 causal reasoning, 1041 causal relationship, 514 causality, xxi, 339, 504, 520, 733, 1003 causation, 18 C-C, viii, 751, 752, 763, 764 CDC, 515, 620, 771 cell, 3, 97, 106, 109, 110, 111, 239, 257, 313, 378, 384, 620, 876 cell phones, 876 cell transplantation, 110, 111 cement, 115, 574 Census, 298, 310, 942 Census Bureau, 298, 310 Centers for Disease Control, 620, 633 central nervous system, 257, 298

CEO, 1100, 1101, 1102, 1103, 1104, 1106, 1107, 1108, 1109, 1110, 1112 cerebral cortex, 227, 932 cerebral hemorrhage, 99 cerebrospinal fluid, 1132 certainty, 69, 80, 85, 166, 181, 558, 559, 594, 638, 679, 681, 698, 704, 711, 898, 917, 919, 929, 1014, 1067, 1069, 1081 certificate, 314 cervical cancer, 471 cesarean section, 577, 590, 794, 795, 796, 797 CFA, 849, 852, 853, 857 CFI, 851, 852, 857, 858 chain of command, 53 changing environment, 1030 channels, 451, 800, 809 chaos, 381 chaotic, 483 cheating, 458 cheese, 874 chemical, 1107 chemotherapy, 32, 97, 110, 111, 164, 281, 287 Chicago, 496, 497, 500, 613, 704, 722, 929, 1096 chicks, 673 Chief of Staff, 56 child abuse, 267, 459, 476, 477, 482, 483, 484, 486, 487, 492, 493, 495, 496, 497, 498, 499, 500, 501 child development, xxi, 482, 483, 498, 517, 1041 child maltreatment, 476, 478, 482, 483, 484, 487, 496, 498 child molesters, 91 child protection, 272, 482, 487, 489, 490, 495 child protective services, 476, 478 child rearing, 487 child welfare, 478, 489 childbearing, 316, 340, 351, 383, 390, 506, 515, 572, 578,620 childbirth, xx, 459, 503, 505, 576, 577, 585, 771, 773, 784, 786, 789, 792, 795, 797 childcare, 477, 573, 578 childhood, xx, 94, 169, 186, 254, 335, 356, 374, 381, 394, 402, 452, 473, 475, 476, 478, 479, 481, 482, 485, 486, 488, 498, 499, 500, 503, 505, 518, 519, 532, 542, 1042, 1133 childhood sexual abuse, 481, 498 childless, 347, 348, 365, 382, 385 childrearing, 572 Chile, 756, 757, 758, 759, 760, 762, 763, 863 China, vii, xxii, 15, 557, 569, 725 Chinese, 611 Chinese women, 471 Chi-square, xx, 504, 508 chorea, 313, 381, 382, 384, 385

Christianity, 568 Christmas, 717 chromosome, 313 chromosomes, 333 chronic, 479, 482, 485, 486, 496, 498, 518, 527 chronic disease, 32, 41, 64, 71, 146, 178 chronic diseases, 32, 41 chronic disorders, 125, 486 chronic fatigue syndrome, 482 chronic illness, 95, 150, 151, 152, 155, 156, 164, 182, 333 chronic pain, xiv, xv, 157, 158, 159, 160, 161, 162, 164, 165, 166, 167, 168, 170, 171, 172, 174, 175, 177, 179, 180, 181, 182 chronically ill, 496, 498 CHS, 462, 463, 468 chunking, 1022 cigarette smoking, 459, 460, 510, 511, 616, 660 cigarettes, 459, 460, 462, 465, 468, 512, 514 cingulated, 233, 234 cirrhosis, 301, 586 citizens, 10, 185, 263, 486, 990, 1007 citizenship, 262 civil liberties, 266, 268 civil rights, 265 classes, 52, 58, 84, 87, 245, 435, 463, 476, 507, 508, 509, 563, 581, 583, 622, 925, 1120 classical, 26, 31, 32, 35, 381, 527, 912, 925 classification, 82, 83, 84, 85, 86, 87, 91, 119, 131, 909, 985, 1009, 1107 classified, 477, 829, 1107, 1108, 1109 classroom, xx, 457, 461, 464, 466, 468, 469, 507, 549, 981, 984, 987, 988, 991, 992, 993, 994, 995, 996, 999, 1000, 1001, 1005, 1006, 1008, 1011, 1050, 1070 classrooms, 563, 992, 1000 cleaning, 120, 458, 524, 586, 588 clients, 80, 87, 179, 265, 269, 270, 317, 352, 443, 566, 568, 709, 737, 869, 874, 877 clinical, xxi, 476, 485, 490, 492, 494, 495, 517, 518, 519, 521, 524, 525, 526, 527, 528, 624, 627 clinical approach, 518 clinical assessment, 521, 1131 clinical depression, 286 clinical examination, 524 clinical judgment, 64, 76 clinical psychology, 89 clinical trial, 29, 30, 31, 32, 38, 41, 42, 43, 107, 109, 156, 791 clinical trials, 29, 30, 31, 32, 38, 41, 42, 43, 107, 109, 156 clinically significant, 790

clinician, xx, 29, 41, 66, 71, 82, 84, 117, 145, 164, 192, 287, 288, 475 clinicians, 490, 495, 514 clinics, xix, 352, 361, 392, 410, 417, 422, 424, 425, 432, 434, 436, 439, 440, 441, 443, 444, 448, 451, 453, 463, 482, 793 cloning, 805 close relationships, 285, 329 closure, 187, 268, 269, 271 clothing, 492 clustering, 33, 34 clusters, 34, 36, 251, 319, 364 CMV, 96, 97, 98, 99 CNAs, 49 CNN, 936 Co, 221, 222, 223, 561, 719, 720, 762, 929 coactors, 1049, 1050, 1064, 1067, 1071 coastal zone, xxvi, 811 cocaine, 212, 228, 241, 249, 252, 254, 255, 257, 259, 507, 511, 535 cocaine abuse, 255 Cochrane, 185, 195, 764, 769, 794, 796 Cochrane Database of Systematic Reviews, 794, 796 codes, 122, 359, 390, 534, 535, 792 coding, 122, 234, 257, 258, 319, 404, 547, 617, 630, 660, 895, 1001, 1031, 1132 coefficient of variation, 605, 648 coercion, 285, 480 co-existence, 264 coffee, 55, 712, 883 cognition, xxvii, 17, 18, 124, 155, 204, 225, 228, 230, 233, 253, 313, 314, 315, 317, 330, 332, 473, 524, 566, 660, 904, 911, 913, 918, 925, 1026, 1029, 1030, 1031, 1069, 1133 cognitive, 484, 489, 490, 505, 519, 521, 522, 523, 524, 525, 632 cognitive abilities, 913, 1009 cognitive ability, xvii, 279, 281, 289, 903, 1036, 1039, 1042 cognitive activity, 198, 221 cognitive biases, 369 cognitive capacities, 1035 cognitive capacity, 286, 1035, 1036, 1038, 1039, 1040 cognitive deficit, 81 cognitive deficits, 81 cognitive development, xxix, 1033, 1041 cognitive dissonance, 344, 383 cognitive domains, 922 cognitive effort, 219 cognitive flexibility, 202 cognitive function, 162, 164, 179, 204, 211, 216, 306

cognitive impairment, 190, 194, 301, 317

cognitive level, 994 cognitive load, 1027, 1035 cognitive models, 385, 389, 568, 1134 cognitive perspective, 1070 cognitive process, 3, 160, 205, 223, 340, 344, 484, 524, 664, 672, 679, 916, 924, 925, 954, 1022 cognitive processing, 340, 344, 484 cognitive psychology, xii, xvi, 15, 16, 161, 237, 248, 664, 979 cognitive science, 554, 931, 979 cognitive style, 1042 cognitive system, 124, 924, 925, 926 cognitive tasks, 220, 1133 coherence, xiv, xv, 157, 158, 159, 160, 161, 168, 169, 171, 263, 762, 912 cohesion, 263, 478, 560, 567 cohort, 8, 31, 99, 105, 108, 172, 775 coitus, 512, 525 collaboration, vii, 12, 46, 47, 51, 52, 53, 57, 61, 62, 98, 188, 192, 284, 373, 403, 450, 452, 517, 708, 769, 794 College Entrance Examination, 567 college students, xxi, 557, 561, 562, 565, 567, 568, 614, 623, 626, 632, 634, 658, 737 colleges, 504 collisions, 856 Colombia, 71, 756, 757, 758, 759, 761, 762, 763, 764 colon, 301, 315, 331, 334 colon cancer, 301, 315, 331, 334 Colorado, 464 colorectal cancer, 335 colors, 476, 666 coma, 159, 194, 306 combined effect, 82, 499 commercial, 479 commitment, 273, 274 commodities, 594, 601, 638, 645 commodity, 594, 595, 601, 639, 644, 645, 1101 common law, 492, 495 communication, xii, xiv, xvi, xvii, 6, 7, 45, 46, 47, 49, 51, 52, 53, 55, 56, 57, 58, 61, 62, 90, 94, 95, 98, 99, 100, 105, 108, 111, 163, 164, 165, 166, 167, 175, 179, 182, 185, 187, 191, 263, 267, 279, 282, 284, 285, 287, 293, 294, 302, 306, 316, 322, 331, 332, 343, 360, 394, 412, 451, 452, 460, 488, 493, 507, 509, 510, 512, 513, 522, 537, 786, 800, 809, 823, 828, 875, 877, 881, 990, 1104 communication skills, xvi, xvii, 57, 166, 175, 279 communication strategies, xiv, 94, 105, 108, 800 communities, xviii, xix, 162, 297, 299, 300, 303, 306, 307, 391, 396, 399, 400, 401, 403, 406, 409,

410, 418, 430, 449, 454, 476, 484, 488, 506, 769, 826 community, xxii, 78, 79, 80, 85, 106, 154, 156, 162, 261, 264, 265, 266, 268, 270, 271, 272, 273, 275, 276, 302, 307, 334, 401, 402, 406, 407, 409, 417, 418, 419, 420, 421, 423, 437, 439, 443, 447, 448, 449, 451, 452, 453, 460, 461, 469, 472, 478, 483, 484, 487, 488, 489, 495, 498, 504, 505, 518, 571, 572, 575, 577, 579, 580, 583, 587, 588, 625, 633, 739, 748, 769, 797, 978, 1026, 1035 community service, 489 communities, 264, 265, 266, 270, 271, 272, 273 community-based, 489, 633 Comparative Fit Index, 852, 858 compatibility, 373 compensation, 200, 829, 833, 1101, 1103, 1104, 1112 competence, 67, 69, 286, 289, 340, 341, 342, 375, 469, 483, 485, 486, 489, 546, 730, 929, 1003, 1042, 1046, 1047, 1064, 1070 competency, xvii, 51, 53, 54, 57, 58, 279, 281, 286, 792, 859 competition, xxii, 221, 226, 546, 557, 728, 729, 989, 1046, 1065, 1112 competitive advantage, 1103 complement, xxv, 751 complementary, 524, 825 complete remission, 38, 106, 110, 111 complexity, xxi, 42, 50, 130, 158, 178, 181, 182, 185, 198, 207, 211, 218, 281, 284, 332, 369, 376, 495, 517, 518, 520, 547, 548, 560, 567, 584, 623, 890, 902, 905, 906, 913, 1007, 1029, 1040, 1077, 1131 compliance, xvii, 7, 58, 69, 71, 72, 77, 94, 96, 150, 170, 279, 466, 483, 583, 774, 789, 843 complications, xix, 95, 109, 391, 393, 394, 450, 772 components, xii, xiv, xxv, 6, 25, 27, 33, 69, 125, 157, 159, 162, 169, 170, 172, 175, 181, 198, 204, 234, 266, 284, 344, 352, 357, 397, 403, 451, 486, 491, 524, 525, 755, 767, 774, 788, 826, 827, 834, 836, 846, 895, 904, 909, 984, 990, 998, 1031, 1049, 1087 composition, xxx, 59, 74, 1099, 1101, 1102, 1103, 1107, 1108, 1110, 1111 comprehension, 109, 168, 952, 1002, 1004 compulsive behavior, 313 computation, 41, 926, 932, 1027, 1035, 1036, 1039, 1040 computer graphics, 186 computer technology, 993 computers, 27, 387, 805, 918, 983, 984, 985, 993, 999, 1000, 1001, 1007, 1009

computing, 27, 29, 41, 1027, 1041, 1050

concave, 836 concentrates, 400, 1006, 1064 concentration, 47, 55, 60, 245, 247, 248, 250, 523, 573, 762, 1133 conception, 96, 190, 328, 338, 394, 547, 553, 555, 584, 589, 595, 639 conceptual model, 111, 461, 813, 814, 826, 829 conceptualization, 23, 170, 179, 529, 567, 986 conceptualizations, 242 concordance, 118, 120, 167, 193, 873 concrete, 22, 188, 190, 221, 242, 685 conditioned stimulus, 208 conditioning, 97, 99, 104, 212, 254, 1031 condom, xx, 437, 504, 511, 513, 535, 628, 629, 636 condoms, 506, 511, 631, 632, 634 conduct, 486 conduct disorder, 486, 520, 521, 522, 523, 524, 525 conductance, 205, 206, 209 conduction, 29, 31 confabulation, 491 confession, 490, 500 confidence, xxvi, 29, 52, 58, 89, 167, 199, 352, 369, 374, 507, 508, 512, 519, 543, 546, 558, 564, 688, 726, 785, 841, 844, 847, 849, 850, 851, 853, 854, 855, 856, 858, 859, 860, 945, 1003, 1004, 1028, 1070, 1085 confidence interval, 29, 507, 508, 512, 688, 1085 confidence intervals, 508, 688, 1085 confidentiality, xix, 59, 327, 354, 392, 404, 405, 449, 451, 488, 493, 1051 confinement, 579 Confirmatory Factor Analysis, 91, 849, 850, 857, 853,863 conflict, xiv, xxv, 10, 13, 14, 46, 52, 57, 78, 94, 107, 162, 163, 170, 198, 210, 211, 223, 224, 225, 230, 267, 270, 287, 289, 299, 389, 416, 420, 426, 427, 445, 446, 458, 481, 485, 489, 493, 560, 564, 567, 569, 726, 728, 731, 767, 770, 773, 774, 775, 776, 779, 783, 786, 787, 789, 790, 792, 796, 931, 1035, 1116, 1119, 1124, 1126 conflict analysis, 14 conflict resolution, 267, 770 conformity, 262, 263, 455 confrontation, 8, 46, 51, 52 confusion, 57, 196, 490, 628, 709, 710, 784, 914 Congress, iv congruence, 170, 177, 480, 774 conjecture, 922 conjunction fallacy, 1043 Connecticut, 569, 733 connectionist, 931 connectivity, 229, 239, 240, 242 conscious awareness, 160, 546

consciousness, xv, 69, 109, 158, 159, 160, 161, 163, 164, 165, 166, 167, 168, 169, 170, 172, 173, 176, 178, 179, 206, 394, 425, 525 Consciousness Model, xiv, xv, 157, 158, 159, 161, 172 consensus, xxiii, 51, 106, 118, 119, 196, 304, 346, 487, 490, 518, 524, 619, 620, 623, 624, 626, 627, 711, 771 consent, 95, 98, 107, 108, 109, 120, 289, 293, 295, 356, 404, 405, 479, 492, 533, 563, 576, 577, 585, 587, 589, 714, 739, 769, 774, 1051 consequence, 262 consequences, 269 conservation, 14 consolidation, 341, 344, 389, 546, 982, 998 constipation, 99, 101, 480 constitution, 272 constraints, xxii, 12, 270, 328, 330, 339, 341, 343, 351, 545, 546, 571, 572, 578, 600, 643, 722, 789, 903, 915, 925, 993, 1038, 1091 construction, 114, 120, 123, 128, 146, 147, 151, 155, 158, 160, 161, 162, 163, 168, 169, 173, 175, 176, 177, 178, 312, 321, 323, 330, 334, 343, 434, 524, 526, 527, 579, 612, 636, 656, 705, 735, 846, 979, 982, 1101, 1107 constructivist, xiv, xv, 157, 158, 159, 160, 161, 169, 170, 176, 178, 987, 988, 989, 991 consultants, 352 consulting, 398, 399, 416, 420, 561 consumer choice, 733, 788, 928, 977 consumer goods, 917 consumerism, 180 consumers, xxvi, 198, 281, 678, 682, 684, 736, 738, 768, 769, 771, 775, 786, 787, 789, 792, 795, 800, 801, 802, 943, 952, 977, 980 consumption, xx, xxiv, 64, 203, 212, 225, 235, 436, 452, 504, 507, 511, 512, 514, 576, 606, 608, 650, 651, 686, 688, 694, 699, 707, 708, 714, 716, 717, 719, 752, 753, 754, 755, 756, 757, 758, 759, 760, 762, 763, 764, 800, 802, 823, 885, 900, 909, 915, 916, 924, 938, 980, 1087 contamination, 791 content analysis, xix, 391, 400, 404, 846, 1001, 1002, 1003, 1009, 1010 context, 261, 267, 268, 477, 480, 484, 490, 498 context-dependent, 221, 256, 598, 642, 675, 925 context-sensitive, 930 contingency, 547, 551 continuity, 596, 640 contraceptives, 395, 411, 506, 514, 576, 578, 583, 584, 592, 622, 631 contracts, 714, 715, 900, 938

control condition, 543, 682, 689, 727, 728, 729

control group, 82, 166, 518, 520, 538, 776, 777, 778, 779, 780, 781, 783, 785, 955, 956, 957, 969 controlled, 483, 1105 controlled trials, 139, 775 conventionalism, 560, 565 convergence, 27, 739 conversational norms, 968 conversion, 38 convex, 199, 595, 597, 598, 599, 639, 641, 642, 643, 891, 893, 896, 898, 900, 903, 959, 965, 966, 976, 977 conviction, 483, 1100 Cook, 264, 276 cooking, 445, 458, 463, 549 coordination, 487, 1104 Coping, 181, 481, 188, 498, 505, 947, 1042 coping strategies, 148, 180, 359, 498 coping strategy, 160, 385, 787 core, 264, 270 coronary heart disease, 486 corporal punishment, 482 corporate finance, 1111, 1112 corporate governance, xxx, 1099, 1100, 1101, 1104, 1111, 1112 corporate performance, 1101, 1104, 1111, 1112 correlation, 3, 42, 115, 139, 535, 562, 694, 740, 752, 756, 758, 759, 760, 762, 763, 764, 782, 803, 805, 809, 817, 819, 820, 833, 846, 847, 850, 854, 855, 871, 993, 1001, 1003, 1004, 1050, 1053, 1057, 1060, 1087, 1088, 1108, 1109, 1110 correlation analysis, 854, 855 correlation coefficient, 536, 694, 752, 756, 763, 764, 809, 846, 850, 1057 correlations, 26, 519, 526, 535, 563, 564, 565, 694, 696, 758, 846, 847, 848, 849, 853, 914, 963, 1002, 1003, 1057, 1060, 1069, 1108, 1109 cortex, xvi, 3, 197, 198, 205, 207, 210, 211, 213, 214, 215, 216, 217, 218, 219, 220, 221, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 237, 238, 240, 241, 242, 243, 244, 245, 247, 248, 249, 250, 251, 252, 253, 255, 256, 257, 258, 614, 615, 616, 617, 658, 659, 661, 932, 1116, 1118, 1119, 1127, 1128, 1130, 1131, 1132, 1133 cortical neurons, 241 cortisol, 485 cost accounting, 1102 cost analyses, 74 cost effectiveness, 73, 748 cost of living, 683 cost-effective, 72, 797, 1009 costs, xxvi, 15, 16, 64, 68, 70, 71, 79, 86, 95, 280, 287, 298, 332, 343, 381, 447, 576, 597, 600, 620, 640, 644, 664, 713, 716, 718, 726, 729, 785, 811,

873, 880, 926, 943, 947, 973, 974, 1036, 1074, 1077, 1078, 1079, 1100, 1104, 1106, 1116 counsel, 448, 449, 566 counseling, xx, xxii, 69, 72, 73, 328, 330, 332, 382, 385, 388, 390, 429, 434, 442, 443, 446, 448, 449, 450, 451, 452, 453, 457, 459, 460, 461, 464, 465, 466, 467, 469, 471, 472, 474, 557, 564, 565, 569, 576, 577, 585, 588, 589, 1009 counseling psychology, 388 couples, 329, 343, 347, 348, 353, 355, 362, 368, 369, 371, 373, 375, 378, 382 courts, 492 covariate, 550, 551 coverage, 264, 336, 575, 871, 936, 937, 938, 939, 943, 945, 946, 947 covering, xix, xx, 50, 391, 399, 504, 507, 576, 622 CPR, 185, 187, 189, 299, 304, 305 CPS, 715 crack, 535 crash rates, 842 craving, 250, 254, 255, 257 CRC, 42, 1029 creativity, 22, 162, 728, 986, 992, 993 creativity tests, 992 credentialing, 54, 58, 59 credibility, 352, 500 credit, 23, 107, 544, 545, 575, 700, 722 crime, 90, 91, 262, 392, 583, 902 crimes, 81, 85, 201, 458, 528, 529 criminal behavior, 81, 484, 519 criminal justice, 90 criminality, 90 criminology, 517 critical analysis, 925, 929 critical period, 394 critical thinking, 1043 critical value, 1084 critically ill, 291 criticism, 76, 121, 147, 202, 363, 481, 718, 895, 912 crops, 799, 809 cross-country, xxv, 751 cross-cultural, 295, 568 cross-cultural comparison, 657 cross-sectional, 8, 84, 99, 108, 111, 119, 182, 331, 464, 471, 506, 514, 1107 cross-sectional study, 111, 464, 471, 506, 514 cross-validation, 78, 80, 88, 89 crowding out, 947 crying, 415, 439, 443, 479, 494 crystallization, 558, 1008 CSF, 1132, 1133 cues, 255, 284, 314, 461, 473, 479, 482, 664, 665, 672, 673, 1071

cultural, xx, 478, 503, 513, 629, 632 cultural differences, 52, 119, 298, 864, 942 cultural factors, 558, 791 cultural norms, xviii, 297, 788 cultural values, 281, 306, 307, 478 culture, 53, 55, 57, 58, 96, 181, 238, 286, 298, 300, 303, 308, 309, 381, 397, 399, 401, 402, 403, 407, 419, 426, 429, 430, 448, 559, 565, 568, 573, 577, 611, 655, 774, 788, 789, 795, 797, 869 cumulative, 265 curiosity, 341, 511, 987, 992 currency, 594, 616, 638, 659, 756 current account, 977 Current Population Survey, 720 curriculum, 360, 538, 622, 990, 991, 992, 999, 1005, 1009 customers, 415 cycles, 240, 241, 250, 378 cyclical process, 162 cycling, 879 cyclists, 1078 cynicism, 266 Cyprus, 96 cystic fibrosis, 312, 371, 379, 386 cystitis, 525 cytomegalovirus, 96 cytosine, 313

### D

daily living, 115, 447, 462, 484, 937 Dallas, 470 damage, 269 dandelion, 1015 danger, xxvi, 153, 202, 208, 385, 399, 464, 544, 830, 841, 843, 845, 853, 854, 855, 856, 857, 858, 859, 860, 861 dangerousness, 78, 82, 83 data analysis, xviii, 311, 404, 405, 708, 862 data collection, xix, 122, 356, 391, 399, 402, 403, 404, 405, 440, 507, 520, 521, 826, 1002, 1052, 1068, 1069, 1075 data processing, 404 data set, 26, 27, 319 database, 47, 55, 185, 740, 862, 993 dating, 458, 623 David Hume, 708 DCA, 830, 840 DD, 614 deafness, 380 death, xi, xxvii, 5, 6, 7, 8, 53, 65, 68, 184, 194, 281, 282, 283, 286, 287, 289, 290, 291, 299, 302, 303, 307, 313, 314, 317, 327, 333, 351, 362, 364, 374,

385, 392, 393, 427, 428, 478, 481, 493, 499, 515, 561, 586, 600, 605, 643, 649, 671, 772, 827, 831, 836, 837, 842, 867, 868, 871, 873, 876, 879, 880, 884, 886, 939 death penalty, 307 death rate, 831, 871, 886 deaths, xi, 5, 45, 68, 104, 305, 318, 827, 828, 842, 863, 871, 946 debt, 303, 350, 752 decay, 602, 604, 646, 648, 1024, 1025, 1026 deciliter, 94 decision, 264, 267, 273 decision control, 1110 decision makers, xii, 10, 13, 25, 26, 70, 195, 283, 289, 314, 344, 357, 396, 399, 598, 642, 726, 730 Decision Support Systems, 14 decision task, 230, 687, 890, 896, 898, 900, 901, 905 decision trees, 357 decision-making process, xxvii, xxxi, 12, 76, 81, 108, 174, 235, 284, 306, 315, 345, 351, 357, 370, 372, 373, 397, 398, 402, 405, 416, 558, 572, 573, 574, 632, 774, 788, 844, 845, 853, 855, 861, 868, 870, 883, 1102, 1115 declarative memory, 1023, 1024, 1026 decoding, 254 Decoding, 258 decomposition, 33, 759 deconstruction, 1042 decoupling, 246 deductible, 736, 737 deductibles, 737 deductive reasoning, 1043 deep-sea, 868 defects, 388 defense, 732 defense mechanisms, 732 defensiveness, 731 deficiency, 252 deficit, 205, 209, 219, 230, 498, 729, 962, 1130 deficits, 81, 209, 228, 229, 233, 313, 484, 485, 1131, 1133, 1134 definition, 67, 86, 106, 116, 117, 184, 315, 343, 362, 401, 416, 477, 479, 518, 621, 622, 623, 624, 625, 626, 632, 672, 803, 804, 869, 904, 985, 986, 991, 992, 994, 995, 1002, 1006, 1008, 1048, 1085, 1103 deflator, 756 deformities, 94 degenerative disease, 313 degradation, 481 degrading, 481 degree, 263, 478, 527, 624, 827, 828, 830, 831, 832, 833, 834

degrees of freedom, 172 delinguency, 498, 513, 538 delivery, 53, 155, 169, 252, 397, 408, 409, 410, 417, 424, 425, 429, 430, 432, 434, 436, 438, 439, 444, 446, 449, 450, 451, 467, 471, 474, 514, 608, 651, 652, 772, 793, 794, 795, 996 Delphi, xiv, 157, 159, 170, 171, 172, 178, 181, 196 demand, 115, 116, 117, 120, 164, 176, 185, 207, 267, 290, 292, 307, 381, 544, 547, 576, 592, 736, 775, 791, 800, 926, 991, 1087, 1106 dementia, xv, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 294, 313 democracy, 14, 989 democratization, 577 demographic, xx, 504, 507, 820, 823, 833 demographic characteristics, xx, 89, 402, 403, 406, 463, 504, 507, 776, 777, 857, 907 demographic factors, 578, 687, 842, 945 demographic transition, 573 demographics, 629, 687, 812, 833 denial, 283, 362, 477, 481, 555, 732 Denmark, 269 density, xxviii, 26, 755, 951, 959, 963, 964, 965, 966, 967, 968, 976, 977, 1028, 1088, 1133 dentate gyrus, 243 Department of Commerce, 748, 840 Department of Health and Human Services, 496, 514, 515, 537, 539, 621, 622 dependent, 738, 740, 744, 746, 813, 814, 816, 817, 827, 834, 836 dependent variable, 82, 83, 84, 550, 563, 740, 744, 746, 781, 813, 814, 816, 817, 834, 1053, 1062, 1083 depersonalization, 528 deposition, 94 depressed, 133, 217, 329, 413, 415, 424, 434, 435, 439, 446, 483, 484 depression, 89, 111, 162, 167, 252, 258, 267, 271, 284, 286, 313, 433, 437, 458, 469, 481, 482, 483, 484, 486, 489, 497, 519, 522, 523, 524, 527 depressive disorder, 469 depressive symptoms, 484 deprivation, xxiv, 156, 265, 266, 268, 284, 476, 497, 707, 715, 719 dermatosis, 525 desire, xx, 13, 16, 70, 141, 169, 282, 285, 299, 301, 322, 327, 342, 457, 459, 462, 466, 553, 558, 578, 581, 585, 601, 645, 687, 708, 714, 727, 728, 731, 792, 843, 847, 851, 854, 855, 859, 942, 1047, 1067, 1094 desires, 107, 281, 300, 306, 360, 479, 573, 580 destruction, 524, 718

detection, 207, 222, 233, 241, 244, 255, 343, 387, 388, 482, 519, 896 detention, 268 determinism, 349, 378 deterrence, 453 developed countries, 757, 868 developing countries, xxii, 392, 394, 571, 575, 757, 870, 872 development, 262, 276 development policy, 999 developmental change, 559 developmental delay, 484 developmental disabilities, xiii, 75, 79, 80, 90, 478 developmental milestones, 483 developmental psychopathology, 496 developmental theories, 1036 deviation, 138, 222, 477, 605, 648, 688, 738, 757, 925, 927, 1052, 1084, 1088, 1089, 1106, 1107, 1108, 1109 diabetes, 95, 125, 151 diagnosis, 267 diagnostic, 476, 481, 494, 520 dialysis, 188, 301, 308 Diamond, 496, 715, 722 diamonds, 958 diaphragm, 584 diarrhea, 99, 101, 189 dichotomy, 176, 352, 366 diet, 140, 169, 470, 610, 653, 1017, 1021 dietary, 330, 952 dieting, 1119 diets, 467, 1017, 1018 differences, 267, 272 differential diagnosis, xxi, 517, 526 differentiation, 242, 341, 344, 890, 1039 diffusion, 799 dignity, 108, 291 dimensions, 266 diplomas, 318 direct measure, 686 direct observation, 11, 482, 543 directives, xvii, 58, 176, 194, 196, 279, 280, 282, 285, 290, 291, 293, 295, 297, 298, 303, 308, 309 disabilities, 263, 269, 478 disability, 38, 68, 72, 89, 90, 114, 115, 135, 136, 154, 156, 194, 269, 392, 393, 863, 880 disabled, 79, 81, 155, 194, 269, 483, 496, 842, 876 disappointment, 283, 419, 1068 disaster, 340, 698, 738, 748, 827, 829, 830, 905 discharges, 259, 1078 discipline, xxiv, 55, 151, 287, 460, 479, 482, 487, 488, 495, 620, 707, 719, 991 disclosure, 8, 299, 300, 303, 480, 490, 491, 493, 494

discordance, xxiii, 619 discount rate, 65, 66, 68, 609, 610, 653, 656 discounting, 1, 64, 65, 66, 68, 72, 73, 74, 307, 600, 601, 602, 603, 604, 605, 609, 610, 611, 613, 614, 615, 616, 617, 644, 645, 646, 648, 653, 654, 655, 657, 658, 659, 660, 923, 1120 discounts, xxix, 72, 708, 939, 1013, 1014, 1022 discourse, xxi, 155, 261, 262, 263, 312, 531, 533, 535, 536, 537, 539, 579, 582, 796, 844, 927 Discover, 61 discretionary, 685, 690, 691, 700, 701 discretionary spending, 685, 690, 691, 701 discrimination, 85, 302, 314, 336, 543, 914, 917, 928, 1024, 1025 discriminatory, 453 disease free survival, 108 disease gene, 332 disease progression, 136, 148, 316 diseases, 32, 68, 106, 125, 189, 191, 192, 380, 482, 535, 538 disinhibition, 313 disorder, xviii, xx, 78, 87, 90, 94, 109, 187, 192, 230, 264, 267, 268, 311, 312, 313, 314, 337, 338, 346, 350, 351, 355, 357, 379, 380, 469, 475 dispersion, 758 displacement, 732 disposition, 50, 60 disputes, 800 dissatisfaction, 46, 164, 165, 179, 181, 584, 943, 946, 947, 979 disseminate, 326 dissociation, 232, 617, 661 distance education, 1010, 1011 distance learning, 985, 1001, 1002, 1004, 1005, 1006, 1007, 1009, 1010 distortions, 199, 901, 908, 956, 969, 1064 distraction, 60, 89, 204 distress, 101, 102, 186, 291, 327, 329, 368, 374, 376, 380, 393, 459, 489, 567, 568, 771 distribution, xii, xxii, xxviii, 25, 26, 27, 28, 29, 32, 33, 34, 37, 41, 67, 70, 89, 224, 349, 505, 557, 605, 648, 681, 684, 688, 698, 699, 737, 738, 740, 741, 743, 744, 747, 754, 803, 817, 897, 903, 918, 924, 927, 939, 951, 957, 959, 961, 962, 963, 964, 966, 967, 968, 969, 970, 976, 977, 1021, 1029, 1047, 1050, 1083, 1088, 1089 distribution function, 817 distributive justice, 70 disutility, 891, 892, 893, 901, 1082, 1084 diurnal, 525 divergence, 244, 518, 718, 826, 827, 828, 995 diversification, 698, 704, 918, 929

diversity, xvii, 2, 11, 52, 57, 161, 227, 297, 298, 307, 353, 462, 533, 575, 990, 1002, 1041 dividends, 756 diving, 868 division, 529, 582, 984, 1006, 1027, 1102 divorce, 323, 340, 344, 458, 506, 709, 729 DNA, 313, 335, 383, 388, 389, 492, 500, 674 DNA testing, 383, 389 doctor, 504, 527 doctor-patient, 165, 793 doctors, 5, 6, 49, 68, 107, 118, 119, 126, 127, 146, 163, 167, 178, 281, 294, 295, 301, 305, 399, 403, 453, 504, 525, 768, 785, 788, 791 dogs, 1017, 1029 do-it-yourself, 263 domain, 534, 535 domestic violence, 477, 478, 481, 482 dominance, xxix, 217, 386, 892, 896, 951, 973, 974, 975, 976, 977, 1040 dominant strategy, 736 donor, xiii, 93, 95, 96, 97, 98, 104, 105, 108, 110, 348 donors, xiii, 93, 94, 95, 96, 97, 100, 101, 105, 108, 110 doors, 1034, 1037, 1039, 1042 dopamine, xvi, 213, 214, 234, 235, 237, 239, 240, 241, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 1118, 1120, 1127, 1130, 1131, 1132, 1133, 1134 dopaminergic, xvi, xxxi, 197, 212, 213, 214, 218, 220, 225, 244, 253, 254, 258, 617, 661, 1115, 1118, 1120, 1122, 1127, 1131, 1132, 1134 dopaminergic modulation, 617, 661 dopaminergic neurons, 212, 254 dorsolateral prefrontal cortex, 205, 614, 658, 1118, 1119, 1128 dosage, 30, 583 DOT, 861 downward comparison, xxix, 951, 971, 977, 1067 draft, 17, 192, 521, 673 dream, 414, 525, 587 drinking, 437, 460, 470, 507, 510, 586, 687, 855 Drought, 283, 284, 286, 293 drowning, 478 drug abuse, 87, 259, 459, 460, 461, 462, 464, 465, 468, 473, 474, 481 drug addict, 221, 224, 225, 229, 258, 412, 869 drug addiction, 229, 258 drug consumption, 507, 512, 514 drug trafficking, 586 drug use, xx, xxi, 250, 457, 459, 460, 461, 462, 466, 468, 469, 472, 473, 474, 479, 484, 511, 531, 538, 621, 1123

drugs, xvii, xx, xxi, 3, 104, 107, 109, 212, 226, 228, 246, 247, 249, 250, 279, 281, 388, 412, 419, 422, 424, 459, 460, 462, 465, 468, 471, 504, 507, 511, 513, 531, 533, 535, 587, 614, 657 DSM, 460, 473, 486, 519 DSM-IV, 473, 486 dual task, 211, 1042 duplication, 879 duration, 6, 38, 40, 68, 71, 159, 435, 485, 524, 604, 606, 612, 648, 650, 655, 753, 871, 879, 927, 936, 939, 1020, 1027, 1031, 1050, 1065 duties, 53, 58, 264, 306, 350, 564, 579, 1051, 1105 dynamic environment, 925 dysfunctional, 1100, 1104 dysphoria, 250 dyspnea, 99, 101 dysregulation, 256, 485, 486 dysregulations, 485 dysthymia, 524

## Ε

earnings, 713, 717, 740, 1103 ears, 247, 1038, 1040 earthquake, 830 East Asia, 303 eating, 432, 900, 937 ecological, xxix, 2, 32, 41, 227, 461, 603, 605, 609, 612, 617, 646, 649, 652, 656, 660, 907, 925, 1013, 1031 Ecological models, 612, 648, 652, 656 ecologists, 664 ecology, xxiii, 234, 469, 498, 594, 596, 599, 600, 609, 610, 617, 637, 638, 640, 643, 644, 652, 653, 656, 660, 665, 1029 economic, xxv, 269, 620, 735, 752, 758, 765, 822, 1100, 1101, 1112 economic behaviour, 758, 916, 919 economic crisis, 415, 424, 575 economic damages, 822 economic growth, 575 economic performance, 1101 economic problem, 410, 415, 679, 916 economic status, 993, 994 economic theory, 679, 686, 708, 716, 758, 912, 915, 916, 919, 924, 928 economically disadvantaged, 533, 576 economics, xvi, xxiii, xxiv, xxvii, 15, 71, 237, 532, 594, 600, 603, 614, 615, 637, 638, 644, 646, 657, 658, 660, 678, 679, 687, 704, 707, 708, 709, 710, 711, 714, 715, 716, 717, 718, 719, 736, 770, 869, 893, 911, 912, 913, 915, 916, 923, 924, 928, 929, 930, 932, 982, 986, 997, 1007, 1026, 1074

economy, 526, 753, 758 ecosystems, 870 Eden, 773, 794 EDSS, 38 educational attainment, 532, 537 educational background, 189 educational policy, 622 educational process, 56 educational research, 1002 educational system, 453 educators, 464, 623, 635, 1066 EEG, 255 effective, 266 efficacy, xx, 475, 621, 629 efficiency level, 1087 egg, 549 eggs, 615 ego, 18, 567, 568, 727, 728, 729, 731, 732 egocentrism, 861 Egypt, 302 elaboration, 460, 524 elasticity, 753, 758, 763 elderly, 119, 151, 188, 189, 294, 300, 301, 303, 304, 429, 467, 470, 472 elders, 273, 304 election, 253, 466, 930, 1031 elective surgery, 152 electricity, 575, 691, 701 electronics, 1107 electrophysiological properties, 240 electrophysiological study, 259 electrophysiology, 3 elementary school, 505, 1009, 1010 elephant, 248 eligibility criteria, 627, 776 email, 93, 157, 531, 635, 889, 941, 943, 1115 emancipation, 788 embryo, 378 embryos, 378 emergence, 17, 29, 546, 549, 628 **Emergency Department**, 47 emerging markets, xxv, 751, 752, 756, 757 emigration, 671, 1133 emission, 207 emotion, xvi, 1, 16, 17, 18, 168, 175, 197, 200, 201, 202, 203, 204, 205, 207, 209, 213, 214, 217, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 239, 240, 241, 252, 315, 376, 393, 485, 486, 573, 574, 774 emotion regulation, 230, 485, 486 emotional, xvi, xvii, xx, 3, 6, 7, 51, 52, 69, 99, 101, 105, 146, 166, 175, 177, 186, 194, 197, 198, 199, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210,

211, 212, 214, 215, 216, 217, 218, 219, 220, 221, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 249, 259, 273, 279, 281, 284, 286, 290, 313, 321, 339, 341, 344, 359, 368, 376, 377, 451, 458, 459, 475, 476, 477, 480, 481, 482, 483, 484, 485, 486, 488, 493, 496, 498, 499, 560, 573, 708, 715, 800, 844, 902, 957, 1127, 1134 emotional abuse, 476, 482, 499 emotional conflict, 481 emotional experience, 105, 211 emotional health, 483 emotional information, xvi, 197, 207, 210, 212, 214, 215, 216, 217, 218, 220, 225 emotional processes, 203, 207, 210, 211 emotional reactions, 7, 204, 205, 220, 234, 359, 498 emotional responses, 206, 207, 708, 715 emotional state, 202, 203, 204, 205, 212, 230, 376 emotional stimuli, 207, 208, 219, 220 emotional well-being, 321, 482 emotions, 3, 16, 17, 46, 53, 57, 60, 160, 166, 168, 201, 202, 203, 204, 205, 207, 208, 212, 214, 216, 217, 219, 220, 224, 226, 227, 231, 232, 233, 340, 347, 385, 394, 412, 428, 485, 486, 573, 869, 870, 902, 906, 1065 empathy, 109, 182, 228 emphysema, 66, 184 Empirical, 816, 818, 824, 839 empirical methods, 80 employees, xxii, 46, 52, 59, 303, 557, 717, 1103, 1105 employers, xxii, 411, 557, 581, 585, 587, 708, 713, 714 employment, xxii, 41, 59, 87, 133, 140, 162, 169, 263, 343, 368, 453, 458, 459, 557, 580, 582, 620, 781, 846, 851 employment status, 133, 140, 846, 851 empowered, 266, 271, 328, 791 empowerment, 263, 266, 269, 276 encapsulated, 375 encoding, 243, 253, 1023 encouragement, 381, 454, 478, 489, 532, 543 endangered, 481 endocrine, 94, 485 endocrine glands, 94 end-of-life care, xvii, 7, 279, 292, 295, 298, 299, 300, 304, 308, 309, 310 end-stage renal disease, 466, 474 endurance, 263 energy, 107, 177, 204, 274, 594, 599, 600, 603, 604, 612, 638, 643, 644, 646, 647, 655, 805, 882, 1017, 1026, 1107, 1130 engagement, 34, 82, 167, 230, 344, 377, 386, 483, 579, 580, 624, 1002, 1132

engineering, xxv, 735, 1107 England, xiv, 18, 152, 154, 157, 159, 265, 337, 382, 384, 454, 497, 722, 789, 795, 862, 932, 940 English, 264, 533, 1110 engraftment, 99, 104 enrollment, 462, 533 enterprise, 272, 708, 1099, 1100, 1101, 1102, 1103, 1104, 1107, 1110, 1111, 1113 enthusiasm, 55, 467 entorhinal cortex, 243, 244 entrapment, 19, 728, 730, 731, 732, 733 entrepreneurs, 732 entrepreneurship, 1112, 1113 envelope, 739 environmental, 482, 632 environmental change, xxix, 218, 1013, 1015 environmental characteristics, 88 environmental context, 169, 170, 926 environmental economics, 708, 714, 718, 1074 environmental factors, 34, 402, 403, 664 environmental impact, 1084 environmental resources, 460, 474 environmental stimuli, 251, 926, 1116 epidemic, 194, 393 epidemiological, xxi, 517, 518, 527 epidemiology, 32, 384, 532, 538 epilepsy, 244 episodic, 738 episodic memory, 1032 epistemological, 155, 518 equality, 298, 451, 714 equilibrium, 6, 924, 928, 930, 1113 equipment, 428, 868, 870, 872, 874, 875, 876, 877, 879, 880, 881 equities, xxiv, 687, 688, 707 equity, xxv, 59, 714, 717, 751, 752, 753, 755, 756, 758, 763, 764, 765, 943, 946, 1065, 1106, 1107, 1108, 1109, 1110 ERIC, 1010 erosion, 266, 672 erythrocyte, 95 estimating, 136, 871, 873, 938, 1058, 1083, 1107 estimator, 31, 1113 estrogen, 1134 Estrogen, 1133 ethical, 492, 493, 622, 629 ethical concerns, 629 ethical issues, xiv, 94, 106, 108, 281, 283, 378 ethical principles, 288, 328, 492 ethical questions, 321 ethicists, 184, 185, 192, 298 ethics, xiv, xvii, 23, 24, 94, 107, 121, 188, 195, 292,

297, 299, 306, 317, 333, 381, 387, 405, 500, 620

ethnic background, 194, 299 ethnic groups, xviii, 297, 298, 300, 305, 476, 478, 482 ethnic minorities, 267 ethnic minority, 300 ethnicity, 52, 60, 162, 193, 300, 308, 309, 310, 463, 464, 534, 634 ethology, 931 etiology, xxi, 80, 254, 470, 477, 504 etiquette, 54 eugenics, 382, 389 Euler, 753 Euro, 1119, 1123, 1126, 1127, 1128 Europe, xi, xvi, 5, 14, 261, 264, 265, 267, 290, 305, 306, 314, 800, 809, 877 European, 264, 274, 275, 276, 1111 European Americans, 300, 304, 305 European Commission, 700 European Union (EU), 264, 800 Europeans, 276, 304, 305 euthanasia, 130, 291, 292, 295, 301 evacuation, xxvi, 748, 811, 812, 814, 817, 820, 821, 823 evening, 435 everyday life, 262, 265 evidence, xxv, xxviii, xxx, 264, 265, 476, 480, 481, 485, 490, 492, 493, 500, 520, 532, 609, 611, 620, 621, 622, 623, 629, 632, 679, 694, 698, 699, 708, 709, 712, 713, 716, 726, 731, 751, 752, 759, 764, 801, 805, 833, 911, 913, 915, 916, 917, 918, 919, 922, 925, 928, 944, 1076, 1099, 1101, 1110, 1111 evolution, 6, 32, 38, 39, 40, 41, 42, 258, 528, 569, 600, 603, 605, 609, 610, 612, 614, 617, 644, 646, 648, 649, 652, 653, 656, 658, 660, 661, 674, 926, 989, 1031 evolutionary process, 672, 675 exaggeration, 905, 906, 907 examinations, 500, 992 exchange rate, 1079 exchange rates, 1079 excitement, 262, 264 exclusion, 80, 362, 368, 389, 537, 575, 1066 excuse, 23, 50 execution, 217, 219, 1066 executive function, 217, 218, 219, 221, 223, 226 executive functions, 217, 218, 219, 221, 223, 226 executive processes, 234 exercise, 66, 89, 137, 140, 148, 151, 156, 160, 169, 170, 263, 270, 272, 289, 291, 323, 330, 350, 461, 566, 758, 800, 884 exertion, 873 expansions, 313 expectations, 271, 538, 539

expenditures, 415, 685, 690, 700, 701, 936 experimental condition, 547, 894, 1066 experimental design, 678, 710, 736, 1075, 1077, 1080, 1093, 1094, 1123 expertise, 107, 192, 264, 342, 374, 381, 518, 542, 543 experts, 477, 828 exploitation, xxxi, 302, 476, 479, 489, 524, 1115, 1117, 1125, 1130 exposure, xx, 31, 195, 250, 251, 349, 393, 477, 479, 485, 486, 499, 503, 514, 515, 566, 581, 665, 686, 694, 736, 800, 826, 827, 828, 829, 830, 833, 834, 835, 842, 870, 871, 876, 1017 expression, 483 external environment, 164 external influences, 988 external locus of control, 366 external validity, 78, 80 extinction, 1024, 1025, 1031 extra help, 508, 509 extraversion, 89, 252 extrinsic rewards, 1066 eye contact, 52 eyes, 242, 438

### F

fabric, 328 facial expression, 52, 186, 208, 209, 229, 523 facilitators, 12, 169, 453 factor analysis, xxvi, 84, 91, 841, 849, 851, 853, 854, 857,863 factorial, 1054, 1061, 1077, 1080, 1093 factors, 261 factual knowledge, 373 failure, 21, 37, 39, 42, 77, 84, 95, 99, 167, 200, 267, 270, 271, 273, 314, 438, 476, 477, 482, 498, 522, 577, 583, 584, 585, 726, 727, 728, 730, 731, 732, 752, 771, 785, 788, 843, 869, 870, 872, 879, 893, 894, 906, 984, 993, 1004, 1006, 1010, 1034, 1066, 1070, 1112 fairness, xxiv, 16, 17, 707, 709, 713, 719 faith, 127, 137, 146 false, 480, 490 false belief, 544, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555 false positive, 86, 544 familial, 330, 380, 480, 483, 519, 527, 565, 579 families, 265 family conflict, 393, 458, 489, 560, 564, 567 family environment, 481, 485, 559 family functioning, 484

family history, 312, 315, 316, 318, 321, 326, 327, 329, 335, 336, 338, 353, 380, 485 family income, 400, 403, 436, 833 family life, 331, 334, 385, 387, 410, 413 family members, xvii, xviii, 162, 186, 192, 193, 194, 273, 279, 284, 285, 288, 290, 291, 292, 294, 300, 303, 306, 311, 312, 313, 315, 316, 317, 318, 319, 321, 324, 326, 327, 328, 330, 332, 337, 338, 345, 347, 349, 351, 353, 354, 356, 361, 363, 367, 370, 372, 373, 374, 375, 376, 397, 407, 420, 421, 426, 480, 483, 486, 489, 560, 579, 942, 946 family planning, xxii, 394, 396, 571, 573, 575, 576, 577, 583, 587, 589, 591 family relationships, xviii, 311, 327, 328, 559, 560 family structure, 356, 507 family support, 105, 356, 507 family system, 334, 353, 388, 484 family therapy, 345, 355, 489 family violence, 482 families, 265, 267, 270, 272, 273 farmers, 716 fat, 412, 418, 437, 439 fatalistic, 560 fatalities, 842, 861, 1077, 1078, 1087 fatigue, 53, 99, 101, 106, 124, 130, 134, 148, 482, 584, 876 faults, 379, 421, 573 fax, 995, 996, 1115 fear, xxvii, 7, 16, 59, 135, 138, 164, 165, 166, 201, 202, 204, 206, 207, 208, 209, 227, 228, 229, 233, 254, 264, 267, 270, 274, 282, 283, 284, 314, 325, 334, 363, 386, 479, 480, 566, 569, 587, 738, 828, 859, 867, 869, 883, 902, 903, 906, 993 fears, 167, 374, 910 February, 61, 62, 178, 455, 514, 634, 636, 810, 947, 948, 949, 1011 feces, 323 Federal Emergency Management Agency (FEMA), 826 federal funds, 298 federal government, 620, 623, 827 federal law, 589 fee, 425, 444, 710, 877 feedback, xiv, 2, 59, 69, 157, 159, 173, 175, 176, 201, 206, 341, 342, 357, 466, 467, 468, 484, 622, 623, 630, 728, 732, 1002, 1003, 1005, 1046, 1050, 1051, 1054, 1056, 1057, 1058, 1064, 1065, 1067, 1068, 1069, 1070, 1071, 1123 feeding, 184, 195, 239, 284, 287, 298, 301, 304, 308, 479, 489, 523, 582, 609, 610, 653, 876 feelings, xx, 18, 60, 73, 167, 168, 182, 190, 202, 214, 216, 227, 232, 284, 285, 312, 315, 316, 317, 319, 320, 330, 348, 365, 374, 399, 401, 406, 410,

426, 440, 480, 483, 488, 503, 505, 573, 715, 779, 781, 785, 788, 921, 983, 988, 1004, 1065, 1074 fees, 594, 638 feet, 132, 881, 882 FEMA, 826, 833, 839 females, xxiii, xxxi, 51, 96, 98, 380, 394, 396, 459, 464, 466, 510, 513, 534, 563, 624, 625, 626, 631, 632, 663, 664, 665, 666, 667, 668, 669, 670, 671, 675, 834, 842, 843, 863, 1115, 1118, 1127 femininity, 580 feminism, 180 feminist, 179, 572, 590 Feminists, xxii, 571 fertility, xxii, 358, 362, 366, 382, 383, 387, 424, 571, 572, 573, 575, 576, 578, 579, 580, 581, 582, 585, 590, 591 fertility rate, 575, 576 fertilization, 378, 584, 674 fetal, 343, 350, 383, 386, 387, 388 fetal abnormalities, 343, 383, 388 fetus, xv, 183, 350, 362, 368, 399, 418, 436, 437, 442, 584, 586 fetuses, 350, 351, 368 fever, 162 fiber, 472 fibers, 241 fibromyalgia, 499 fibrosis, 95, 96, 97, 312, 371, 379, 386 fidelity, 579 field theory, 921, 922, 931 Fife, 164, 179 filial piety, 559, 560, 561, 563, 564, 565, 566, 567, 569 Filipino, 539 film, 188, 191, 192, 193, 196 films, 187, 263 filters, 874 finance, 698, 913, 1111, 1112 financial loss, xxv, 725, 728, 729, 731 financial markets, 708, 758 financial performance, 1104 financial problems, 410, 432, 435, 446 financial support, 235, 412 financing, xxviii, 935, 945 fines, 392 fingerprinting, 674 fingerprints, 500 Finland, 9, 1100, 1101, 1105 fire, 160 fires, 245 firm management, 1102 firm size, 1107, 1108, 1109, 1110

firms, xxxi, 716, 717, 912, 1099, 1100, 1101, 1103, 1104, 1105, 1106, 1107, 1110, 1111, 1112 first aid, 875 first language, 1037 first-time, 459 fish, 549, 1015 fishing, 1087 fitness, 140, 594, 600, 603, 605, 617, 638, 644, 646, 649, 660, 672, 870, 874, 876, 879, 880 fixed costs, 1106 fixed rate, 1026 flavors, 973 flexibility, xxx, 86, 121, 125, 202, 215, 216, 217, 218, 375, 490, 716, 789, 927, 946, 992, 993, 1073, 1074, 1075, 1076, 1077, 1078, 1079, 1080, 1083, 1085, 1089, 1090, 1095 flight, 208, 973 flood, xxvi, 161, 738, 811, 813, 815, 816, 818, 819, 820, 821, 822, 823 flooding, 813, 822, 823, 1000 flow, xii, xxvii, 45, 241, 246, 247, 249, 255, 357, 574, 584, 777, 867, 868, 869, 870, 873, 874, 880, 881, 882, 1017, 1102, 1103 fluctuations, 240, 258, 259, 722, 752, 1017, 1021, 1066, 1132 fluid, 18, 492, 1132 fluid intelligence, 18 fluoxetine, 1132 fMRI, xi, 1, 2, 4, 207, 208, 213, 215, 217, 218, 220, 227, 230, 231, 232, 243, 244, 250, 251, 257, 1130, 1132 focus group, 56, 146, 154, 169, 192, 406, 409, 418, 810, 845 focus groups, 56, 146, 154, 192 focusing, 28, 170, 186, 189, 259, 264, 265, 298, 339, 343, 347, 459, 461, 547, 685, 890 follicular, 1122, 1123, 1124, 1132 food, 139, 188, 212, 241, 256, 257, 363, 422, 424, 477, 519, 581, 582, 594, 600, 603, 605, 606, 607, 608, 610, 613, 616, 617, 638, 643, 647, 649, 650, 652, 653, 656, 657, 660, 667, 668, 685, 700, 800, 801, 829, 830, 870, 872, 877, 879, 881, 882, 952, 980, 987, 1014, 1015, 1016, 1018, 1019, 1020, 1021, 1022, 1023, 1024, 1025, 1029, 1031, 1107, 1116, 1119, 1123, 1131, 1132 food industry, 801 food intake, 257, 603, 647, 952, 980 footwear, 875 Ford, 395, 454, 577, 585, 592 forebrain, 252, 256, 257 forecasting, xii, 25, 27, 28, 41, 69, 306 foreclosure, 562 foreign travel, 870, 872

forensic, xiii, 75, 78, 85, 87, 89, 477, 490, 491, 492, 495, 518, 520 Forest Service, 13 forestry, xi, 9, 11, 12 forgetfulness, 190 forgetting, 199, 926, 1024, 1026, 1029, 1126 formal education, 119, 575, 576 Fox, 347, 383, 729, 895, 898, 899, 900, 905, 909, 1017, 1030 fracture, 68 fractures, 68, 478, 485, 496 framework, 262, 263, 268, 269, 270 framing, 71, 73, 196, 200, 201, 209, 220, 226, 234, 286, 370, 578, 617, 661, 678, 696, 712, 715, 797, 805, 894, 896, 898, 902, 907, 909, 910, 1043 France, 62, 264, 265, 505, 517, 518, 520 free choice, 682, 683, 688, 689, 692, 693 free recall, 490 freedom, xxii, 106, 109, 121, 138, 172, 262, 263, 266, 341, 346, 350, 363, 395, 413, 557, 564, 580, 591, 843, 991, 1004, 1078 freedom of choice, 346, 350 freedoms, 481 frequency distribution, 630 frequency theory, 681, 697, 699, 705, 921, 922, 927, 928, 931, 955, 964, 978 Freud, 479 friction, 876, 882, 912 Friday, 423 friends, xix, 52, 265, 273, 330, 336, 392, 393, 396, 407, 409, 410, 414, 418, 419, 420, 422, 423, 424, 425, 429, 430, 432, 435, 436, 437, 439, 440, 441, 443, 444, 445, 446, 447, 483, 506, 507, 510, 512, 513, 514, 583, 586, 790, 847, 864, 873, 971, 989 friendship, 510, 512 frog, 543, 549 frontal cortex, xvi, 198, 207, 214, 215, 220, 223, 230, 237, 240, 241, 245 frontal lobe, 228, 234, 253, 1131 fruit juice, 2 fruits, 432, 594, 610, 638, 653, 708 frustration, xii, 45, 47, 60, 164, 167, 182, 373, 790 fuel, 60, 323, 1087 fuel type, 1087 fulfillment, 365, 460, 561, 580, 1051 functional changes, 461 functional imaging, 232 functional magnetic resonance imaging, xi, 1, 207, 230, 257 funding, 29, 272, 380, 577, 627, 628, 629, 948 funds, 298, 622, 634, 698, 709, 713, 918, 943, 946, 947 fungi, 1020

fusiform, 208 future myopia, 221 futures, 319, 322, 327, 343, 426, 583

### G

GABA, 255 GABAergic, 253 gait, 313 gambling, 2, 64, 65, 206, 207, 213, 217, 230, 232, 234, 250, 257, 259, 262, 611, 614, 655, 658, 681, 697, 730, 909, 919, 1131, 1133 game theory, 207, 234, 919, 920, 929 games, 1, 13, 200, 204, 213, 221, 222, 223, 224, 226, 335, 617, 661, 681, 709, 713, 868, 910, 912, 919, 920, 930, 932, 1009 gamete, 362, 366, 378 gametes, 349, 365, 369 gamma-aminobutyric acid, 258 ganglia, 225, 226, 239, 240, 241, 249, 1134 gas, 691, 701, 828 gasoline, 426 gastrointestinal, 94 gauge, xxiii, 619, 628 gay marriage, 620 gay men, 629 GDP, 756, 757 GDP deflator, 756 gender, 39, 51, 57, 58, 60, 70, 73, 114, 118, 119, 131, 143, 144, 145, 149, 150, 162, 169, 192, 193, 286, 294, 309, 353, 354, 366, 394, 451, 506, 508, 509, 510, 511, 536, 562, 563, 565, 576, 578, 579, 581, 582, 586, 589, 625, 626, 629, 630, 631, 632, 694, 746, 804, 806, 807, 808, 809, 834, 842, 859, 862, 942, 956, 1067, 1117, 1124, 1125, 1126, 1127, 1128, 1129, 1132, 1134 gender differences, 114, 118, 131, 143, 145, 149, 509, 510, 511, 564, 625, 629, 632, 694, 842, 862, 1067, 1117, 1124, 1127, 1128, 1129, 1132, 1134 gender effects, 294 gender equality, 451 gender identity, 581 gender ideology, 581 gender role, 576 gender-sensitive, 451 gene, 300, 313, 318, 321, 322, 333, 335, 336, 362, 371, 378, 379, 380, 388, 664, 979, 1127, 1133 genealogy, 333 general practitioner, 171, 182 general practitioners, 171, 182 generalizability, 79, 88, 331 generalization, 88, 231, 892, 897, 954, 957, 958, 980, 1022, 1029, 1050

generalizations, 300, 979 generation, 52, 187, 189, 191, 259, 316, 320, 321, 353, 354, 360, 361, 362, 365, 370, 372, 373, 375, 378, 490, 575, 581, 627, 672 generators, 343 genes, xxiii, 312, 338, 371, 380, 663 genetic, 485, 492 genetic counselling, 350, 352 genetic defect, 313 genetic disease, 312, 387 genetic disorders, 348, 380, 383, 1133 genetic endowment, 861 genetic factors, 338, 376 genetic information, 327, 332, 333, 352, 377 genetic marker, 492 genetic mutations, 94 genetic screening, 334 genetic testing, 312, 313, 315, 316, 318, 320, 321, 328, 329, 332, 333, 334, 335, 338, 340, 343, 351, 354, 355, 356, 362, 363, 365, 366, 367, 368, 369, 374, 375, 379, 383, 385, 388 genetics, xix, 252, 312, 315, 317, 319, 326, 327, 328, 330, 331, 332, 334, 337, 338, 349, 351, 354, 355, 357, 374, 386, 387, 388, 389, 390 Geneva, 863, 910, 948 genital herpes, 634 geographical variations, 35 geography, 471 Georgia, 580 geriatricians, 192 Germany, 237, 264, 265, 309, 350, 504, 720 germination, 987 gestation, xxv, 767 gestures, 847, 864, 1004 Ghost, 276 Gibbs, 42, 43 gift, 385, 563, 624 gifted, 190, 533, 1009, 1041 girls, 393, 405, 420, 439, 508, 509, 510, 511, 512, 513, 518, 520, 521, 538, 549, 580, 623, 859 glaciers, 874, 876 glass, 912 glasses, 912 global education, 538 global warming, 1076 globalization, 925 globulin, 97 globus, 239, 257 glucose, 882 glutamate, 250, 258 glutamatergic, 247 goal-directed, 238, 241, 248, 249, 254, 984, 1130 goal-directed behavior, 241, 254

goals, xi, xxxi, 9, 12, 17, 31, 64, 69, 106, 164, 167, 170, 187, 192, 217, 218, 243, 250, 269, 286, 289, 298, 299, 341, 344, 347, 365, 367, 373, 374, 447, 461, 477, 532, 558, 561, 565, 581, 686, 699, 716, 926, 982, 983, 984, 990, 991, 994, 995, 996, 997, 1004, 1005, 1006, 1007, 1008, 1017, 1047, 1048, 1070, 1115, 1130, 1132 God, 107, 301, 302, 324, 581, 588 going to school, 508 goods and services, 579 Gore, 316, 333, 458, 459, 469 gossip, 55, 412, 420, 421, 422, 423, 429, 432, 434, 575 governance, xxx, 1099, 1100, 1101, 1104, 1111, 1112, 1113 government, iv, xxii, xxv, 114, 116, 119, 176, 264, 266, 392, 395, 396, 399, 400, 411, 415, 417, 448, 450, 451, 488, 557, 575, 578, 620, 623, 678, 735, 769, 791, 801, 805, 812, 827, 831, 884, 939, 1064 government policy, 114, 119 GPs, 182, 264 grades, 131, 406, 460, 620, 991, 1004, 1037, 1038, 1039, 1050, 1051 grading, 980, 992 gradings, 262, 263, 264, 265, 269, 273 graduate students, 611, 655 grain, 262, 1018 grains, 987 grandparents, 581 grants, 589, 633, 634 graph, 38, 671, 741, 744, 745 grass, 139 grazing, 1029 Great Britain, 181 Great Depression, 719 Greece, 264, 541, 751 greed, 147, 172, 301 grief, 454 grounding, 353 group activities, 453 group identity, 726 group interactions, 238 group work, 272, 446 growth, xxi, 39, 84, 94, 272, 458, 504, 505, 541, 575, 752, 754, 755, 757, 758, 759, 760, 762, 763, 764, 801, 989, 1006, 1042, 1103 guanine, 313 guardian, 85, 493, 535 guessing, 229 guidance, xix, 85, 381, 391, 569, 1091 guidelines, xxvi, 12, 181, 290, 298, 332, 334, 354, 357, 386, 387, 401, 402, 404, 405, 451, 492, 622, 634, 768, 769, 793, 795, 811, 992

guilt, 284, 285, 303, 361, 367, 414, 480 guilty, 133, 283, 427, 432, 434, 443 gustatory, 242 GVHD, 97, 98, 99, 100, 102, 105, 106 gyri, 208, 224 gyrus, 223

#### н

H1, 836, 837, 838 habitat, 1017 hair loss, 164 haloperidol, 251, 253, 255 handicapped, 129, 498, 628 handling, 23, 58, 59, 69, 492, 983, 1021 hands, 431, 435, 877, 881, 882, 884, 900, 990 hanging, 130, 355, 579 haplotypes, 110 happiness, 201, 202, 231, 505, 978, 979, 1067 harassment, 46, 520 hardships, 140, 726 harm, xiv, 29, 46, 77, 83, 94, 107, 223, 230, 233, 264, 272, 286, 288, 308, 316, 478, 482, 801, 837 harmful, 266, 478, 621, 829, 835, 838 harmful effects, 569, 584, 802 Harvard, 71, 150, 183, 382, 390, 590, 720, 732, 863, 886, 932 harvest, 594, 638, 1014, 1023 harvesting, 1014, 1022 hazards, 384, 505, 825, 827, 831, 847, 848, 865, 873, 876, 904 head injury, 161 head trauma, 876 headache, 28, 584 healing, 95, 108, 151, 153, 301, 403 Health and Human Services, 496, 514, 515, 537, 539, 621, 622, 634 Health and Retirement Survey (HRS), xxviii, 935, 937 health care professionals, 71, 136, 143, 146, 280, 281, 285, 288, 289, 291, 350, 361, 373, 490, 575, 581 health care system, 194, 397, 454 health care workers, 299, 488, 589 health clinics, 482 health education, 71, 73, 151, 412, 622 health effects, 65, 72, 485 health expenditure, 936 health information, 162, 176, 336 health insurance, 477, 576, 937, 939, 941, 943, 947 health problems, xx, 65, 266, 270, 272, 273, 322, 343, 392, 395, 396, 397, 448, 458, 475, 482, 485, 489, 585, 621

health psychology, 155, 332, 334, 339, 377, 387 health services, xx, 265, 268, 274, 387, 395, 396, 397, 398, 399, 440, 448, 450, 463, 467, 475, 504, 513, 576, 589 health status, 102, 149, 385, 499, 826, 829, 836, 837, 838, 939, 942 heart, 81, 94, 99, 138, 142, 161, 203, 206, 229, 231, 232, 314, 322, 427, 485, 486, 871, 1131 heart disease, 186, 195, 322, 485, 486, 871 heart failure, 99, 314 heart rate, 161, 203, 206, 229, 231, 485 heat, 137, 251, 701, 828 hedonic, 199, 239, 256, 957, 963, 978, 979, 980, 1067 hegemony, 182, 263 height, 271, 349, 955 helix, 334, 386, 388, 389 helmets, 868, 876 helplessness, 284 hematological, 106 hematoma, 496 hematopoiesis, 104 hematopoietic, 110 hematopoietic stem cell, 110 hematopoietic stem cells, 110 hemisphere, 209, 210, 217, 219, 226, 245 hemodynamic, 229 hemoglobin, 94 hemorrhage, 99 hepatitis, 95, 96, 472 hepatitis C, 96 hepatomegaly, 95, 96 hepatosplenomegaly, 94 herbivores, 1017, 1018 heredity, 382 heroin, 535 herpes, 525, 634 heterogeneity, 33, 34, 79, 80, 462, 477, 704, 1088 heterogeneous, xiii, xxvi, 75, 77, 79, 80, 148, 468, 799, 805, 928 heteroscedasticity, 1088, 1107 heteroskedasticity, 1113 heterozygote, 1127 heuristic, 203, 389, 608, 651, 698, 884, 904, 909, 918, 922, 1034, 1035, 1036, 1038, 1039, 1040, 1041, 1042 heuristic processing, 1042 HHS, 633, 634 high blood pressure, 322 high resolution, 34 high risk, xiii, xix, xxiii, 38, 40, 41, 71, 82, 83, 84, 87, 94, 95, 105, 243, 266, 273, 315, 333, 351, 353, 354, 362, 368, 369, 393, 457, 459, 460, 462,

466, 467, 472, 473, 474, 483, 484, 487, 498, 513, 619, 699, 763, 801, 813, 817, 833, 842, 860, 861, 862, 864, 885, 1100, 1106 high school, xxix, 118, 189, 193, 318, 418, 423, 458, 459, 460, 462, 463, 464, 465, 466, 467, 468, 469, 470, 473, 474, 513, 535, 536, 567, 622, 636, 957, 981, 993, 1010, 1033, 1037, 1038, 1039, 1040, 1041 high scores, 846, 851, 993 high tech, 333, 385 higher education, xxii, 532, 538, 557, 821 high-frequency, 905 high-level, 242, 244 hip, 115, 118, 119, 135, 138, 139, 149, 151, 152, 153, 154, 156, 181, 451, 537, 718, 1110 hip replacement, 181 hippocampal, 244, 249, 253, 254, 255, 256, 257, 258, 259, 613, 656 hippocampus, xvi, 207, 214, 237, 238, 239, 240, 241, 242, 243, 244, 245, 247, 248, 249, 250, 251, 253, 254, 255, 258, 259 hips, 483 hiring, 716 Hispanic, xviii, xxi, 297, 298, 299, 300, 304, 305, 306, 309, 531, 533, 534, 536, 537 Hispanics, 304 histochemical, 239 histogram, 969 history, 263, 270, 274 HIV, xx, 369, 380, 382, 393, 450, 470, 472, 499, 503, 505, 506, 515, 534, 622, 628, 634 HIV infection, xx, 503, 505, 628 HIV/AIDS, 393, 634 HLA, xiii, 93, 95, 96, 100, 101, 104, 105, 110 HLA-B, 96 holistic, 169, 396, 1008 Holland, 81, 90, 562, 567, 680, 704, 764, 915, 930 home-care, 72 homeowners, 816, 833 homes, 271, 488, 738, 816, 818, 820, 823, 826 homicide, 264 homogeneity, 162 homogenous, 67, 247, 385, 1085 homozygote, 1127 honesty, 24, 167, 547 honey, 617, 660, 1014 honey bees, 617, 660 Hong Kong, 21, 567, 569 hopelessness, 299 horizon, 11, 67, 566, 685, 759, 762, 763, 1047, 1048, 1054, 1056, 1057, 1058, 1060, 1064, 1066, 1067, 1068 hormone, 229, 796, 1134

hormones, 588 horse, 127 horses, 1018, 1030 hospice, 189, 194, 291, 293, 300 hospital, xiv, 46, 49, 55, 56, 60, 85, 87, 88, 89, 113, 115, 119, 120, 146, 164, 167, 193, 268, 269, 271, 291, 307, 309, 351, 355, 392, 394, 415, 429, 438, 439, 467, 478, 526, 575, 576, 587, 771, 777, 787, 788, 789, 797 hospital beds, 394 hospital care, 119 hospitalization, 46, 81, 264, 266, 272, 394 hospitalizations, 87 hospitalized, 300, 309, 483 hospitals, xix, 47, 49, 85, 89, 266, 298, 392, 417, 436, 450, 451, 463, 585, 776, 777 host, 110, 624, 627, 729, 1028, 1034, 1037 hostility, xii, 45, 375, 478, 486 hotels, 870 House, 455, 470, 799, 839, 1009, 1010, 1011 household, 115, 485, 499, 575, 579, 580, 586, 587, 756, 804, 812, 814, 815, 816, 820, 821, 936, 941, 943, 946 household income, 575, 815, 816, 943 household tasks, 580 households, 582, 812, 815, 1087 housing, 83, 169, 574, 668, 669, 718, 811 HPA, 161, 167 HPV, 620 HRS, xxviii, 935, 937, 939, 946 HSC, 313, 314, 317, 319 human actions, 476 human agency, 574 human animal, xxxi, 600, 602, 644, 645, 1115, 1130 human behavior, xxiv, 22, 58, 707, 708, 709, 719, 826, 925, 1046 human brain, 4, 17, 228, 229, 233, 235, 252, 257, 499, 617, 1131 human capital, 65 human cognition, xxvii, 911, 913, 925 human condition, 153, 194 human development, 469, 477 Human Genome Project, 333, 385, 390 human motivation, 16 human nature, 708 human resources, 59 human rights, 629 human subjects, 601, 645 human welfare, 22 humans, xvi, xxxi, 11, 22, 42, 198, 209, 213, 228, 234, 235, 237, 239, 241, 242, 244, 249, 257, 259, 596, 603, 609, 611, 612, 640, 646, 652, 655, 827,

890, 896, 897, 899, 900, 905, 906, 1026, 1027, 1028, 1115, 1120, 1127 humiliation, 481, 519, 526 hunting, 1130 Huntington disease, xviii, 311, 312, 315, 331, 332, 333, 334, 335, 361, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390 Huntington's chorea, 381, 382, 384, 385 hurricane, xxv, xxvi, 735, 748, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823 Hurricane Andrew, xxv, 735 hurricanes, xxvi, 748, 811, 812, 813, 814, 815, 816, 818, 819, 820, 821, 822, 823, 836 husband, 120, 129, 136, 322, 323, 361, 371, 419, 424, 426, 429, 436, 579, 582, 585, 586, 587, 588, 785, 787 hydration, xvii, 279, 287, 305, 879, 882 hygiene, 482 hyperactivity, 230 hyperbolic, 602, 603, 604, 605, 611, 614, 615, 616, 617, 646, 648, 654, 655, 657, 658, 660, 923 hypersensitivity, 221, 224, 225, 525 hypertension, 486 hypnotic, 202 hypocrisy, 583 hypoparathyroidism, 95 hypothalamic, 485 hypothalamus, 208, 209, 210, 212, 214, 216, 220, 239 hypothesis, xvi, 26, 27, 30, 68, 69, 197, 205, 206, 209, 222, 228, 232, 234, 239, 246, 249, 250, 254, 468, 518, 537, 542, 551, 564, 565, 603, 604, 605, 609, 612, 617, 647, 648, 649, 652, 653, 656, 660, 693, 695, 698, 717, 786, 851, 863, 894, 923, 978, 993, 1001, 1026, 1031, 1039, 1058, 1066, 1084, 1111, 1120, 1122, 1127, 1130, 1134 hypothesis test, 30 hysterectomy, 772

### I

#### IASP, 180

iatrogenic, 165, 167 ice, 404, 609, 652, 868, 870, 872, 873, 874, 875, 876, 880, 881, 882, 886 ICU, 49, 309 id, xxvi, 126, 347, 369, 512, 525, 579, 635, 677, 738, 805, 852, 948, 1004 idealization, 912, 1041 identical sibling, xiii, 93, 95, 104 identification, 39, 41, 86, 175, 190, 200, 312, 352, 369, 374, 463, 492, 495, 579, 630, 680, 705, 708,

774, 836, 844, 860, 861, 909, 914, 915, 922, 927, 928, 931, 932, 1088 identification problem, 1088 identity, 17, 114, 125, 148, 163, 168, 179, 180, 265, 275, 303, 306, 325, 328, 329, 335, 342, 348, 349, 351, 365, 370, 372, 394, 395, 449, 494, 519, 524, 526, 528, 544, 558, 566, 567, 568, 579, 581, 591, 625, 714, 726, 732, 736, 739 ideology, xxii, 263, 571, 573, 578, 579, 581 idiographic approach, 319 idiosyncratic, 172, 912 IGT, 205, 206, 207, 209, 216, 217, 219, 221, 224, 1116, 1117, 1118, 1119, 1120, 1121, 1122, 1123, 1124, 1125, 1126, 1127, 1128, 1129 Illinois, 571, 613, 656, 1045 illusion, 203, 263, 353, 553, 714, 715, 1036, 1041, 1064 illusions, 885 imagery, 166 images, 172, 186, 188, 189, 191, 192, 193, 195, 199, 203, 220, 248, 249, 343, 344, 345, 348, 453, 860, 979, 1074, 1077 imagination, 190, 891 imaging, xi, 1, 207, 228, 230, 232, 235, 249, 252, 256, 257, 259, 1133 imaging techniques, xi, 1 imbalances, 768, 789 imitation, 228 Immanuel Kant, 990 immigrants, 300, 303 immigration, 576 immortality, 374 immune system, 161 immunohistochemical, 254 impairments, xvi, 237, 250, 313 implementation, xii, 21, 23, 83, 196, 298, 341, 343, 372, 373, 374, 462, 463, 558, 577, 773, 774, 795, 929, 982, 998, 999, 1023, 1102, 1112 impotence, 189 imprinting, 673, 674, 675 impulsive, 220, 226, 241, 243, 483, 607, 609, 610, 614, 615, 617, 651, 653, 657, 658, 661, 1118 impulsiveness, 216, 313, 523, 613, 656, 660, 844 impulsivity, xvi, 237, 239, 522, 614, 657, 658, 843, 844 in situ, 201, 204, 210, 211, 218, 219, 225, 226, 327, 482 in vitro, 253, 378 in vitro fertilization (IVF), 378 in vivo, 254, 255, 258 inadmissible, 493 inattention, 477, 482, 770 incarceration, 81

incentive, 226, 231, 251, 254, 256, 360, 378, 494, 577, 710, 920, 937, 942, 1104 incentives, 225, 228, 577, 681, 710, 739, 823, 919, 1104 incest, 479, 480, 517, 518, 519, 520, 524 incidence, 27, 46, 68, 87, 104, 105, 136, 150, 151, 350, 476, 482, 487, 523, 526, 576, 844 inclusion, 81, 192, 263, 287, 402, 404, 406, 550, 624, 774, 847, 973, 974 Income, 815, 816, 818, 819, 820, 821, 822, 834, 835 incomes, 304, 400, 415, 819, 821, 822, 943, 971, 972 incurable, 293, 473 indebtedness, 560, 564 indecisiveness, 566 independence, 280, 348, 458, 559, 560, 596, 597, 598, 611, 615, 640, 641, 642, 655, 659, 738, 801, 851, 857, 893, 925, 929, 1087, 1112 independent variable, 82, 550, 551, 563, 740, 744, 746, 816, 819 India, 303, 455, 475, 499, 515, 591 Indian, 298, 303, 308, 497 Indians, 828 indication, 7, 34, 291, 360, 365, 367, 369, 467, 731, 740, 762, 772, 1047 indicators, 37, 39, 68, 76, 80, 286, 398, 481, 524, 774, 849, 850, 852, 943 indices, 31, 489, 558, 559, 560, 687, 694, 756, 851, 852, 857, 920 indifference curves, 917, 925, 928 Indigenous, 564, 568, 767 indirect effect, 565 indirect measure, 1040 individual character, 770, 939, 941, 1073 individual characteristics, 770, 939, 941, 1073 individual development, 505, 560 individual differences, 1, 199, 234, 243, 363, 364, 380, 558, 610, 653, 1067, 1068, 1119 individual perception, 800 individual personality, 282 individual rights, 107, 299 individualism, 168 individualization, 573 Indonesia, 454 induced bias, 952, 966, 968 induction, 255 industrial, 13, 573, 932, 1107, 1112 industrial revolution, 1112 industrialization, 575 industrialized countries, 842 industry, 55, 57, 60, 678, 688, 716, 801, 936, 939, 941, 1106, 1107, 1108, 1109, 1110, 1113 inequality, 119, 169, 395, 712, 714

inequity, 1065

infancy, 238, 350, 476, 478, 479, 488 infants, 301, 309, 496, 498, 585 infection, xx, 138, 299, 314, 481, 497, 503, 505, 588, 629 infections, 467, 470, 480, 514, 534, 620, 628, 634 inferences, 69, 200, 542, 543, 545, 554, 555, 624, 1030, 1043, 1102 inferiority, 1004 infertile, 348, 363, 378 infertility, 346, 347, 348, 360, 386, 387 infinite, 759, 763 inflammatory, 499, 525 inflammatory bowel disease, 499 inflation, 616, 659, 685, 688, 714, 715 influence, 484 influenza, 162 informal sector, 402, 403 Information, 839 information exchange, 165, 768 information processing, 15, 16, 17, 66, 203, 207, 233, 366, 930, 1002, 1100, 1103 information seeking, 164, 294, 365, 366, 386 information sharing, 168, 773 Information Technology, 1009 informational access, 547, 551, 555 informed consent, 95, 98, 107, 289, 479, 563, 576, 585, 769, 774 infrastructure, 627 ingestion, 882 ingroup bias, 732 inheritance, xxiv, 350, 377, 379, 663, 673, 674 inherited, 315, 316, 331, 352, 362, 380, 664, 665 inhibition, 207, 210, 211, 215, 217, 218, 219, 220, 224, 225, 226, 234, 259, 524, 525, 1035, 1133 inhibitor, 253, 1132 inhibitors, 480 inhibitory, 31, 217, 218, 220, 226, 239, 1024 initiation, 505, 506, 514, 580, 629 injection, 127, 411, 438, 443 injections, 38, 127, 137, 576 injuries, 68, 198, 228, 471, 479, 482, 491, 520, 826, 842, 863, 876 injury, iv, xxvii, 83, 161, 164, 166, 289, 478, 480, 698, 826, 842, 861, 863, 867, 868, 870, 871, 873, 879, 884 inmates, 77, 80, 81 innocence, 495 innovation, ix, xxvi, 57, 185, 275, 276, 799, 800, 803, 804, 805, 806, 897, 912, 981, 999, 1001 input, xxi, 23, 51, 53, 160, 192, 201, 240, 242, 244, 249, 253, 257, 259, 368, 541, 885, 1005 insects, 610, 653, 1015, 1023 insecurity, 53, 525

insemination, 348 insertion, 451 insight, 22, 55, 120, 506, 543, 545, 548, 550, 715, 748, 784, 860, 924, 968, 977, 1104 insomnia, 101, 102, 106, 519 inspection, 505, 550, 552, 818, 846, 852 inspiration, 23 instability, xxx, 6, 522, 744, 844, 883, 1099, 1101, 1110 institutionalization, 81, 269, 270 institutions, xvii, 78, 297, 298, 302, 305, 307, 397, 402, 460, 461 instruction, 459, 461, 474, 583, 689, 690, 907, 930, 989, 992, 994, 999, 1000, 1009, 1010, 1125 instructional activities, 1000 instructors, 1003, 1004 instruments, xiii, 75, 76, 77, 78, 79, 81, 82, 187, 191, 402, 470, 518, 562, 752, 982, 993, 1001, 1010 insurance, xxvi, xxviii, 67, 140, 198, 326, 336, 463, 477, 576, 687, 691, 701, 735, 736, 737, 738, 739, 748, 771, 825, 828, 829, 830, 839, 890, 900, 901, 904, 907, 908, 909, 910, 935, 936, 937, 938, 939, 940, 941, 942, 943, 945, 946, 947 integration, xvi, xvii, 168, 204, 212, 215, 219, 222, 225, 226, 228, 256, 270, 276, 279, 461, 981, 983, 984, 985, 992, 996, 997, 1000, 1001, 1004, 1006, 1007 integrity, 18, 24, 167, 622, 733, 882 intellectual development, 524 intelligence, 18, 216, 971, 994, 1009, 1038 intensity, xxviii, 3, 47, 130, 159, 160, 163, 167, 180, 206, 216, 218, 219, 233, 248, 485, 525, 527, 679, 680, 805, 811, 812, 844, 873, 904, 911, 913, 914, 915, 922, 979, 985, 1078, 1101, 1106, 1107, 1108, 1109, 1110 intensive care unit, 189, 301, 309 intentionality, 553 intentions, 52, 343, 345, 347, 360, 369, 538, 555, 582, 624 interaction, 3, 56, 68, 82, 91, 121, 147, 155, 164, 203, 213, 234, 251, 253, 257, 344, 367, 397, 398, 399, 401, 416, 440, 444, 446, 447, 461, 482, 488, 498, 560, 567, 584, 674, 770, 873, 904, 917, 929, 987, 988, 989, 995, 996, 1001, 1002, 1003, 1005, 1010, 1056, 1058, 1059, 1061, 1062, 1064, 1080, 1093, 1094, 1095, 1103, 1123 interaction effect, 82, 1080, 1095 interaction effects, 82, 1080, 1095 interaction process, 416 interactions, xxi, 3, 52, 53, 72, 78, 136, 162, 164, 167, 168, 222, 223, 231, 238, 248, 249, 253, 254, 257, 258, 352, 353, 401, 408, 440, 478, 482, 483,

489, 520, 526, 531, 532, 536, 537, 667, 924, 995,

996, 1001, 1002, 1003, 1004, 1005, 1006, 1010, 1050, 1058, 1095, 1133 interactivity, 12, 1009, 1011 interdependence, 47 interdependent self-construal, 559 interdisciplinary, xi, 1, 3, 331, 383, 495 interest, 264 interest groups, 10, 14 interface, 254, 256, 674 interference, xx, 475, 524, 917, 1029, 1030 intergenerational, 559, 560, 565, 566, 568, 569 intermittent explosive disorder, 87 internal consistency, 404, 534, 791, 847, 855, 926 internal influences, 988 international, xxv, 532, 751, 752, 756 International Monetary Fund, 756 internet, 176, 185, 186, 193, 395, 455, 623, 941, 948, 983, 995, 1010, 1121 interneurons, 240 interpersonal communication, 990 interpersonal factors, 352 interpersonal interactions, 78 interpersonal relations, 165, 384, 394, 397, 451, 483, 484, 489, 559, 560 interpersonal relationships, 165, 394, 397, 483, 484, 489, 559, 560 interpersonal skills, 488 interpretation, 3, 30, 76, 121, 122, 147, 166, 167, 177, 220, 226, 227, 319, 376, 515, 623, 627, 630, 731, 773, 891, 917, 923, 987, 988, 1004, 1024, 1025 interstitial, 99 interstitial pneumonia, 99 interval, 10, 11, 13, 29, 30, 31, 34, 96, 375, 507, 512, 601, 602, 604, 606, 608, 645, 646, 647, 648, 650, 651, 757, 1015, 1020, 1023, 1029, 1031, 1078 intervention, xxv, xxvii, 68, 95, 109, 111, 117, 139, 140, 166, 177, 184, 185, 186, 188, 189, 191, 193, 195, 207, 221, 224, 266, 268, 269, 272, 286, 288, 289, 377, 463, 467, 468, 470, 471, 472, 478, 484, 487, 489, 500, 533, 544, 573, 578, 767, 776, 777, 778, 780, 782, 791, 805, 841, 860, 861, 988 intervention strategies, xxvii, 841, 860 interventions, 266 interview, xix, 7, 120, 121, 122, 129, 137, 139, 147, 294, 312, 318, 319, 320, 322, 323, 347, 354, 356, 391, 400, 401, 402, 404, 405, 410, 416, 430, 443, 449, 450, 452, 453, 464, 465, 466, 490, 491, 493, 521, 527, 627, 812 interview methodology, 627 interviews, xiv, xviii, xxvii, 11, 73, 113, 121, 122, 137, 139, 140, 142, 147, 150, 186, 188, 189, 192, 266, 301, 302, 311, 317, 318, 319, 335, 356, 375,

448, 449, 450, 451, 453, 464, 482, 490, 491, 500, 575, 624, 803, 812, 826, 867, 870, 873, 874, 882 intestinal obstruction, 99 intimidating, 490 intimidation, 883 intracerebral, 206 intravenous, 298 intravenous fluids, 298 intravenously, 97 intrinsic, 203, 256, 1070, 1077 intrinsic motivation, 203 introspection, 203 intuition, 22, 342, 708, 1034, 1035, 1039 invasive, 146, 148 inventors, 22 investment, xxiv, 64, 65, 222, 368, 505, 532, 575, 578, 582, 594, 638, 677, 678, 681, 682, 683, 685, 686, 687, 688, 689, 690, 693, 694, 696, 697, 698, 699, 700, 702, 705, 725, 726, 727, 728, 730, 731, 733, 748, 753, 826, 918, 919, 984, 1005, 1008, 1076, 1100, 1106, 1107 investors, xxiv, 698, 725, 726, 731, 752, 753, 1105 involuntary unemployment, 716 Iowa Gambling Task, ix, xxxi, 2, 205, 228, 1115, 1116, 1118, 1120, 1123, 1130, 1131, 1132, 1134 IQ, 1004 Ireland, 810 iron, xiii, 93, 94, 95, 96, 105, 437 irrationality, 227 irritability, 313, 1132 irritable bowel syndrome, 167, 169, 179, 180, 499 ischemic, 186, 195, 485 ischemic heart disease, 186, 195, 485 island, 248, 814, 816, 818, 820 isolation, xvi, xxx, 53, 132, 162, 167, 175, 237, 263, 271, 323, 346, 445, 458, 483, 484, 486, 487, 489, 498, 523, 525, 598, 642, 666, 732, 902, 1073, 1075 isoleucine, 1118 Israel, 302, 710, 981 issues, xvi, 261, 264, 265, 266, 269, 271, 274 Italy, xi, 5, 25, 34, 42, 93, 108, 264, 276, 979, 1041

376, 401, 404, 405, 406, 416, 418, 419, 440, 447,

#### IVF, 378

# J

JAMA, 60, 111, 182, 196, 499 January, 635, 812, 824 Japan, 302, 303 Japanese, 303, 308, 567, 611, 655, 675, 721, 1112 Jerome Bruner, 987 Jerusalem, 1009 job position, 463 job training, 463 jobs, xxii, 132, 348, 435, 459, 460, 461, 464, 465, 557, 573, 575, 880, 971 joints, 155 Jordan, vii, 180, 619, 624, 625, 626, 627, 635 judge, xxi, xxvii, 35, 66, 541, 581, 680, 681, 685, 699, 774, 845, 848, 855, 856, 865, 874, 889, 900, 904, 911, 914, 916, 927, 955, 958, 960, 968, 986, 990, 991, 1038, 1050, 1053, 1078, 1079 judges, 407, 845, 859, 982, 1049 judgment, xxviii, 16, 64, 66, 69, 76, 96, 106, 191, 281, 304, 347, 483, 546, 677, 680, 681, 682, 698, 699, 705, 711, 844, 901, 902, 903, 905, 907, 908, 909, 911, 913, 914, 915, 916, 917, 918, 919, 921, 922, 925, 927, 928, 931, 932, 933, 951, 952, 953, 954, 959, 963, 976, 978, 979, 980, 989, 991, 1007, 1042, 1043, 1047, 1048, 1049, 1050, 1052, 1053, 1064, 1071 judiciary, 488 junior high, 193, 470 junior high school, 193 jurors, 307 jury, 492, 495 justice, 70, 90, 166, 288, 289, 308, 495 justification, 23, 163, 202, 290, 352, 460, 576, 577, 582, 714, 730, 733, 786, 891

# Κ

Kaiser Family Foundation, 622, 627, 634 Kant, 987, 990 Katrina, 949 keiretsu, 1112 Keynes, 708, 714, 715, 721 kidnapping, 85 killing, 426, 833 kindergartners, 553 King, 213, 222, 231, 234, 308, 680, 705, 915, 931 kinks, 599, 643 kinship network, 299 knee arthroplasty, 149, 150, 152, 156 knee replacement, xiv, 113, 115, 123, 138, 139, 141, 150, 151, 152, 153, 154, 155, 156 knees, 129, 133 knowledge, 269, 274, 482, 483, 495 knowledge acquisition, 988 knowledge economy, 800 knowledge transfer, 924, 928 Korea, 568 Korean, 299, 300, 302, 560, 567

### L

labeling, 492 labo(u)r, xxii, xxiv, 399, 419, 453, 557, 573, 594, 638, 707, 708, 709, 714, 715, 716, 772, 776, 787, 788, 790, 794, 795, 1106 labor force, 399, 573, 715 laboratory studies, 710 lack of confidence, 369 lack of control, 1066 land, 302, 575, 884 Land, 838 language, xxii, xxiii, 1, 57, 58, 186, 188, 189, 193, 194, 244, 267, 269, 356, 461, 490, 523, 524, 571, 589, 619, 627, 628, 788, 789, 952, 953, 962, 963, 985, 1037 language barrier, 58 large-scale, 394 latency, 245, 257 late-onset, xviii, 337, 338, 339, 351, 352, 357, 380 later life, 351 laterality, 228, 1131 Latin America, viii, xxv, 272, 567, 575, 751, 752, 756, 757, 758, 763, 764 Latin American countries, 757, 758 Latino, 189, 462, 463, 464, 470, 534, 536, 537 Latinos, 194 laughter, 257 law, xvii, 90, 107, 252, 290, 297, 303, 359, 383, 390, 392, 395, 397, 399, 413, 415, 423, 434, 435, 445, 453, 458, 477, 478, 487, 492, 493, 495, 500, 521, 577, 580, 586, 587, 589, 605, 620, 648, 680, 705, 729, 843, 856, 913, 914, 932, 954, 957, 958, 960, 980, 1030, 1103, 1112 law enforcement, 477, 478, 487, 492, 843 law suits, 729 laws, xxvi, 198, 448, 482, 490, 492, 493, 603, 646, 841, 843, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 858, 859, 860, 864, 865, 989 lawsuits, 1110 lawyers, 495, 709, 884, 1040 lead, 263, 271, 272, 274, 477, 480, 482, 483, 488, 490, 491, 506, 518, 520, 525, 532, 620, 752, 815, 823, 825, 1104 leadership, 53, 56, 272, 495 leakage, 705 learning difficulties, 269, 380, 382, 1037 learning environment, 987, 988, 990, 991, 1001, 1002, 1004, 1005, 1008, 1010 learning outcomes, 1001, 1002, 1010 learning process, 31, 208, 524, 525, 981, 983, 984, 985, 986, 987, 988, 989, 991, 993, 995, 996, 997, 1001, 1003, 1005, 1006

learning skills, 458 learning task, 214, 232, 990, 1002 left hemisphere, 209, 217 legal, 262, 264, 267 legal issues, 195, 495 legality, xvii, 279, 290 legislation, 261, 265, 267, 577, 929 leisure, xx, 131, 132, 145, 263, 504, 507, 685, 700, 716, 885, 886 leisure time, xx, 504, 507, 716 lens, 169, 188, 193 lenses, 575 lesioning, 248 lesions, 4, 38, 205, 206, 215, 216, 219, 221, 222, 224, 225, 226, 227, 241, 243, 244, 252, 255, 257, 478, 1116, 1130, 1131 leucine, 1118 leukemia, 111 Levant, 91 liberal, 108, 270, 350 liberalization, 272 liberation, 264 libido, 162 lice, 521 licenses, 739, 740, 743 lien, 528 life changes, 163 life cycle, 65, 362, 722 life expectancy, 95, 98, 101, 102, 105, 186, 272, 936, 942,943 life experiences, 53, 57, 396, 461 life span, 134, 153, 569, 609, 653, 697, 960, 1036 lifestyle, 67, 69, 72, 413, 460, 461, 878 lifestyle decisions, 878 lifestyles, 468, 507, 678, 800 life-threatening, xiii, xix, 93, 95, 391 lifetime, xxviii, 68, 71, 312, 313, 394, 507, 686, 717, 935 light, 264 likelihood, xii, xxv, 25, 26, 27, 29, 32, 33, 46, 50, 60, 67, 117, 148, 160, 189, 199, 200, 201, 202, 204, 211, 212, 265, 299, 306, 314, 340, 380, 486, 490, 506, 768, 769, 772, 820, 821, 885, 902, 907, 920, 938, 939, 942, 943, 957, 958, 960, 1014, 1083, 1084, 1088, 1089, 1103 Likert scale, 47, 562, 846, 851 limbic system, 220, 246, 247, 256 limitation, 2, 16, 108, 147, 272, 339, 518, 536, 543, 572, 791, 1075, 1103 limitations, xii, 15, 16, 37, 76, 77, 81, 83, 89, 128, 156, 169, 227, 331, 339, 368, 370, 374, 468, 495, 499, 514, 566, 612, 656, 708, 869, 894, 913, 929, 1026, 1041, 1048, 1065, 1066, 1081

Lincoln, 121, 151, 155, 177, 178, 182, 387, 591 linear, xiii, 6, 14, 65, 75, 82, 83, 84, 341, 374, 594, 597, 599, 601, 638, 641, 643, 645, 744, 754, 756, 830, 834, 898, 960, 964, 965, 977, 1082, 1083, 1091 linear function, 594, 599, 638, 643, 744, 964 linear programming, 14 linear regression, 82, 83, 834 linkage, 46, 47, 60, 313, 354, 363, 367, 566 links, 13, 265, 328, 345, 359, 372, 461, 523, 845, 916, 1001 lipid, 524 liquefied natural gas, 828 liquor, 535 listening, 51, 223, 788 literacy, xvi, 184, 185, 186, 189, 193, 194, 196, 272 literature, xxv, xxx, 11, 263, 270, 495, 525, 536, 537, 602, 627, 679, 686, 696, 736, 751, 752, 753, 805, 1073, 1099, 1100, 1101 Lithuania, 272 litigation, xx, 475, 493, 771, 772 liver, 94, 96, 99, 110, 137, 301, 485 liver disease, 485 liver failure, 99 livestock, 987 loans, 715 lobbying, 272 lobectomy, 244 local community, 271, 739 localised, 43, 137, 139 location, 242, 405, 460, 466, 561, 748, 826, 942, 1001, 1021, 1027, 1116 locomotion, 239, 251 locus, 96, 299, 366, 377, 389 Locus of Control, 390 logical reasoning, 543, 554, 555 logistics, 403, 467 lognormal, 754, 1090 London, 152, 178, 181, 182, 230, 233, 265, 271, 272, 275, 276, 277, 293, 295, 332, 333, 334, 378, 381, 382, 383, 384, 385, 386, 387, 388, 389, 454, 496, 497, 504, 538, 554, 592, 636, 674, 675, 677, 700, 704, 721, 722, 764, 794, 796, 797, 799, 889, 909, 911, 929, 930, 931, 1011, 1096, 1097, 1111, 1112, 1131, 1133 loneliness, 483 long period, 250, 392, 443, 467 long run, 68, 647, 1015 longitudinal studies, 524 longitudinal study, xxx, 334, 472, 473, 1045 long-term, 266, 478, 484, 485, 486, 752, 1103 long-term care insurance, 937, 938, 943 long-term memory, 243, 248, 256, 924

long-term potentiation, 244, 249, 253, 257 Los Angeles, 62, 499, 567 loss of control, 351 losses, xxv, xxvi, 67, 73, 198, 199, 200, 205, 206, 208, 209, 211, 214, 215, 216, 218, 221, 227, 228, 230, 252, 348, 481, 594, 595, 598, 599, 603, 614, 638, 639, 642, 643, 646, 657, 698, 712, 717, 718, 735, 737, 738, 811, 834, 891, 895, 896, 897, 898, 899, 900, 901, 903, 905, 959, 962, 963, 966, 972, 976, 1104, 1118, 1122, 1123, 1125, 1126, 1127, 1129 lotteries, 687, 736, 897 Louisiana, 455 love, 132, 138, 262, 323, 374, 381, 413, 422, 423, 434, 454, 458, 482, 511, 535, 633 low back pain, 180, 181, 182 low risk, 34, 40, 84, 362, 368, 369, 452, 699, 817, 901 low temperatures, 599, 643 lower-income, 584 low-income, xix, xxii, 196, 391, 399, 400, 401, 402, 406, 424, 443, 447, 448, 571, 572, 574, 575, 581, 589, 590, 592 LTC, xxviii, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 949 LTD, 258 LTP, 258 luggage, 890 lung, 485 lung disease, 485 lying, 896, 927

# Μ

MacArthur Violence Risk Assessment Study, 78 machine learning, 246 machines, 282, 304 macroeconomic, 756, 926 macroeconomics, 594, 638, 659, 708 magazines, 145 magical thinking, 195 magnetic, iv, 252, 256 magnetic resonance, xi, 1, 207, 230, 252, 256, 257 magnetic resonance imaging, xi, 1, 28, 207, 230, 252, 256, 257 magnetoencephalography, 3 Maine, 622 mainstream, 79, 708, 709, 711, 716 maintenance, 10, 81, 168, 405, 472, 581, 691, 701, 769 major cities, xxix, 981 major decisions, 882 major depression, 89

major depressive disorder, 469 maladaptive, 161, 164, 167 males, xxvi, xxxi, 51, 96, 98, 331, 380, 393, 451, 459, 464, 466, 510, 513, 534, 563, 624, 625, 626, 631, 632, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 675, 676, 841, 842, 843, 844, 845, 853, 854, 855, 857, 859, 860, 861, 863, 1115, 1118, 1120, 1127 malignant, xiv, 94, 107, 108 malingering, 91 maltreatment, 476, 477, 479, 481, 482, 483, 484, 487, 488, 489, 493, 496, 497, 498, 500 mammals, 1015, 1038 mammography, 312, 316, 461, 469 management, xi, xvii, xix, xxvii, 9, 10, 11, 13, 15, 16, 32, 54, 56, 57, 72, 120, 148, 149, 151, 152, 153, 154, 166, 170, 180, 182, 211, 216, 269, 271, 272, 275, 276, 279, 281, 282, 291, 312, 316, 317, 328, 331, 392, 403, 451, 453, 463, 469, 471, 477, 489, 495, 814, 824, 826, 838, 867, 869, 870, 871, 873, 874, 875, 876, 877, 878, 881, 884, 885, 1002, 1102, 1103, 1111 management committee, 272 manipulation, 3, 584, 681, 699, 728, 736, 799, 901, 919, 1048, 1127 manners, 177 MANOVA, 563 manpower, 988 Maori, 273 mapping, 42, 73, 254, 293, 471, 680, 913 marginal utility, 64, 594, 595, 639, 884, 1082 marginalization, 165 marijuana, xxi, 459, 460, 462, 465, 468, 473, 507, 511, 531, 535, 536, 587 marital distress, 329 marital status, 87, 396, 463, 942, 946 market, xxii, xxviii, 2, 68, 557, 576, 678, 689, 699, 702, 708, 710, 712, 716, 718, 722, 727, 736, 738, 739, 744, 747, 752, 755, 756, 757, 758, 829, 832, 909, 935, 938, 941, 942, 946, 1103, 1105, 1106, 1113 market penetration, 946 market prices, 718, 738, 739 market share, 1103 market value, 1103 marketing, xvi, 16, 237, 452, 678, 800, 925, 1077 marketplace, 711, 909 markets, xxv, 678, 694, 699, 708, 710, 714, 751, 752, 756, 757, 758, 763, 764, 765, 907, 912, 924, 928, 1066, 1103 Markov, 27, 33, 38, 42, 72 Markov chain, 27, 33, 38, 42

marriage, xx, 247, 316, 322, 323, 324, 327, 329, 331, 345, 360, 382, 386, 387, 390, 393, 407, 419, 458, 459, 503, 505, 513, 561, 573, 579, 580, 582, 588, 620, 621, 629, 945 marriages, 580 married women, 407, 410, 411, 412, 414, 415, 420, 444, 576, 577, 580 marrow, vi, xiii, 93, 95, 96, 97, 100, 109, 110, 111 Mars, 51 Maryland, 464, 619 masculinity, 580, 860 mask, 815, 1120, 1123 mass, xx, 487, 503 mass media, xx, 487, 503 Massachusetts, 61, 183 mastery, 461, 465, 468, 542, 1047 maternal, 497, 537, 538, 771 Mathematica, 621, 635 mathematicians, 198, 1034 mathematics, 543, 554, 992, 1027 matrix, 87, 303, 402, 545, 546, 755, 756, 833, 846, 851, 852, 857, 920, 1113 maturation, xxix, 1033, 1036, 1038, 1040 McGill Pain Questionnaire, 130, 154 mean, 270, 816 meanings, xiv, 113, 121, 123, 147, 160, 169, 180, 286, 319, 357, 383, 408, 582, 952, 988 measurement, 10, 82, 154, 496, 609, 613, 616, 653, 657, 660, 704, 706, 718, 791, 826, 827, 830, 844, 845, 849, 850, 851, 852, 857, 860, 907, 930, 932, 982, 983, 984, 986, 990, 991, 1001, 1007, 1077, 1078, 1082, 1112, 1117 measures of perception, xviii, 311 mechanical ventilation, 187, 188, 189, 192, 301 media, xx, xxv, 130, 145, 169, 187, 192, 264, 267, 445, 487, 503, 581, 694, 735, 800, 860, 871, 885, 1107, 1110 medial prefrontal cortex, 213, 233, 252, 253, 257 median, 34, 36, 37, 68, 96, 97, 100, 393, 681, 833, 927, 955, 956, 1087 mediation, 256 mediators, 76, 499 Medicaid, 939, 943, 946, 947 medical care, xiii, xv, 8, 47, 68, 93, 95, 106, 107, 151, 152, 176, 183, 184, 185, 187, 194, 280, 286, 298, 301, 338, 454, 477, 768, 939 medical school, 53, 56 Medicare, 154, 937, 939, 946, 947, 949 medication, xix, 49, 77, 78, 137, 164, 299, 301, 391, 392, 394, 410, 430, 448, 461, 466 medications, xx, 77, 79, 83, 475, 478, 482 medicine, xiv, 42, 70, 71, 94, 95, 107, 108, 109, 151, 152, 169, 178, 180, 181, 182, 184, 185, 307, 333,

355, 385, 387, 389, 403, 425, 443, 448, 492, 495, 504, 525, 575, 768, 982 Medline, 459, 628, 635 MEG, 232 membership, xiv, 157, 159, 171, 1104 memory capacity, 1035, 1042 memory deficits, 1131 memory formation, 244, 249, 253 memory retrieval, 1029 men, 52, 115, 118, 119, 120, 128, 129, 132, 135, 136, 137, 143, 144, 145, 146, 147, 148, 149, 152, 266, 267, 302, 312, 322, 348, 351, 365, 367, 386, 393, 395, 407, 409, 412, 420, 422, 513, 563, 564, 565, 573, 577, 579, 580, 582, 624, 629, 635, 687, 694, 739, 746, 748, 805, 807, 833, 859, 861, 863, 942, 955, 957, 971, 1067, 1117, 1118, 1120, 1124, 1125, 1126, 1127, 1128, 1129, 1130, 1132, 1134 menarche, 505 menopause, 179, 361, 528, 796 menstrual cycle, xxxi, 583, 1115, 1118, 1120, 1122, 1123, 1124, 1131, 1132, 1133, 1134 menstruation, 396, 416, 417, 418, 419, 425, 432, 436, 439, 442, 450, 535, 583 mental disorder, 78, 90, 500 mental health, xvi, xx, 18, 76, 81, 82, 85, 90, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 457, 461, 464, 465, 466, 475, 477, 478, 481, 482, 485, 487, 488, 489, 733 mental health professionals, xx, 76, 82, 265, 267, 269, 274, 475 mental illness, 79, 80, 81, 263, 264, 265, 266, 268, 270, 272, 274, 486 mental image, 199 mental life, 555 mental model, 357, 382 mental processes, 17, 923 mental representation, 199, 612, 655, 697, 978 mental retardation, 91, 483 mental simulation, 341, 342, 373, 374 mental state, 226 mental states, 226 mergers, 1112 messages, 70, 71, 73, 148, 246, 249, 451, 554, 628, 629, 802, 828, 996, 1004 messenger RNA, 255 meta-analysis, 150, 344, 384 metabolic, 2, 211, 599, 643, 1124 metabolic rate, 599, 643, 1124 metabolism, 605, 609, 649, 653, 1134 metaphor, 166, 342 metaphors, 360

metastases, 189

metastatic, xviii, 293, 297, 300, 304 metastatic cancer, xviii, 293, 297, 300, 304 methamphetamine, 250, 251 methodological implications, 980 methodology, 515 Methotrexate, 73, 97, 98 metric, 1020, 1048 metropolitan area, 845, 846, 850, 857 Mexican, 299, 304, 534, 536, 537, 568 Mexican Americans, 304 Mexican-Americans, 299, 304, 568 Mexico, 300, 592, 748, 756, 757, 758, 759, 761, 762, 763, 867 Mexico City, 592 MHD, xxix, 1033, 1034, 1035, 1036, 1037, 1038, 1039, 1040, 1041 mice, 604, 605, 614, 647, 649, 658, 1117, 1127 microdialysis, 1134 microeconomics, xxiii, 234, 637 midbrain, xvi, 214, 237, 239, 240, 244, 245, 246, 248, 249, 254, 257 middle class, 549, 578, 581, 586 middle schools, xxi, 531, 532, 533, 636 middle-class families, 581 midwives, 785, 788, 791, 792 migration, 599, 643 migratory birds, 599, 643 military, 459, 575, 1064 milk, 436, 447, 586 mimicking, 480 minimum wage, 575, 586, 589 Ministry of Education, xxii, 235, 557 Minnesota, 90, 515 minorities, 196, 267, 300, 568 minority, xvi, 184, 185, 188, 193, 194, 248, 262, 265, 300, 318, 747 minority groups, 300 mirror, 525, 666, 792, 1020 miscarriage, 368, 438 miscarriages, 495 miscommunication, 165 misconceptions, 187, 506 misinterpretation, 52, 577 misleading, 33, 116, 547, 822, 823 missiles, 826 missions, 468 Mississippi, 968, 969, 970 Mississippi River, 968, 969, 970 Missouri, 85 misunderstanding, 69, 107 MIT, 254, 258, 554, 555, 929, 930, 931, 932, 1030, 1095, 1096 Mixed Logit, 1095, 1096

mixing, 12, 448 mobile phone, 800 mobility, 115, 132, 146, 162, 581, 714 modalities, xiv, 94, 208, 210 modality, 213, 460, 467, 901 mode, 268, 484 Model, 276, 743, 745, 816, 836 model specification, 1090 model system, 2 modeling, 32, 34, 37, 41, 73, 76, 84, 383, 486, 489, 538, 857, 959, 961, 967 moderates, 172 moderators, 76, 173 modern society, 1006 modernity, 262, 263 modernization, 569, 801 modulation, 209, 212, 221, 228, 229, 253, 254, 258, 259 mold, 52, 982 molecular mechanisms, 255 molecules, 95, 986 momentum, 291 money, xxv, 16, 19, 65, 67, 71, 128, 201, 210, 212, 218, 222, 241, 242, 250, 326, 411, 414, 415, 418, 422, 425, 426, 427, 430, 434, 436, 440, 441, 442, 444, 447, 580, 582, 585, 594, 638, 679, 685, 688, 694, 698, 701, 708, 709, 713, 714, 715, 725, 726, 727, 728, 729, 730, 731, 733, 736, 737, 739, 756, 880, 894, 917, 918, 922, 923, 924, 936, 948, 949, 956, 1084, 1117, 1120, 1122, 1123, 1124, 1125, 1127 money illusion, 714, 715 monitoring, 485 monkeys, 252, 253, 255, 257, 604, 605, 609, 616, 617, 647, 649, 652, 653, 660, 1031 monoamine, 1134 monoaminergic, 233, 1133 monograph, 497 monotone, 1078 Monte Carlo, 27, 33, 38, 42, 688 mood, 6, 53, 124, 130, 146, 148, 162, 199, 201, 202, 213, 217, 227, 842, 844, 846, 1068, 1123 mood disorder, 1123 moral behavior, 1005, 1008 moral judgment, 989 moral standards, 480 morale, 265 morality, 428, 620 morals, 333, 344, 421, 929 morbidity, xiii, 63, 68, 72, 94, 313, 526, 771, 772, 842 morning, 126, 411, 437, 508, 509, 620, 881, 973, 1023

mortality, xiii, xiv, 8, 47, 55, 63, 68, 72, 93, 94, 95, 98, 101, 102, 104, 105, 313, 379, 486, 538, 610, 654, 771, 823, 827, 842, 863, 1132 mortality rate, 104 mortality risk, xiv, 94, 98, 101, 102, 105, 538, 610, 654 MOS, 851, 857 Moscow, 505 motherhood, 383, 387, 388, 578, 579, 581, 582, 586 mothers, 316, 330, 348, 350, 380, 382, 383, 407, 426, 428, 459, 471, 483, 486, 489, 497, 498, 499, 539, 582, 633, 635, 772 motion, 208, 216, 486 motivation, xiv, xxvii, 16, 17, 18, 66, 67, 78, 82, 94, 95, 98, 99, 100, 102, 105, 108, 203, 211, 229, 241, 255, 256, 257, 315, 332, 334, 339, 340, 473, 474, 480, 560, 714, 726, 727, 728, 729, 730, 859, 862, 867, 914, 928, 988, 989, 1002, 1005, 1008, 1047, 1065, 1066, 1069, 1111 motives, xxvi, 10, 128, 622, 841, 844, 845, 846, 849, 850, 851, 852, 853, 854, 855, 856, 857, 860, 1048, 1070 motor activity, 220 motor skills, 90, 1124 motor system, 256 motor vehicle crashes, 842 mountains, xxvii, 867, 871, 874, 876, 880, 882, 883 mouse, 1123 mouth, 453, 626 movement, 46, 166, 266, 272, 298, 303, 313, 418, 620, 768, 788, 1021, 1106 MRI, xi, 1, 28, 207, 244, 256 MSC, 867 MSI, 79 multicultural, 307 multidimensional, 37, 99, 845, 847, 860, 906, 1009 multidisciplinary, xiv, 59, 157, 159, 171, 261, 355, 487, 488 multi-ethnic, 298 multiple factors, 50 multiple persons, xxx, 1045 multiple sclerosis, 34, 40, 42 multiples, 1079 multiplicity, 43, 53, 285, 339, 347, 372, 376 multivariate, xiii, 75, 76, 83, 85, 86, 118, 508, 563, 755, 812 murder, 85, 163, 583 muscle, 162, 313 muscular dystrophy, 368, 387 musculoskeletal, 125, 150, 156, 178, 180 musculoskeletal pain, 178, 180 music, 177, 848, 864

mutation, 312, 313, 321, 322, 325, 334, 351, 362, 367, 368, 380 mutations, 94, 338, 527 mutual, 266 myopia, 221, 660 myopic, 605, 612, 615, 649, 658

## Ν

NAc, 212, 213, 214, 222, 225, 227 narratives, xviii, 114, 121, 149, 150, 153, 166, 186, 192, 193, 297, 319, 321, 323, 324, 325, 337, 343, 345, 357, 362, 363, 376, 382, 575 national, xxiii, 476, 506, 619, 620, 622, 624, 627 National Academy of Sciences, 228, 229, 230, 232, 233, 674 National Health Service, 114, 115, 154, 163, 381 National Hurricane Center, 823, 824 National Institutes of Health, 538 national parks, 877 national policy, 12 National Science Foundation, 673, 951 National Youth Survey, 534 nationality, 300 Native American, 299 Native Hawaiian, 298 natural, xi, xxix, 9, 10, 11, 12, 13, 28, 32, 37, 38, 39, 40, 41, 124, 135, 151, 153, 189, 212, 259, 285, 343, 347, 370, 377, 384, 411, 572, 573, 580, 609, 610, 623, 652, 654, 665, 671, 683, 686, 758, 826, 827, 828, 829, 830, 831, 838, 870, 873, 880, 900, 901, 905, 924, 925, 926, 929, 966, 986, 987, 988, 1001, 1013, 1014, 1019, 1029, 1046, 1050, 1064, 1065 natural disasters, 826, 827, 900, 901 natural environment, xxix, 870, 873, 880, 924, 926, 987, 988, 1013, 1019 natural evolution, 38 natural gas, 828 natural habitats, 1029 natural hazards, 831 natural resource management, 10, 11, 12 natural resources, xi, 9, 10, 13, 838, 1014 natural science, 986 natural sciences, 986 natural selection, 1014 nausea, 99, 101, 411, 584 Nazi Germany, 350 NBC, 623, 627, 635 needs, 265, 269, 272, 476, 477, 478, 480, 482, 483, 484, 487, 488, 489, 494 negative attitudes, 396, 408, 444, 629, 859

1167

negative consequences, 165, 205, 210, 213, 216, 225, 426, 458, 576, 585, 628, 855, 880 negative emotions, 16, 202, 203, 204, 208, 217, 219 negative experiences, 582, 786 negative mood, 162, 201, 204 negative outcomes, 66, 186, 216, 244, 283, 617, 660, 843, 859, 901 negative relation, 409, 483, 537, 718, 1110 negativity, 210, 790 neglect, xx, 328, 475, 476, 477, 478, 479, 481, 482, 483, 484, 485, 486, 487, 489, 492, 493, 495, 496, 497, 498, 499, 500, 501, 544, 545, 546, 550, 551, 552, 555, 1036, 1040 negligence, 163 negotiating, xviii, 283, 337, 340, 350, 371, 634 negotiation, 168, 283, 287, 346, 347, 368, 373, 377, 477, 573, 578 neocortex, 258 neonatal, 301, 309, 772 neonatal intensive care unit, 301, 309 neostriatum, 258 nerve, 313 nerves, 161, 584 nervous system, 257, 298, 313, 485 nesting, 672 Netherlands, 264, 265, 292, 305, 386, 593, 634, 637, 721, 1115, 1132, 1134 network, xvi, 83, 119, 128, 160, 168, 207, 237, 248, 254, 272, 299, 343, 402, 451, 453, 460, 461, 526, 561, 912 networks, 274 neural function, 610, 653 neural mechanisms, xi, 1, 209, 238, 1029 neural network, 912 neural systems, 214, 228, 232, 234, 616, 659, 1133 neural tissue, 238 neuroanatomy, 3, 229, 231, 256 neurobiological, xvi, 237, 238, 1128 neurobiology, 252 neurochemistry, 3 neurodegenerative, 380 neurodegenerative disorders, 380 neuroeconomics, 254, 615 neuroendocrine, 161, 205, 499 neuroimaging, xvi, 2, 3, 207, 210, 227, 229, 231, 237, 239, 256 neuroimaging techniques, 2, 227, 239 neurological disease, 32 neurologist, 28 neuronal circuits, 206 neuronal plasticity, 259

neurons, xvi, 3, 212, 237, 240, 241, 245, 246, 247, 249, 250, 251, 253, 254, 255, 256, 257, 258, 259, 615, 616, 659, 1132 neurophysiology, 256, 258 neuropsychology, xii, 3, 15, 17, 231, 256 neuroscience, xvi, xxiii, 227, 229, 230, 231, 234, 237, 238, 246, 247, 248, 251, 594, 615, 637, 638, 658, 659, 1030, 1041 neuroscientists, xi, 1 neurotic, 127 neuroticism, 89, 233 neurotransmission, 251 neurotransmitter, xxxi, 1115, 1118 neurotransmitters, 3 neutrophil, 31 New England, 152, 384, 795, 940 New Jersey, 229, 617, 622, 660, 661, 794, 824, 864 New Mexico, 748, 867 New Orleans, 455 New South Wales, 275, 1099 New World, 610, 617, 653 New York, iii, iv, 14, 18, 61, 62, 71, 90, 111, 178, 180, 195, 227, 229, 230, 231, 233, 234, 253, 292, 308, 334, 382, 383, 384, 385, 387, 388, 389, 390, 500, 515, 538, 566, 567, 568, 569, 591, 592, 615, 616, 633, 634, 635, 658, 659, 674, 704, 705, 716, 720, 721, 722, 732, 733, 794, 795, 824, 838, 839, 840, 862, 864, 886, 907, 908, 909, 929, 930, 931, 932, 948, 978, 979, 980, 1008, 1009, 1030, 1042, 1043, 1070, 1095, 1096, 1111, 1131 New York Times, 635, 948 New Zealand, 266, 273, 275, 794 newspapers, 130, 145, 901, 905 Newton, 264, 1035, 1043 Newtonian, 912 next generation, 316, 320, 321, 365, 372, 378 NFI, 851, 852, 857, 858 NGO, 400, 453 NGOs, 392, 399 NHS, 116, 177, 182 nicotine, 249, 474 Nielsen, 196 nightmares, 519 NIS, 476 NIST, 748, 840 NMDA, 259 Nobel Prize, 711 nociception, 210 nodes, 38, 39, 1075 noise, 664 non-biological, 1028 non-clinical, 193, 769, 771, 775 nondisclosure, 327, 332

non-emergency, 286 nongovernmental, 576 nongovernmental organization, 576 non-human, xxiii, xxxi, 600, 602, 644, 645, 663, 1115, 1130 non-invasive, 148 non-linearity, 599, 643 non-renewable, 1017 non-smokers, 939 nonverbal, 52, 1001, 1004, 1010 nonverbal communication, 52 normal, xxix, 77, 105, 108, 109, 124, 128, 135, 137, 138, 140, 145, 148, 155, 205, 206, 207, 216, 219, 221, 233, 248, 263, 285, 313, 347, 348, 351, 352, 360, 361, 362, 374, 381, 421, 423, 436, 477, 480, 485, 505, 542, 549, 555, 585, 588, 605, 648, 737, 740, 747, 790, 803, 817, 872, 959, 961, 964, 965, 966, 967, 969, 977, 981, 1017, 1089, 1133 normal distribution, 737, 747, 803, 817, 964, 965, 966, 967, 969, 977, 1090 Normal Distribution, 961, 965, 967 normalization, 181 norms, xviii, xxii, 71, 79, 80, 111, 162, 198, 200, 223, 224, 297, 302, 306, 307, 347, 382, 397, 404, 419, 421, 426, 429, 448, 571, 573, 770, 788, 789, 859, 860, 929, 968, 1104 North America, 152, 264, 314, 471, 885 North Carolina, 622, 635, 811, 812, 814, 822, 840 Northeast, 576, 579, 590, 591 Norway, 96, 269 novel stimuli, 248 novelty, xvi, 165, 237, 241, 244, 246, 248, 249, 255, 258, 999, 1091 nuclear, 828 nuclei, 208, 209, 216, 220, 225 nucleus, xvi, 208, 210, 213, 214, 225, 229, 230, 231, 237, 238, 239, 240, 241, 242, 243, 244, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 1119, 1131, 1133, 1134 nucleus accumbens (NAc), xvi, 210, 231, 237, 238, 239, 240, 241, 242, 243, 244, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 1131, 1133, 1134 nudity, 525 null hypothesis, 27, 851, 1084 nurse, 46, 47, 49, 55, 56, 58, 132, 163, 171, 287, 295, 449, 463, 793 nurses, 47, 49, 51, 52, 56, 57, 58, 59, 171, 265, 281, 287, 386, 403, 450 nursing, 46, 55, 56, 58, 59, 136, 162, 188, 193, 196, 288, 294, 295, 301, 305, 492, 788, 792, 797, 936, 939, 942 nursing care, 936

nursing home, 136, 188, 193, 196, 301, 305, 936, 939, 942 nutrition, xvii, 196, 279, 301, 305, 461, 875, 879, 882 nutrition education, 196 nuts, 1014, 1030 nylon, 875

# 0

obedience, 559, 560, 561, 565, 566 obesity, 470 object recognition, 257 objective criteria, 117 objective reality, 177 objective tests, 142 objectivity, 12, 176, 178, 192, 1008 obligation, 160, 285, 312, 316, 322, 324, 325, 330, 559, 560, 565, 581, 988 obligations, xviii, 311, 312, 316, 322, 325, 328, 329, 330, 567, 569 observable behavior, 890 observations, xxvii, 11, 26, 37, 41, 55, 189, 241, 347, 348, 356, 369, 374, 448, 482, 669, 670, 672, 719, 740, 744, 756, 817, 882, 890, 904, 905, 939, 1016, 1028, 1050 observed behavior, 679 obsessive-compulsive, 313 obstetricians, 386, 577, 585, 784, 789, 793 obstruction, 99 occupational, 73, 171, 181, 558, 562, 565, 567, 936 occupational health, 73 occupational therapists, 181 occupational therapy, 936 odds ratio, 84, 508, 512 Oedipus, 526 offenders, 79, 80, 85, 86, 90, 91 Ohio, 569 oil, 137 oils, 137 Oklahoma, 826, 838, 1013 old age, 690 older adults, 114, 151, 154, 195, 937 older people, 688 olfactory, 110, 214, 257 Oman, 301, 309 omission, 253, 375, 476, 477, 478, 481, 1042 oncology, 31, 110, 293, 309, 469 one dimension, 11, 916, 973, 974, 975, 1020 online, 42, 110, 152, 182, 188, 196, 244, 331, 334, 372, 538, 690, 720, 919, 1002 oophorectomy, 333, 334 open-mindedness, 565

openness, 287, 560, 565, 993, 1102 operant conditioning, 212 operator, 463, 1031 opposition, 211, 307, 350, 360, 523, 582 oppression, 163, 302, 788, 886 optimism, 365, 726, 856, 859, 1071 optimization, 937 oral, xvii, 110, 185, 198, 234, 279, 290, 291, 302, 390, 411, 415, 480, 518, 522, 523, 524, 578, 583, 592, 622, 623, 624, 625, 626, 627, 629, 636 oral contraceptives, 411, 578, 583, 592 orbit, 242 orbitofrontal cortex, xvi, 3, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 223, 225, 226, 229, 231, 232, 233, 237, 238, 242, 243, 248, 249, 256, 257, 258, 617, 661, 1119, 1127, 1128 ordinal data, 1078 organism, 1116 organization, 55, 56, 59, 60, 99, 202, 242, 253, 257, 265, 267, 272, 384, 392, 447, 450, 451, 526, 586, 812, 932, 1030 organizational behavior, 733, 1104 organizational culture, 381 organizations, xii, 23, 45, 46, 56, 58, 59, 60, 175, 177, 186, 330, 392, 450, 451, 487, 506, 526, 576, 768, 873, 900, 1110 orgasm, 256, 629 orientation, 461, 466, 560, 565, 806, 859, 902, 1067, 1070, 1103 originality, 24, 993, 994 orthodox, 182, 919, 920 orthopaedic, xiv, 113, 116, 120, 139 oscillations, 241, 255, 256, 259, 753 osteoarthritis, xiv, 113, 114, 118, 119, 120, 123, 125, 126, 129, 130, 133, 134, 135, 136, 138, 139, 142, 143, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 179 osteoporosis, 68 osteosarcoma, 43 outcome of interest, 77, 84 out-of-pocket, 585 outpatient, 77, 78 outpatients, 167 output, xxix, 208, 238, 239, 240, 241, 259, 1005, 1008, 1013, 1015, 1016, 1024, 1027 outrage, 801 ovarian cancer, 315, 326, 332, 333, 380 ovarian cancers, 380 ovaries, 179, 312 overload, 94, 206, 343 overtime, 1049 overweight, 128, 748, 827, 896, 898 ownership, 273, 1105, 1106, 1107, 1110, 1112, 1113

ownership structure, 1112 oxygen, 876, 886 oxytocin, 251

#### Ρ

Pacific, 89, 298, 310, 455, 940, 944 Pacific Islander, 298, 310 Pacific Islanders, 310 Pacific Region, 455 packaging, 492 packets, 882 pain management, 170 pain score, 136 palliative, xvi, xvii, 6, 7, 190, 279, 280, 281, 283, 285, 286, 287, 288, 292, 293, 294, 295, 306, 474 palliative care, 6, 190, 280, 283, 285, 286, 292, 293, 294, 295, 474 paper, xxv, 262, 275, 527, 622, 633, 635, 736, 751, 753, 811, 812, 824, 826, Paper, 275, 633, 824 paradigm shift, 43, 472 paradox, 611, 655, 733, 893, 897, 914, 1111 paradoxical, 893, 919 parallel performance, 1048 paramedics, 403, 861 parameter, 26, 27, 67, 743, 745, 758, 764, 803, 816, 852, 898, 899, 906, 953, 954, 955, 957, 958, 961, 962, 964, 967, 969, 975, 976, 1082, 1084, 1085, 1087, 1088, 1089, 1090, 1126, 1127 parameter estimates, 852, 899, 1089 paraventricular, 208 parental care, 609, 653, 1021 parental control, 537 parental involvement, 481, 532, 539 parental support, 458, 559 parent-child, 482, 483, 484, 498, 532, 560, 569 parenthood, 348, 382, 383, 387, 459, 483, 515 parenting, xviii, 272, 325, 337, 351, 356, 357, 360, 363, 365, 366, 370, 372, 377, 380, 396, 398, 399, 401, 403, 405, 406, 482, 484, 487, 488, 489, 498, 532 parietal cortex, 219, 221, 226, 251 parietal lobe, 161, 224 parietal lobes, 161 Paris, 528, 529 Parkinson, 252, 375 Parliament, 797 parole, 529 partnership, 179, 269, 373, 533, 768, 769, 792, 793, 794 passenger, 1077, 1078

passive, xxii, 51, 148, 163, 173, 232, 280, 282, 283, 284, 346, 483, 560, 571, 573, 580, 774

pasture, 1018 paternalism, 107, 284, 295, 788 paternity, 672, 673, 674 path analysis, xxvi, 84, 841, 860 path model, 569, 857 pathogenic, 525, 526, 528 pathological gambling, 250 pathologists, 477 pathology, xiii, 93, 255, 268, 269, 520, 523 pathways, 161, 254, 342, 485, 486, 532 patient care, 45, 46, 47, 48, 49, 53, 55, 56, 57, 60 patient-centered, 303 pay off, 262 PCR, 96 PDG, 378 peak concentration, 245 peanuts, 1023 pears, 594, 638 pedagogical, 981 pedestrian, 843 pediatric, 467, 469, 478, 479, 483, 496 pedigree, 389 peer, xx, xxi, xxvii, 23, 412, 413, 458, 460, 463, 467, 470, 471, 480, 504, 507, 508, 511, 512, 513, 624, 860, 863, 867, 870, 884 peer group, 480, 860 peer influence, 863 peer review, 23 peers, xix, xxiii, 47, 49, 391, 394, 395, 397, 403, 408, 411, 413, 414, 418, 422, 448, 452, 465, 479, 480, 484, 504, 505, 506, 515, 537, 560, 619, 620, 715, 716, 717, 844, 856, 860, 971, 1053, 1070 pelvic, 498, 525 pelvic pain, 498 penalties, 215 penalty, 307, 392, 736 penetrance, 313, 380 penicillin, 302 penis, 623, 624, 625, 626 Pennsylvania, 622, 721, 737, 838, 1112 pension, 682, 684, 685, 690, 694, 701, 901 pension plans, 694, 901 per capita, 756, 757 perceived control, 66, 339, 794, 826 perceived outcome, 895 percentile, 40, 964, 1049, 1052, 1053, 1058, 1060, 1065, 1069 perceptual judgments, 680, 915 performance, xxx, 11, 508, 509, 533, 710, 728, 914, 1045, 1046, 1047, 1048, 1049, 1050, 1051, 1052, 1053, 1054, 1055, 1056, 1057, 1058, 1060, 1061, 1062, 1063, 1064, 1065, 1066, 1067, 1068, 1069,

1070, 1071, 1100, 1101, 1102, 1103, 1104, 1110, 1111, 1112 performers, 69, 1046, 1047, 1049, 1050, 1066 perinatal, 498, 771 periodic, 576, 738 periodicity, 752 peripheral nerve, 161 permit, 205, 361, 402, 404, 480, 524, 1015 perseverance, 216, 604, 647 personal, 263, 269, 274, 477, 479, 489, 497, 629, 631, 632, 685, 700, 731, 732, 756, 757, 803, 831, 936, 939, 943, 946, 947, 1048, 1050, 1071 personal accounts, 329, 353 personal communication, 343, 1050 personal control, 163, 291 personal efficacy, 831 personal history, 18, 398, 732 personal identity, 164, 179 personal problems, 479 personal relations, 873 personal relationship, 873 personal responsibility, 263, 322, 327 personal values, 60, 286, 769, 774, 775, 785 personality, xxi, xxvi, 1, 50, 57, 58, 60, 81, 87, 264, 267, 268, 282, 284, 313, 314, 473, 481, 517, 565, 566, 616, 659, 727, 728, 841, 842, 845, 855, 856, 862, 985, 988, 992, 998, 999, 1000, 1007, 1009, 1047, 1067, 1068, 1069, 1097 personality characteristics, 855 personality disorder, 87, 264, 267, 268, 314 personality factors, 862 personality research, 616, 659 personality traits, 50, 998, 999, 1097 personality type, 51, 1067 perspective, 263, 265, 267, 270, 477, 496, 497, 532, 537 persuasion, 384 perturbations, 255 Peru, 756, 757, 758, 759, 762, 763 pessimism, 299 pesticide, 415 PET, 207, 209, 230, 249, 1128, 1131 pets, 815, 818, 1017 PFC, 205, 206, 207, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227 pharmaceutical, 30 pharmaceutical companies, 30 pharmacies, 576 pharmacists, 47 pharmacogenetics, 1133 pharmacokinetic, 31, 43 pharmacokinetic parameters, 31, 43

pharmacological, 30 pharmacology, 614, 657 phenomenology, 233, 319, 342 phenotype, 671, 672 phenotypes, 672 phenotypic, 672, 675 phenotypic plasticity, 675 phenylketonuria, 312 Philadelphia, 195, 633, 737, 1112 Philippines, 455 philosophers, 198 philosophical, 96, 107, 117, 147, 163, 987 philosophy, 41, 121, 180, 181, 266, 269, 280, 283, 769, 992 phobia, 519, 527 Phoenix, 383 phone, 49, 157, 465, 471, 881 photoperiod, 666 PHS, 471 phylogenetic, 242 physical abuse, 476, 478, 479, 485, 486, 489, 496, 499 physical activity, 472 physical aggression, 513 physical environment, 266, 269 physical fitness, 870, 879 physical health, 458, 485, 486, 499 physical sciences, 708 physical therapist, 47 physical therapy, 936 physicians, xiv, xv, xviii, 7, 8, 29, 31, 47, 51, 53, 55, 56, 57, 58, 59, 61, 94, 102, 106, 107, 109, 111, 153, 167, 169, 179, 183, 184, 185, 187, 193, 194, 286, 287, 290, 295, 297, 300, 301, 304, 305, 306, 307, 309, 470, 477, 479, 505, 576, 577, 579, 585 physics, 238, 912 physiological, xvi, 3, 107, 160, 161, 164, 167, 203, 206, 234, 237, 334, 497, 519, 873, 1116, 1117 physiological correlates, 234 physiology, 167, 394, 664 physiotherapy, 122, 129, 137, 150 Piagetian, 524 pilot studies, 355 pilot study, 95, 474, 499 pitch, 872, 875 pituitary, 161, 485 pku, 15, 725 placebo, 30, 31, 82 placenta, 789 planets, 912 planning, xi, xv, xx, xxii, 9, 10, 12, 13, 14, 29, 42, 83, 183, 184, 185, 186, 187, 194, 195, 216, 281, 284, 286, 309, 316, 317, 330, 343, 347, 381, 394,

396, 460, 461, 464, 483, 503, 542, 558, 571, 573, 574, 575, 576, 577, 581, 583, 587, 589, 591, 823, 880, 882, 884, 936, 991, 997, 1009, 1023 planning decisions, 381, 574, 884 plants, 1015 plastic, 218, 664, 665, 666, 667 plasticity, 249, 250, 254, 259, 528, 664, 675 Plato, 987 plausibility, xii, 25, 26 play, xv, xvi, xix, 50, 51, 58, 76, 128, 131, 132, 140, 158, 159, 177, 197, 198, 208, 212, 218, 220, 223, 227, 282, 287, 298, 306, 313, 338, 354, 376, 450, 452, 453, 488, 489, 559, 560, 574, 595, 604, 639, 648, 673, 679, 681, 698, 727, 768, 773, 784, 786, 800, 828, 855, 918, 920, 926, 959, 962, 968, 1004, 1056, 1101 pleasure, 16, 64, 166, 198, 200, 226, 232, 594, 638, 705, 924, 927, 979 plurality, 628 pneumonia, 99, 360 poison, 1100, 1111 poisonous, 477 Poland, 1132 polarity, 628 polarization, 309, 385 polarized, 620 police, 415, 488, 490, 859, 861 Policy, 838, 839 policy makers, 68, 76, 298, 300, 622, 769, 990 policies, 261, 264, 266, 267, 271, 272, 275, 276 politeness, 1104 political, 267, 513, 620, 622, 629, 1110, 1112 politicians, 10, 264, 577, 579, 632 politics, 22, 23, 24, 182, 262, 513, 636 pollen, 1014 polygamy, 673 polygraph, 79 polymorphism, 1132 polynomial, 87 pools, 1036 poor, xxii, 5, 38, 42, 46, 49, 55, 56, 57, 72, 94, 98, 99, 101, 102, 103, 104, 106, 148, 216, 266, 267, 287, 298, 299, 325, 424, 430, 459, 480, 483, 485, 486, 513, 532, 571, 572, 578, 579, 580, 581, 589, 595, 621, 639, 843, 844, 870, 914, 922, 926, 1014, 1015, 1018, 1027, 1052, 1053, 1102, 1103, 1120 poor health, 325, 486, 532 poor performance, 216 population, xi, xiii, xxiii, 5, 27, 32, 34, 35, 64, 70, 75, 77, 79, 80, 81, 83, 84, 85, 89, 108, 111, 114, 115, 118, 119, 125, 138, 155, 165, 171, 196, 298, 299, 308, 330, 335, 343, 353, 380, 401, 405, 440,

462, 464, 467, 468, 471, 476, 487, 499, 505, 508, 523, 573, 574, 575, 576, 578, 590, 624, 626, 663, 664, 671, 673, 678, 682, 684, 687, 688, 694, 718, 748, 756, 803, 804, 805, 809, 831, 833, 835, 851, 852, 857, 986, 1005, 1009, 1086, 1087, 1088 population growth, 505, 575 pornography, 395 portfolio, xi, 9, 686, 704 portfolios, 717, 1100 positive attitudes, 404 positive correlation, 833, 871, 1003 positive economics, 929 positive emotions, 212, 214, 217, 232 positive feedback, 466, 467 positive interactions, 489 positive mood, 201, 202, 213, 217, 223 positive regard, 304 positive reinforcement, 577 positive relation, xxx, 408, 564, 834, 835, 837, 1099, 1101, 1107, 1110, 1111 positive relationship, xxx, 408, 564, 834, 835, 837, 1099, 1101, 1107, 1110, 1111 positivism, 177 positivist, 176, 177 positron, 207 positron emission tomography, 207 posterior cortex, 221 postmenopausal, 796 postmenopausal women, 796 postmodernism, 181 postpartum, xxv, 450, 577, 592, 767 postsynaptic, 255 post-traumatic stress, xx, 163, 475 post-traumatic stress disorder, xx, 475, 519 potassium, 258 poverty, 163, 270, 477, 478, 484, 487, 581, 588, 620 powder, 436, 447 power, xviii, 6, 27, 51, 52, 70, 86, 89, 106, 107, 108, 167, 190, 264, 265, 266, 282, 285, 297, 299, 326, 342, 385, 395, 397, 469, 505, 543, 575, 578, 612, 656, 680, 695, 698, 714, 753, 754, 768, 782, 784, 789, 828, 836, 913, 914, 987, 1043, 1074, 1075, 1091 power relations, 282 powers, 267, 913 pragmatic, 621 prayer, 301 predators, 600, 644, 1021 predictability, 212 prediction, xiii, xvi, xxi, 3, 32, 42, 69, 70, 75, 76, 80, 81, 83, 84, 85, 86, 88, 90, 91, 106, 209, 213, 214, 232, 233, 237, 244, 246, 247, 248, 251, 255, 256, 258, 317, 330, 473, 532, 548, 553, 565, 604, 608,

612, 648, 651, 655, 680, 683, 709, 716, 770, 794, 898, 901, 924, 992, 1019, 1036, 1042, 1047, 1054, 1120 predictive accuracy, 87 predictive model, 535 predictive validity, 88, 1057 predictor variables, 82, 84 predictors, xxi, 38, 39, 41, 83, 84, 116, 117, 118, 156, 231, 252, 293, 471, 531, 536, 567, 686, 697, 705, 806, 811, 820, 842, 930, 1021 pre-existing, 519, 526, 665, 672, 939 prefrontal cortex (PFC), 3, 205, 213, 216, 228, 229, 230, 231, 232, 233, 235, 240, 242, 243, 247, 248, 250, 251, 252, 253, 255, 256, 257, 258, 614, 615, 658, 1116, 1118, 1119, 1128, 1130, 1131, 1132, 1133 pregnancy test, 396, 417, 418, 419, 432, 435 pregnant, xv, 183, 364, 365, 369, 383, 393, 407, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 432, 434, 435, 436, 437, 438, 439, 440, 441, 446, 447, 471, 506, 514, 535, 581, 583, 584, 586, 587, 588, 794, 1130 pregnant women, 383, 412, 423, 424, 425, 440, 794 prejudice, 204, 336 premature death, 289, 374 prematurity, 483 premenopausal, 333 premenopausal women, 333 premium, xxv, 685, 686, 717, 736, 737, 751, 752, 753, 754, 755, 756, 758, 759, 763, 764, 765, 874, 902, 938, 939, 941, 943, 948 premiums, 714, 737, 901, 936, 943, 947 prenatal care, 352 preparation, xxvi, 482, 811, 814, 817, 823 preparedness, 408, 811, 812, 814, 817, 820, 823 prepubertal, 500 preschool, 542, 548 preschool children, 542, 548 preschoolers, 543, 545, 548, 549, 554 prescription drug, 388 prescription drugs, 388 present value, 67 President Clinton, 623 pressure, xix, xxii, 51, 129, 215, 284, 290, 322, 324, 382, 387, 389, 392, 412, 413, 414, 443, 446, 480, 485, 490, 557, 560, 565, 566, 610, 622, 653, 672, 727, 728, 774, 790, 914, 986, 989 prestige, 265, 269, 896 presumption of innocence, 495 presynaptic, 248, 255, 256 preterm infants, 301 prevention, 67, 178, 384, 394, 450, 452, 459, 460,

461, 462, 463, 464, 467, 468, 469, 470, 471, 472,

1173

473, 474, 486, 487, 488, 495, 497, 500, 513, 514, 515, 532, 533, 534, 538, 622, 624, 629, 632, 634, 842, 863 preventive, xiii, xx, 38, 63, 64, 65, 66, 67, 68, 70, 71, 72, 73, 264, 268, 271, 272, 274, 316, 468, 487, 500, 503, 796 preventive programs, 66, 67, 70 price changes, 756, 1090 prices, xxviii, 16, 203, 714, 718, 719, 738, 739, 743, 901, 924, 935, 939, 958, 965, 1085, 1090, 1113 primacy, 107, 184, 572, 1030, 1116, 1132 primary care, 119, 155, 181, 182, 474, 477 primary school, 543, 549, 554 primate, 230, 240, 259, 616, 659 primates, 242, 244, 250, 254, 1132, 1133 priming, 968 principle, 269, 270 prior knowledge, 26, 27, 28, 41, 50, 697, 988 prisoners, 1042, 1043 privacy, 163, 289, 405, 406, 451 private, xix, xxviii, 10, 13, 175, 340, 346, 349, 359, 372, 378, 392, 395, 396, 400, 410, 411, 416, 417, 418, 421, 432, 434, 436, 437, 438, 442, 448, 449, 450, 453, 467, 469, 518, 573, 576, 784, 785, 793, 797, 800, 802, 935, 939, 945, 946, 947 privatization, 590 proactive, 366, 377, 378, 488, 885 probability distribution, xii, 25, 27, 37, 41, 901 probability judgment, 1043 probation, 478 probe, 108, 306, 403, 560 probit models, 817, 820 problem behavior, 514 problem solving, 164, 167, 179, 223, 346, 458, 483, 498, 560, 562, 989, 990, 1028, 1031 problem space, xi, 9, 11, 12 problem-solving skills, 483 problem-solving task, 560 procedures, 482, 492, 826, 838 procrastination, 341 procreation, 372, 582, 622 product attributes, 952 production, 192, 479, 520, 708, 716, 942, 985, 995, 1043 productivity, 160, 715, 716, 717, 1019, 1104 profession, 109, 270, 361, 700, 769, 982 professions, 7, 272, 274, 488, 492, 768, 788, 791 profit, 16, 222, 573, 726, 986, 1005, 1103 profitability, 1102, 1104 profits, 1104, 1106 prognosis, xviii, 5, 8, 37, 38, 40, 297, 305, 307, 359, 479, 481, 774 prognostic factors, 39

prognostic value, 39 program, xix, 60, 64, 68, 151, 196, 304, 312, 322, 335, 389, 392, 399, 451, 452, 457, 459, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 489, 532, 533, 575, 577, 621, 622, 626, 658, 727, 794, 851, 857, 860, 913, 919, 1002, 1088, 1091 programming, 14, 87, 459, 460, 461, 462, 464, 471, 620, 621 progressive, 524 promote, 532, 1065 propagation, 924 property, iv, xxvi, 22, 241, 272, 478, 718, 738, 748, 763, 764, 811, 814, 815, 818, 819, 822, 823, 1087 prophylactic, 316, 333, 334 prophylaxis, 97, 98 proposition, xxx, 175, 547, 580, 716, 1099, 1110, 1111 prosocial behavior, 478 prosperity, xxix, 981, 1101 prostate, xv, 183, 185, 186, 189 prostate cancer, 185, 189 prostate specific antigen, xv, 183 prostitution, 479, 482 protection, 18, 180, 264, 272, 340, 405, 471, 477, 482, 487, 488, 489, 490, 495, 524, 532, 537, 622, 630, 631, 632, 738, 747, 748, 830, 834, 836, 837, 839, 868, 872, 875, 876, 901 protective factors, 69, 471, 506 protective mechanisms, xxvi, 825 protein, 313 Protestants, 305 protocol, 6, 104, 378, 451, 490, 492, 506, 507, 533, 739, 776, 777 protocols, 451, 490, 519 prototype, 119, 733, 963, 965, 966 proxy, xxviii, 299, 309, 935, 1106, 1107 prudence, 143, 903, 906, 908 psychiatric diagnosis, 498 psychiatric disorder, xx, 475, 482, 486, 499 psychiatric disorders, xx, 475, 482, 486, 499 psychiatric hospitals, 85, 89 psychiatric institution, xiii, 75, 78 psychiatric patients, 78, 81, 87 psychiatrists, 264, 272, 492 psychiatry, xx, 475, 526 psychological, xxi, 263, 476, 478, 481, 482, 489, 492, 498, 505, 517, 519, 520, 524, 528, 621 psychological distress, 380, 393, 568 psychological phenomena, 699, 919 psychological processes, 3, 199, 238, 844, 845, 923 psychological well-being, 328, 331

psychologist, 98, 102, 171, 248, 527

psychologists, 492 psychopathology, xxi, 496, 517, 519, 526, 527 psychophysics, xxiv, xxviii, 677, 679, 680, 681, 704, 911, 913, 914, 915, 917, 929, 979, 1078 psychophysiology, 229 psychosis, 77, 78 psychosocial, xx, xxi, 485, 500, 504, 507, 513, 517, 620, 629 psychosocial factors, 117, 118, 119, 179, 398, 402, 403 psychosocial stress, xvii, 279 psychosomatic, 387, 519, 524, 525 psychostimulants, 235 psychotherapeutic, 525 psychotic, 77 psychotropic drug, 873 PsycINFO, 459 puberty, xx, 95, 395, 418, 432, 503, 526, 1124 public, xx, 264, 274, 475, 478, 479, 486, 487, 507, 513, 518, 622, 623, 625, 627, 628, 632, 634, 739, 827, 830, 838 public awareness, 479 public domain, 306 public health, xx, 155, 330, 395, 473, 475, 478, 486, 513, 575, 576, 583, 585, 627 public policy, xxiv, 43, 67, 68, 707 public relations, 449 public safety, 86 public schools, 507, 622, 628, 634 public sector, 450, 886 public service, 487 publishers, 153, 659 Puerto Rican, 300, 591 punishment, 205, 213, 214, 217, 218, 224, 229, 232, 243, 257, 480, 482, 713, 1122, 1129, 1131 punitive, 59 P-value, 104 pyramidal, 39, 256 Pyszczynski, 1067, 1070

# Q

QLQ-C30, 98, 99, 103, 104, 106 quail, 671, 675, 676 qualitative concept, 1006 qualitative research, xix, 22, 23, 120, 147, 155, 182, 320, 334, 382, 384, 386, 387, 388, 389, 391, 403, 561 quality control, 404 quality of life, xiii, xiv, xxix, 93, 94, 95, 98, 99, 101, 102, 103, 104, 105, 106, 108, 109, 110, 111, 134, 135, 152, 271, 280, 281, 286, 288, 291, 298, 299, 350, 351, 360, 364, 365, 463, 471, 484, 582, 697,

- 805, 806, 807, 879, 981
- quality of service, 451

quality-adjusted life years, 66

quantitative concept, 1006

- quantitative research, 22
- quartile, 40
- questioning, 121, 144, 285, 320, 360
- questionnaire, xiv, xx, 98, 99, 100, 102, 106, 108, 116, 120, 137, 157, 159, 171, 172, 180, 356, 375, 386, 464, 504, 507, 508, 533, 534, 562, 567, 627, 635, 636, 685, 687, 689, 690, 697, 836, 846, 849, 851, 857, 979, 993, 998, 999, 1075, 1076, 1077, 1080
- questionnaires, xiv, 94, 98, 100, 157, 159, 170, 563, 625, 626, 687, 998, 1080
- quizzes, 1048, 1050, 1051, 1054, 1056, 1057, 1058, 1062, 1066, 1068, 1069

# R

race, xxi, 73, 119, 149, 160, 192, 193, 247, 248, 249, 250, 309, 505, 531, 634, 848, 864, 942, 945 radio, 245 radio station, 245 radiological, 136 radiologists, 477 radiotherapy, 287, 295 radon, 382 rail, 166, 378 rain, 140, 225, 240, 691, 701, 822, 1088 rainwater, 1018 random, xxx, 32, 33, 35, 36, 37, 203, 312, 466, 536, 672, 779, 812, 817, 819, 820, 883, 939, 1015, 1024, 1025, 1026, 1027, 1037, 1073, 1075, 1076, 1083, 1087, 1088, 1089, 1090, 1125, 1127 rape, 85, 392, 479, 494, 499, 505, 518, 519, 520 rapists, 91 rat, 41, 239, 240, 244, 253, 254, 255, 257, 258, 259, 1031, 1127, 1131, 1132, 1134 rate of return, 753 ratings, 172, 349, 1077 rationalisation, 770 rationality, 17, 18, 73, 200, 220, 343, 452, 573, 574, 590, 605, 609, 612, 617, 648, 649, 652, 656, 660, 738, 891, 893, 894, 895, 898, 908, 913, 919, 925, 926, 929, 930, 931, 932, 1043 rats, xvi, 237, 240, 241, 244, 247, 249, 250, 251, 252, 253, 255, 256, 257, 259, 602, 604, 605, 613, 614, 616, 645, 647, 649, 656, 657, 658, 659, 660, 1017, 1029, 1031, 1117, 1118, 1127, 1131, 1132 reaction time, 204, 224 reactivity, 233, 485

reading, 23, 145, 201, 307, 318, 319, 355, 429, 523, 526, 583, 880, 988, 992, 1037 real estate, xxiv, 707, 714, 718, 719, 943 real time, xv, 53, 54, 58, 60, 184, 926 real wage, 714, 715 realism, 156, 1064 reality, 57, 107, 119, 121, 159, 168, 169, 177, 180, 182, 187, 190, 192, 227, 263, 280, 283, 334, 378, 525, 553, 612, 655, 687, 697, 768, 792, 869, 912, 925, 981 reasoning, xxix, 69, 72, 77, 198, 200, 221, 224, 273, 306, 543, 552, 554, 555, 583, 907, 919, 926, 929, 930, 931, 1033, 1034, 1035, 1036, 1038, 1039, 1040, 1041, 1042, 1043, 1103 recall, 98, 108, 198, 202, 244, 254, 259, 313, 324, 375, 376, 490, 554, 690, 903, 904, 954, 971, 973, 994 recall information, 994 recalling, 490 reception, 214 receptors, 235, 245, 246, 252, 255, 256, 257, 258 recession, 715 recidivism, 77, 81, 83, 90, 91 recidivism rate, 81 reciprocity, xxiv, 213, 222, 223, 560, 616, 660, 707, 713, 714, 719 recognition, xxx, 22, 51, 184, 206, 208, 209, 217, 227, 254, 257, 270, 286, 325, 340, 342, 347, 377, 477, 490, 580, 620, 1099, 1104 recollection, 199 reconcile, 752, 914, 1036, 1127 reconstruction, 1111 recovery, xiii, 94, 95, 136, 486, 585, 790, 1022, 1024, 1030, 1031 recreation, 460, 461, 465, 868, 869, 880, 884, 885 recreational, xx, xxvii, 446, 457, 464, 466, 484, 594, 638, 843, 868, 870, 878, 879, 880, 885, 1087 recruiting, 400, 404, 405 recurrence, 280, 487, 519 recurrent respiratory papillomatosis, 634 red light, 847, 856, 864 reduction, 65, 66, 67, 68, 70, 263, 266, 458, 628, 714, 779, 790, 801, 822, 823, 837, 876, 917, 1030, 1042, 1058, 1085, 1086 refining, 319 reflection, 18, 176, 202, 288, 315, 339, 347, 580, 786, 790, 924 refuge, 347 regenerate, 1014 regional, 10, 12, 115, 308, 355, 942, 1090 Registry, 96 regression, xiii, xx, 75, 82, 83, 84, 85, 86, 87, 89, 90, 504, 508, 512, 740, 744, 745, 746, 764, 781, 782,

784, 812, 826, 830, 834, 852, 898, 1015, 1018, 1027, 1106, 1107 regression analysis, 86, 512, 812, 826, 830, 834, 898, 1107 regression equation, 834, 1106 regression method, 84 regression weights, 852 regressions, xxviii, 86, 744, 781, 805, 834, 935, 1107 regular, xiii, 93, 94, 95, 96, 272, 348, 412, 423, 436, 439, 452, 460, 462, 511, 726, 1129 regulation, 217, 222, 230, 232, 239, 241, 249, 254, 263, 339, 399, 424, 442, 453, 485, 486, 636, 705, 732, 947, 1049, 1127, 1133 regulations, 448, 453, 482, 577 regulators, 450, 884 regulatory bodies, xxxi, 1099 rehabilitation, xx, 167, 266, 270, 475 Reimann, 255 reinforcement, 205, 208, 212, 213, 214, 215, 216, 218, 219, 220, 222, 226, 241, 243, 252, 253, 256, 259, 388, 577, 613, 614, 615, 616, 657, 658, 659, 1027, 1133 reinforcement contingencies, 216, 226 reinforcement learning, 253, 258 reinforcers, 212, 214, 218, 221, 223, 225, 242, 246, 615, 616, 658, 659 rejection, 97, 99, 101, 105, 172, 216, 314, 351, 352, 481, 486, 709, 713, 1041, 1084 relapse, 38, 39 relapses, 38, 39 relationship quality, 470 relative income hypothesis, 717 relatives, xix, 69, 73, 98, 101, 102, 105, 265, 267, 284, 286, 304, 315, 316, 321, 326, 327, 328, 333, 334, 362, 374, 392, 409, 410, 412, 413, 414, 416, 418, 419, 423, 429, 430, 432, 435, 436, 437, 440, 444, 445, 447, 583, 586, 790 relativity, 476, 527, 567, 678, 681, 682, 684, 687, 690, 693, 694, 696, 697, 704, 705, 919, 920, 921, 926, 928, 930, 932 relevance, xiv, xxix, 3, 27, 58, 90, 155, 157, 158, 215, 221, 233, 506, 605, 648, 678, 682, 684, 699, 788, 791, 916, 1003, 1013 reliability, xxix, 196, 369, 402, 473, 494, 500, 562, 566, 567, 584, 626, 628, 630, 674, 849, 850, 855, 996, 1013, 1014, 1015, 1021, 1028 religion, 300, 620, 804, 807, 809 religions, 476 religiosity, 301, 801 religious, 482, 513, 620, 622, 629 religious belief, 52, 60, 303 religious beliefs, 52, 60, 303

religious groups, 429

remission, 38, 106, 110, 111 renal, 466, 474 renal disease, 466, 474 rent, 685, 700, 973, 974, 975 repair, 463 repetitions, xxx, 709, 1045, 1050, 1066, 1088 replacement, xiv, 113, 115, 116, 118, 119, 120, 123, 125, 136, 138, 139, 140, 141, 144, 150, 151, 152, 154, 155, 156, 474, 1036, 1041, 1134 replication, 23, 545, 1120 repression, 362 reproduction, xxii, 316, 317, 322, 338, 347, 348, 350, 353, 356, 359, 360, 361, 362, 364, 365, 366, 369, 372, 375, 377, 378, 379, 385, 387, 571, 572, 573, 578, 579, 582, 584, 589, 591, 592 reproductive age, 393, 572, 573, 574, 576 reptile, 1015 reputation, 49, 50, 56, 58, 270, 395, 560, 587, 629, 709 research, xx, xxi, xxiii, xxv, 261, 265, 272, 274, 475, 476, 483, 496, 497, 500, 504, 506, 507, 513, 514, 517, 518, 519, 520, 521, 524, 526, 527, 528, 532, 533, 536, 537, 538, 619, 620, 621, 623, 624, 626, 627, 630, 632, 735, 738, 826, 827, 833, 836, 1104, 1105 research and development, 43, 539 research design, 76, 382, 621, 1066 researchers, xxiii, xxvi, 267, 273, 274, 519, 619, 621, 623, 624, 626, 627, 628, 825, 827 resentment, 329 reservation, 308, 894 reserves, 15, 594, 599, 600, 639, 643, 644 resettlement, 269, 271 residential, 549 residuals, 852 resilience, 498, 517, 538 resistance, 151, 158, 167, 299, 480, 584, 795, 801, 804, 886, 1000, 1006 resolution, 34, 56, 59, 95, 160, 161, 163, 169, 267, 489, 514, 770, 927 resource allocation, 726 resource management, 10, 11, 12, 54 resources, xi, xx, 6, 9, 10, 13, 15, 32, 34, 36, 41, 42, 59, 68, 70, 71, 72, 169, 175, 176, 180, 218, 226, 268, 283, 306, 372, 392, 394, 395, 452, 460, 461, 463, 464, 465, 466, 471, 474, 475, 476, 478, 484, 487, 488, 574, 580, 594, 599, 605, 638, 643, 649, 686, 773, 792, 793, 834, 838, 839, 925, 938, 988, 990, 1014, 1015, 1017, 1018, 1019, 1021, 1022, 1023, 1027, 1034, 1035, 1041, 1130 respiration, 161, 162 respiratory, 47, 286, 634 respiratory arrest, 286

respiratory therapist, 47 response, 272 response time, 49 responsibility, 263, 267, 482, 487 responsiveness, 80, 82, 481 restaurant, 415, 422, 435 restaurants, 16 restructuring, 340, 362, 371, 378 resuscitation, 196 retail, 1077 retaliation, 59 retardation, 91, 483 retention, 55, 86, 266, 471, 904, 1002, 1023 retirement, xxiv, 134, 162, 326, 677, 678, 682, 683, 684, 686, 687, 688, 689, 690, 691, 698, 699, 702, 704, 717, 918, 929, 936, 937, 949, 960 retirement age, 682, 683, 687, 688, 936, 960 retirement pension, 684 retrenchment, 274 returns, xxx, 245, 600, 644, 678, 685, 699, 717, 752, 754, 755, 756, 757, 758, 759, 760, 762, 763, 764, 765, 1014, 1015, 1016, 1026, 1099, 1101, 1106, 1107 revenue, 60 reversal learning, 252 revolutionary, 268 rewards, xvi, xxiv, 1, 2, 3, 207, 209, 210, 212, 213, 214, 215, 218, 220, 221, 223, 225, 226, 232, 233, 237, 238, 239, 240, 241, 242, 243, 246, 249, 250, 252, 480, 595, 597, 600, 601, 602, 603, 604, 607, 608, 609, 610, 611, 612, 613, 614, 616, 639, 641, 644, 645, 646, 647, 648, 649, 651, 652, 653, 654, 655, 657, 659, 710, 713, 725, 1020, 1025, 1030, 1066, 1102, 1103, 1118, 1122, 1126, 1133, 1134 Reynolds, 600, 611, 616, 644, 655, 660 rheumatoid arthritis, 179 rhythm, 1131 rice, 744, 747, 832, 942 right hemisphere, 210, 217, 219, 226 rights, 476, 488, 495 rigidity, 313, 560 rings, 213, 223 risk assessment, 78, 80, 81, 90, 91, 881, 885 risk aversion, xiii, xxiv, xxv, 16, 63, 65, 66, 70, 203, 214, 227, 595, 597, 611, 639, 641, 654, 677, 686, 689, 696, 704, 717, 737, 746, 751, 752, 753, 754, 755, 758, 759, 762, 763, 764, 830, 834, 871, 884 risk behaviors, xxi, 499, 504, 505, 512, 513, 514, 531, 532, 533, 534, 535, 536, 537, 624, 627, 860 risk factors, 26, 30, 32, 41, 69, 72, 78, 482, 483, 486, 487, 506, 634, 811, 812, 863 risk management, xxvii, 12, 56, 312, 316, 867, 869, 870, 873, 874, 875, 876, 877, 878, 884, 885

risk perception, xxvi, 66, 67, 69, 72, 73, 339, 379, 695, 698, 706, 799, 806, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 825, 826, 827, 828, 833, 834, 843, 859, 861, 862, 886, 910 risk profile, 696, 772, 876 risk society, xxvi, 327, 799 risk-taking, 66, 230, 233, 256, 270, 273, 480, 628, 846, 860, 861, 863, 998, 999, 1000, 1010, 1100, 1106, 1107, 1111, 1120, 1122, 1131, 1133 risky driving, xxvi, 841, 842, 843, 844, 845, 853, 855, 856, 858, 859, 861, 862, 864 Rita, viii, 841 RMSEA, 851, 852, 857, 858 RNA, 96, 97, 255 robbery, 85, 1023 rodent, 1127, 1131 rodents, xxxi, 1016, 1023, 1115 rods, 620 ROE, 1106, 1108, 1109 romantic relationship, 486 Royal Society, 674, 675, 1096, 1097, 1134 rumination, 89 rural, 87, 474, 581 Russia, 309 Russian, 305, 390, 872 Rwanda, 371

# S

sacred, 292 sacrifice, 458, 709, 713, 727, 731 sadness, 202, 204 safe room, 826, 829, 832, 833, 834, 835, 836, 837, 838 safeguard, 24, 320, 493 safety, xxx, 46, 47, 49, 55, 56, 57, 60, 86, 110, 144, 266, 287, 388, 428, 484, 489, 735, 800, 826, 829, 831, 860, 862, 863, 870, 874, 876, 877, 879, 880, 881, 885, 886, 905, 1020, 1073, 1074, 1075, 1076, 1077, 1078, 1079, 1083, 1084, 1085, 1086, 1089, 1090, 1091, 1092, 1093, 1095 salaries, 714, 971, 972 salary, 265, 411, 415, 434, 589, 682, 685, 689, 700, 972 sales, 678, 697, 1106 saliva, 492 sample design, 507 sampling, xiii, xxviii, 27, 33, 42, 75, 76, 77, 79, 80, 81, 85, 89, 105, 108, 120, 354, 507, 561, 679, 705, 833, 903, 904, 907, 909, 911, 913, 923, 924, 925, 932, 979, 980, 1021, 1027

sand, 912

Sarin, 1113 satellite, xxi, 504, 876, 877, 881, 995, 996 satisfaction, xxv, 46, 52, 55, 95, 98, 99, 100, 105, 139, 167, 186, 222, 226, 332, 558, 566, 594, 638, 681, 767, 774, 775, 776, 777, 789, 795, 797, 869, 878, 879, 880, 927, 943, 1064 saturation, 405 Saturday, 423 saving rate, xxiv, 677, 690, 697, 699, 920 savings, xxiv, 415, 677, 678, 683, 685, 687, 688,

689, 690, 691, 692, 694, 695, 696, 697, 699, 700, 701, 702, 704, 705, 717, 900, 901, 929, 949,

1075, 1076, 1090

savings account, 900

savings rate, 683, 685, 690, 692, 701

SBU, 1112

scalar, 605, 648, 1030

scaling, 11, 705, 884, 917, 930, 931

scandalous, 923

scarcity, 744

scheduling, 577

Scheffe test, 999

schema, 163, 344

schemas, 360

schizophrenia, 77, 78, 79, 247, 254, 256, 268, 276

Schmid, 515

scholarship, 22

school failure, 522

school learning, 525

school performance, 508, 509

schooling, 581, 582, 815

sciatica, 133

science, xxvi, 492, 538, 614, 708, 711, 799, 800,

803, 804, 805, 806, 807, 809, 912, 929, 931, 932 scientific, 479, 492, 495, 622, 629

scientific knowledge, 175, 176, 806, 983

scientific validity, 69

scientists, 622

sclerosis, 34, 40, 42

scores, 99, 101, 102, 103, 104, 120, 172, 468, 561, 562, 694, 695, 776, 778, 779, 780, 781, 782, 784, 830, 846, 849, 851, 853, 858, 860, 964, 965, 992, 993, 1004, 1050, 1051, 1052, 1053, 1057, 1058, 1062, 1066, 1121, 1122

screening programs, 64, 70

scripts, 490, 903

SCUBA, 868

search, xiii, 6, 15, 78, 93, 95, 96, 117, 164, 165, 177, 199, 210, 216, 341, 344, 357, 386, 389, 425, 459, 490, 605, 632, 633, 649, 679, 827, 863, 904, 924, 968, 1021, 1023, 1070, 1132 searches, 394, 770

sanctions, 263, 630

searching, 165, 430, 994 Seattle, 97, 110, 180, 886 second language, 58, 186 secondary, 267, 271 secret, 332, 414, 416, 419, 421, 437, 439 secretion, 213 secrets, 273 security, xxviii, 77, 85, 87, 88, 89, 266, 575, 690, 877, 878, 879, 881, 935 security services, 266 sedentary, 472, 1130 seed, 666, 667, 1022 seeds, 1014, 1015, 1016, 1020, 1022, 1023 segmentation, 450 segregation, 895 selecting, 76, 86, 151, 172, 200, 215, 402, 664, 682, 684, 688, 689, 831, 846, 918, 957, 1027, 1040, 1041, 1080 selective attention, 211 selectivity, 1068 Self, viii, ix, xvii, xxi, xxiv, xxx, 17, 19, 113, 123, 124, 143, 149, 161, 172, 176, 180, 263, 264, 265, 267, 268, 270, 272, 275, 276, 297, 298, 388, 418, 426, 474, 479, 481, 486, 488, 496, 499, 522, 532, 533, 535, 536, 558, 562, 566, 567, 650, 722, 725, 726, 730, 731, 733, 839, 1045, 1046, 1047, 1066, 1069, 1070, 1071, 1131 self worth, 488 self-actualization, 868 self-assessment, 1064, 1067 self-awareness, 166, 180 self-care, 458 self-concept, xxv, 165, 725, 726, 728, 730, 731 self-confidence, 167, 519, 1004 self-consistency, 979 self-control, 16, 220, 460, 606, 608, 609, 614, 615, 649, 650, 651, 652, 658, 659, 716, 717, 843, 1118, 1119, 1131 self-destructive behavior, 479, 499 self-discipline, 620 self-efficacy, 163, 164, 179, 339, 466, 469, 470, 558, 559, 560, 561, 563, 564, 565, 566, 567, 568, 569 self-employed, 716 self-esteem, xxiv, 17, 19, 348, 496, 519, 725, 726, 727, 728, 729, 730, 731, 732, 733, 863, 1047, 1068 self-evaluations, 1004, 1047, 1048, 1049, 1053, 1057, 1064, 1065, 1067, 1068, 1069, 1070 self-expression, 561 self-help, 453, 467, 472 self-identity, 335, 348, 394, 732 self-image, xv, xxiv, 158, 159, 162, 163, 171, 172, 725, 726, 727, 730, 731, 1010

self-improvement, 465 self-interest, 328, 713 self-knowledge, 166 self-management, xix, 148, 149, 180, 392, 453 self-monitoring, 1066 self-mutilation, 528 self-organized criticality, 929 Self-protection, 839 self-regulation, 732, 1049 self-relevant feedback, 1071 self-report, xxi, 468, 532, 533, 535, 536, 627, 686, 694, 863, 939, 941, 943, 1119 self-verification, 1067 self-worth, 481, 726 seller, 425, 938 SEM, 1117, 1119, 1121, 1122, 1123, 1124, 1125, 1126, 1129 semantic, 230, 527, 904 semen, 584 semi-structured interviews, xiv, xviii, 113, 121, 142, 311 Senate, 771, 773, 792, 797 sensation, xxvi, 159, 225, 459, 704, 841, 842, 845, 851, 853, 854, 855, 856, 857, 858, 859, 860, 861, 863, 864, 915, 930 sensation seeking, xxvi, 459, 841, 842, 845, 851, 853, 854, 855, 856, 857, 858, 859, 860, 861, 863, 864 sensations, 209, 843 sense of coherence, 263 sensitivity, 11, 22, 57, 85, 221, 248, 489, 505, 605, 611, 613, 614, 615, 616, 648, 654, 655, 657, 658, 660, 740, 758, 808, 900, 927, 955, 960, 962, 963, 966, 1017, 1031, 1106, 1126, 1129, 1131 sensitization, 161, 254 sensory cortices, 242 sensory data, 161 sensory modalities, 208, 210 sensory modality, 213 sensory systems, 208 sentences, 223, 392, 952 separation, 481, 566, 573, 1110 septum, 212 sequelae, 478 sequencing, 1080 Serbia, 503 series, xii, xxv, 25, 27, 41, 46, 47, 50, 55, 57, 85, 170, 200, 206, 207, 248, 263, 273, 401, 490, 573, 574, 578, 580, 582, 583, 608, 623, 628, 652, 666, 689, 701, 751, 752, 755, 756, 757, 759, 762, 763, 765, 781, 792, 874, 877, 991, 1020, 1023, 1129 serology, 96, 97

serotonergic, xxxi, 485, 617, 661, 1115, 1118, 1120, 1127 serotonergic dysfunction, 485 serotonin, 251, 258, 1118, 1127, 1130, 1132, 1133 serum, 485 service provider, xiv, xv, 157, 158, 159, 167, 168, 169, 171, 172, 173, 175, 176, 177, 399, 401, 442, 451, 495 SES, 559 settlements, xxix, 981 severity, xiv, 67, 80, 81, 83, 113, 118, 119, 152, 179, 250, 304, 313, 330, 379, 477, 478, 485, 519, 826, 828, 829, 831, 833, 834, 836, 837, 838 sex differences, 1133 sex offenders, 79, 90 sex ratio, 1027, 1031 sex role, 485 sex steroid, 1132 sexual abuse, xx, 475, 476, 479, 480, 481, 488, 489, 490, 492, 494, 495, 496, 497, 498, 499, 500, 505, 520, 526, 529 sexual activities, 479, 625, 631, 636 sexual activity, xx, 213, 393, 480, 489, 503, 504, 507, 512, 513, 514, 518, 537, 620, 621, 623, 625, 628, 629, 633 sexual assault, 492, 494, 500, 622 sexual behavior, xiii, xx, xxi, xxiii, 75, 395, 452, 503, 505, 506, 508, 511, 512, 513, 515, 531, 619, 620, 622, 623, 624, 626, 627, 628, 629, 633 sexual contact, 492, 628 sexual experiences, 589 sexual harassment, 46 sexual health, xxi, 395, 396, 406, 412, 499, 531, 532, 533, 534, 535, 536, 537, 627, 629, 631 sexual intercourse, xx, 393, 411, 415, 424, 451, 504, 507, 511, 512, 513, 514, 523, 579, 620, 621, 623, 625, 626, 627, 629, 632 sexual recidivism, 91 sexual risk behavior, xx, xxi, 394, 503, 504 sexual transmitted diseases, xx, 503, 514 sexual violence, 81, 519 sexuality, xx, xxiii, 394, 399, 401, 404, 405, 407, 452, 503, 505, 506, 513, 517, 537, 579, 581, 582, 586, 589, 619, 620, 621, 622, 623, 624, 627, 628, 629, 632, 634, 636 sexually abused, 481, 490, 497, 498, 521, 524, 525, 526, 528 sexually transmitted disease, 494, 505, 506, 511, 535 sexually transmitted diseases, 505, 511, 535 Sexually Transmitted Infections (STIs), xxiii, 467, 470, 514, 534, 619, 620 shade, 1021 shame, 202, 285, 519, 587

shape, 65, 81, 106, 149, 162, 191, 234, 339, 344, 352, 366, 428, 437, 442, 448, 478, 525, 572, 574, 597, 632, 641, 737, 740, 747, 891, 899, 901, 908, 927, 959, 960, 1005 shaping, 17, 342, 364, 366, 376, 452, 610, 653, 893, 989, 1046, 1056 shareholders, 1103, 1110, 1111 shares, 201, 454, 689 sharing, 168, 222, 280, 284, 302, 345, 446, 453, 773 shelter, 403, 409, 410, 415, 419, 422, 424, 426, 429, 430, 432, 434, 436, 439, 445, 446, 448, 449, 451, 453, 477, 826, 830, 832, 833, 834, 835 shock, 414, 418, 432, 434, 439 shocks, 233, 246, 960 short period, 440 short run, 752, 763, 1102, 1104, 1106 shortage, 46, 55, 265, 962 short-term, 13, 16, 18, 41, 78, 89, 205, 230, 392, 464, 602, 604, 605, 607, 608, 609, 646, 648, 649, 651, 652, 653, 717, 752, 756, 757, 762, 763, 1103, 1104, 1130, 1133 short-term interest rate, 756 short-term memory, 18 shy, 51, 58, 350, 442 sibling, xiii, 93, 95, 104, 108, 321, 366, 372, 376, 388 sibling support, 372 siblings, 104, 321, 354, 362, 366, 373, 375, 376, 513, 586 sickle cell, 109, 384 side effects, 31, 186, 395, 411, 414, 436, 584, 585, 800 sign, xv, 99, 158, 159, 164, 405, 417, 542, 583, 587, 739, 746, 936, 1085, 1093, 1110 signaling, 216, 258 signals, 204, 205, 206, 208, 210, 212, 216, 223, 225, 240, 241, 247, 249, 250, 252, 290, 614, 658, 674, 800, 802, 809 significance level, 30 signs, xxi, 199, 413, 416, 417, 432, 476, 517, 521, 582, 847, 864, 1080, 1084, 1085, 1093, 1094, 1110 similarity, 201, 247, 343, 402, 546, 572, 609, 653, 834, 845, 846, 902, 906, 909, 917, 930, 978, 990, 1017, 1018, 1066 simulation, 2, 38, 341, 342, 373, 374, 554, 924, 952, 962, 963, 967, 975, 976, 977, 993, 997 simulations, 341, 965, 968, 1021, 1030 sine, 669 single test, 1017 singular, 625 sites, 156, 247, 395, 403, 775, 776, 788, 791, 1116, 1124

situation, 265 skeptics, 710, 805 skilled labor, 463 skills, xvi, xvii, xxvi, 49, 53, 56, 57, 58, 90, 153, 166, 175, 194, 269, 274, 279, 283, 287, 360, 402, 404, 452, 458, 459, 460, 474, 483, 485, 486, 487, 488, 489, 632, 727, 773, 792, 841, 843, 847, 848, 849, 850, 851, 853, 854, 855, 856, 859, 860, 862, 865, 982, 989, 1005, 1124 skills training, 153 skin, 137, 205, 206, 229, 349, 524, 1116, 1131 skin conductance, 205, 206, 229, 1116, 1131 skin conductance responses, 205, 206, 1116 slavery, 302 slaves, xxiii, 663 sleep, 99, 175, 284, 523, 527, 881, 883, 973 sleep deprivation, 284 sleep disturbance, 99 Sm, 656 small firms, 1104, 1112 smoke, 507, 829, 900 smokers, 66, 472, 507, 510, 939 smoking, 66, 70, 71, 72, 323, 327, 329, 330, 452, 459, 460, 465, 467, 471, 472, 507, 510, 511, 512, 514, 616, 660, 687 smoking cessation, 72, 329, 330, 460, 471, 472 smoothing, 32, 36, 756 smoothness, 758 SNAP, 538 SNc, 239, 245 SNS, 1111 sociability, 519 social, xvi, 261, 262, 263, 264, 265, 268, 269, 270, 271, 272, 274, 476, 477, 478, 479, 481, 483, 484, 485, 486, 487, 488, 489, 492, 497, 498, 500, 505, 519, 523, 525, 527, 528, 621, 630, 632, 636 social activities, 115, 356, 481 social attitudes, 321 social behavior, 18, 609, 653, 1064 social behaviour, 205, 216, 217, 614, 658 social benefits, 160 social care, 476 social category, 1064 social change, 575 social class, 476 social cognition, 230, 314, 315, 317, 330, 332, 904, 1069 social cognitive model, 568 social cohesion, 263 social comparison, xxix, xxx, 16, 951, 952, 956, 959, 972, 973, 976, 977, 978, 980, 1045, 1046, 1047, 1048, 1049, 1050, 1052, 1053, 1058, 1059, 1060, 1061, 1064, 1065, 1066, 1067, 1069, 1070, 1071

social competence, 485, 486 social consequences, 162 social construct, 146, 149, 162, 163, 178, 636 social context, xxii, 162, 163, 164, 349, 375, 377, 572, 573, 574, 586, 589, 613, 657, 1030, 1046, 1049, 1064 social control, 262, 269 social desirability, 468 social development, 505 social environment, 15, 168, 486, 770 social evaluation, 378 social events, 166 social exclusion, 575 social factors, xxi, 517, 519 social group, 160, 223, 303, 573 social identity, 265 social influence, 162, 314, 315, 330, 340, 473, 577, 676 social influences, 162, 314, 315, 340, 676 social institutions, 460 social integration, 461 social isolation, 132, 167, 458, 483, 484, 487, 489 social learning, 389 social learning theory, 389 social life, 164, 685 social network, xxii, 83, 119, 128, 168, 271, 281, 340, 402, 471, 561, 572, 581 social networks, 271 social norms, 198, 419, 429, 573, 788, 859 social performance, 1046, 1047, 1049, 1050, 1052, 1053, 1054, 1055, 1056, 1057, 1058, 1061, 1062, 1063, 1064, 1065, 1067, 1069 social phenomena, 913 social policy, 500 social problems, 95 social psychology, xii, 10, 15, 16, 17, 18, 314, 319, 385, 632, 732, 733, 980, 1070, 1071 social regulation, 486 social relations, 173, 481, 485, 486, 574, 1001 social relationships, 173, 481, 485, 486 social roles, 565, 990 social sciences, xxii, 382, 571, 913, 1006 social security, 575, 690 Social Security, 621, 636 social services, 271, 272, 273, 275, 398, 401, 460, 466, 478, 497, 498 social situations, 204 social skills, 49, 459, 632, 990 social status, 422, 861 social stress, 489 social structure, 142, 572

social support, 339, 358, 359, 364, 372, 375, 461, 470, 472, 486, 560, 573, 576, 583, 773

social welfare, 505 social withdrawal, 134 social work, xvi, 261, 264, 265, 268, 269, 270, 271, 272, 274, 317, 400, 403, 406, 415, 429, 439, 449, 460, 492 social workers, 261, 264, 265, 269, 270, 271, 272, 274, 400, 403, 449, 492 socialisation, 169, 266 socially, 262, 267, 269, 270, 479, 486, 488, 505, 514 society, xxvii, xxix, 5, 52, 64, 70, 71, 96, 107, 119, 141, 162, 163, 166, 185, 262, 263, 269, 270, 271, 280, 298, 306, 307, 335, 336, 383, 397, 407, 409, 423, 427, 448, 476, 484, 505, 559, 574, 620, 621, 633, 636, 659, 678, 800, 801, 804, 831, 868, 885, 924, 979, 981, 982, 989, 990, 999, 1005, 1007 sociocultural, 165, 505 socioeconomic, 180, 459, 478, 482, 576, 578, 589, 777, 801, 829 socioeconomic status, 459, 478, 482 sociological, 119, 262, 518, 870, 981, 982, 1007, 1008 sociologists, 263 sociology, 71, 151, 158, 319, 355, 382, 532, 997, 1006 Socrates, 275, 987 sodium, 258 sodomy, 479 soft drinks, 406 software, 35, 78, 357, 690, 817, 995, 1009 solar, 805 solar energy, 805 solidarity, 264, 560, 568 Solow, 715, 722 solution, 273 somatic marker, xvi, 197, 205, 206, 209, 210, 216, 220, 228, 232, 234, 1134 somatic symptoms, 524 sorting, 917 sounds, 914, 917 South America, 300 South Asia, 591 Southeast Asia, 455 space exploration, 805 Spain, viii, xxvi, 1, 197, 264, 799, 801, 805 spatial, 2, 13, 32, 33, 36, 47, 162, 178, 245, 246, 247, 542, 554, 609, 610, 653, 660, 912, 1022, 1025, 1030, 1031 spatial information, 2 spatial location, 1022 spatial memory, 1030, 1031 spatial representations, 1022 spatiotemporal, 232 spawning, 869

special education, 477 specialization, 217 species, xxix, xxxi, 241, 609, 610, 652, 653, 659, 664, 665, 671, 672, 673, 925, 1013, 1015, 1017, 1018, 1022, 1028, 1115 specificity, xiii, xxi, 40, 75, 77, 185, 207, 517, 526, 568, 627, 894, 1043, 1085, 1116 spectra, 752, 755, 756, 759, 762 spectral analysis, 752, 755, 764 spectrum, 8, 666, 755, 762, 763, 764, 991 speculation, 67 speech, 313 speed, 27, 115, 198, 360, 542, 842, 843, 845, 847, 848, 849, 855, 861, 864, 865, 879, 881, 882, 1021, 1122, 1123, 1124 speed limit, 849, 864, 865 sperm, 380, 492, 632 spheres, 204 sphincter, 39, 523 spin, 912 spinal cord, 161 spiritual, 7, 166, 286, 291, 302, 307, 308, 309, 990 spirituality, 8, 301, 302 sponsor, 947 spontaneity, 494 spontaneous recovery, 1024, 1031 sporadic, 511, 587 sports, 711, 712, 868, 882, 1034, 1050, 1064 spouse, 285, 312, 323, 329, 384, 483, 942, 946 spreadsheets, 688 springs, 233 SPSS, 630, 669, 851, 857, 864 SRH, 448, 450, 451 Sri Lanka, 591 S-shaped, 599, 897, 899, 900 SSI, 1096 SSS, 851 stability, 38, 459, 460, 539, 620, 672, 744, 764, 883, 893, 900, 1022 stabilize, 924, 928 staff, 266, 272 staff development, 999 stages, 29, 89, 168, 171, 180, 187, 188, 190, 193, 208, 340, 341, 344, 354, 361, 362, 370, 372, 374, 375, 451, 452, 482, 558, 857, 990, 991, 996, 1007, 1070, 1118, 1119 stakeholder, 276 stakeholders, 345, 800 standard deviation, 138, 605, 648, 688, 757, 1084, 1085, 1088, 1089, 1106, 1107, 1108, 1109 standard error, 103, 104, 692, 693, 1085, 1107

standards, xii, 22, 45, 54, 55, 56, 58, 59, 60, 191, 334, 344, 405, 478, 480, 490, 495, 580, 589, 794, 916, 926, 982, 1034, 1041, 1070 starvation, 600, 605, 609, 643, 649, 653 statistical analysis, 27, 37, 621, 636 Statistical Package for the Social Sciences, 864 statistics, xii, xx, 25, 26, 28, 32, 40, 43, 99, 115, 152, 264, 314, 352, 463, 477, 487, 497, 504, 508, 596, 640, 757, 781, 803, 820, 833, 842, 860, 871, 872, 873, 876, 881, 884, 909, 930, 932, 939, 1050, 1078, 1105, 1106 statutory, 265 STD, 472, 506, 629, 633 steel, 876 stem cell research, 620 stem cell transplantation, 110, 111 stem cells, 110 stereotype, 1040 stereotypes, 135, 145, 168, 956, 1040 stereotypical, 670, 1040 sterilization, xxii, 350, 369, 571, 572, 573, 575, 576, 577, 578, 579, 581, 582, 583, 585, 586, 587, 588, 589, 590, 592 steroid, 137, 1132 STI, 627, 629 stigma, 166, 269, 295, 318, 335, 444, 624 stigmatization, 399, 401, 410, 444, 448 stigmatized, 333, 385, 393, 430 stimulant, 228 Stimuli, 987 stimulus, 206, 207, 208, 213, 217, 218, 220, 242, 243, 246, 247, 253, 256, 573, 612, 655, 667, 668, 914, 915, 916, 917, 922, 930, 978, 987, 988, 1074, 1078, 1079 STIs, xxiii, 619, 620, 622, 624, 628, 629, 632 stochastic, 37, 230, 764, 765, 921, 922, 924 stock, xxx, 2, 389, 666, 678, 689, 698, 702, 717, 719, 727, 752, 754, 755, 757, 759, 763, 764, 918, 1099, 1101, 1105, 1106, 1107, 1110, 1111 stock markets, 752, 757, 763 stock price, 719, 1110 stomach, 522, 524 storage, 823, 904, 926, 995 storm resistant construction, 735 storm shutters, 815, 816, 819 Storm surge, 815, 816, 818, 819 storms, xxv, xxvi, 90, 735, 811, 815, 823, 871 strategic, 1101, 1102, 1103, 1106, 1111 strategic management, 1111 strategies, xiv, xx, xxvi, xxvii, xxix, 21, 33, 50, 53, 54, 76, 80, 91, 94, 105, 108, 109, 148, 180, 219, 235, 284, 287, 328, 341, 342, 348, 359, 362, 365, 370, 373, 374, 378, 396, 459, 475, 485, 487, 498,

527, 546, 582, 613, 657, 702, 704, 733, 768, 769, 770, 774, 775, 788, 789, 792, 793, 800, 841, 860, 862, 929, 951, 968, 988, 1003, 1015, 1030, 1032, 1102, 1103, 1112, 1116, 1118 strategy use, 220 streams, 59 strength, xxix, 39, 80, 121, 143, 147, 149, 161, 170, 301, 671, 672, 746, 775, 783, 787, 791, 833, 854, 861, 981, 1006, 1026 stress, xx, 47, 51, 53, 55, 57, 60, 105, 163, 169, 263, 284, 330, 348, 373, 376, 389, 429, 434, 440, 445, 449, 460, 469, 475, 482, 483, 485, 487, 493, 499, 544, 688, 770, 787, 805, 863, 870, 926, 1041, 1100 stress level, 263, 284 stress reactions, 499 stressors, xvii, 279, 482 stress-related, 330 striatum, xvi, 207, 210, 212, 213, 214, 220, 222, 225, 229, 232, 233, 237, 239, 242, 243, 245, 246, 247, 248, 250, 251, 252, 253, 254, 255, 256, 257, 1118, 1119, 1120, 1131 strictures, 679 strikes, 709, 729 stroke, 31, 42, 185, 191 strokes, 161 structural equation model, 861, 863, 1091 Structural Equation Modeling, 863 structuring, xii, 9, 11, 195, 376, 574, 992 student group, 460, 737 study, 265, 266, 267, 273 subcortical structures, 241 subdural hematoma, 496 subgroups, 30 Subjective Expected Utility theory, 892 subjective experience, 206, 633, 904 subjective judgments, 629 subjective well-being, 980 subjectivity, 23, 28, 168, 624 subscribers, 937 subsidization, 578 subsidy, 13, 736 substance abuse, xvi, xx, 53, 197, 207, 237, 247, 249, 452, 458, 460, 471, 472, 473, 475, 482, 483, 485, 486, 487, 489, 506, 515, 532, 863 substance addiction, 1131 substance use, 458, 459, 468, 469, 471, 511, 538 substances, 161, 198, 212, 221, 224, 225, 462, 492, 507, 831 substantia nigra, 212, 239, 249, 253, 257, 258 substitutes, 945 substitution, 714, 753, 1082, 1083, 1085, 1087 substrates, 4, 18, 230

subtraction, 1027 suburban, 464, 942, 1037 success rate, 105, 138, 187 suffering, 28, 163, 166, 187, 190, 191, 194, 264, 268, 288, 290, 291, 292, 298, 302, 303, 307, 414, 458, 496, 525, 1116 sugar, 882 suicidal, xx, 475, 486, 499, 519, 524 suicidal behavior, xx, 475, 486, 499 suicidal ideation, 524 suicide, xvii, 79, 264, 279, 280, 290, 291, 292, 301, 415, 479, 486, 499 suicide attempters, 486 suicide attempts, 486, 499 sulphur, 139, 140 summaries, 1002 summer, 812, 881, 958 Sun, 259, 464, 473, 474 Sunday, 423 sunflower, 1016 superior parietal cortex, 226 superiority, 76, 107, 170, 217, 992, 1065 superposition, 756 superstitious, 560, 565 supervision, 266, 395, 477, 480 supervisor, 107 supervisors, 59, 381 supplemental, 947, 1103 supplements, 185 suppliers, 938 supply, xxiv, 109, 115, 161, 394, 577, 582, 707, 716, 800, 823, 907, 988, 989, 990, 1002, 1003 supply chain, 800 support services, 60, 489 support staff, 51 suppression, 231, 393, 565, 1127 surfing, 868 surgeons, 49, 62, 127, 128, 138, 154, 262 surgery, xv, 47, 59, 61, 115, 116, 117, 118, 119, 120, 121, 123, 130, 135, 136, 138, 139, 140, 141, 145, 146, 147, 151, 152, 153, 154, 155, 162, 183, 185, 186, 189, 244, 293, 316, 333, 482, 576, 587, 588, 787 surgical, xxii, 47, 116, 139, 302, 571, 572, 573, 575, 576, 577, 578, 586, 772, 794 surgical intervention, 573 surplus, 962 surprise, 126, 187, 222, 242, 565 surrogates, 185, 293 surveillance, 312, 379, 380, 470, 624, 627, 634, 1102 Survey, 277, 824, 825, 829, 830, 838, 839, 840 survival, xvi, xvii, xxvii, 94, 95, 99, 100, 104, 106, 108, 109, 164, 186, 208, 279, 280, 281, 287, 306,

476, 578, 582, 599, 600, 610, 643, 644, 654, 867, 869, 877, 879, 881, 882 surviving, xiv, 94, 98, 100, 873 survivors, 52, 106, 472 susceptibility, 34, 340, 485, 843 suspects, 583 swallowing, 185, 313 Sweden, 63, 265, 269, 309, 388, 504, 1073, 1075, 1097 swelling, 126 switching, xxix, 550, 1033, 1034, 1035, 1036, 1039, 1040, 1041 Switzerland, xi, 5, 389 syllogisms, 1036, 1038, 1039 symbolic, 2, 210, 212, 213, 215, 221, 241, 319, 524, 978 symbols, 374, 397 symmetry, 665, 669, 672, 675 sympathetic, 205, 421, 485, 585 sympathetic nervous system, 485 symptom, xvii, 7, 79, 99, 101, 103, 104, 115, 116, 117, 118, 119, 120, 124, 127, 144, 146, 148, 279, 280, 281, 282, 481, 520 symptoms, xiv, 77, 78, 101, 102, 106, 113, 114, 117, 118, 119, 120, 123, 124, 125, 126, 135, 137, 139, 147, 148, 149, 156, 180, 250, 281, 286, 291, 313, 330, 351, 379, 380, 413, 416, 417, 432, 476, 480, 481, 484, 485, 489, 490, 491, 497, 519, 524, 527, 979 synapse, 240, 245, 247 synapses, 257 synaptic strength, 1026 synchronization, 259 syndrome, xv, 38, 94, 110, 167, 169, 179, 180, 183, 252, 254, 380, 480, 482, 483, 490, 497, 499, 518, 526, 748 syntactic, 527, 926 synthesis, 1, 28, 94, 500, 538, 843, 862, 989 syphilis, 302, 308 systematic, 490, 1106

## Т

tactics, 200, 770 talent, 539, 989 tangible, 221, 1000 target behavior, 461 target population, 440, 463, 464, 467 targets, 180, 241, 550, 716 tariff, 266 task demands, 927, 1048 task force, 56, 139, 154 task load, 1035

task performance, xxx, 1045, 1047, 1048, 1050, 1051, 1054, 1058, 1060, 1065, 1070, 1071, 1119, 1126, 1129 taste, 2, 110, 209, 214, 215, 218, 232, 242, 952, 1088 tax deduction, 1089 taxonomy, 339, 346, 352, 994 teacher attitudes, 1010 teachers, 395, 399, 407, 452, 477, 482, 508, 509, 514, 622, 981, 985, 999, 1000, 1001, 1002, 1003, 1004, 1005, 1007, 1009, 1010 teaching, 120, 372, 452, 490, 621, 628, 632, 635, 981, 982, 983, 984, 985, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000, 1001, 1003, 1005, 1006, 1007, 1008, 1009, 1051 teaching process, 985, 997, 998, 1000, 1003, 1006 teaching strategies, 1003 team members, 287, 495, 630 technological, 262, 828, 1106, 1113 technological change, 369, 1106 technological developments, 369, 372, 801 technology, xxvi, 43, 52, 146, 185, 186, 280, 298, 301, 302, 308, 333, 357, 366, 369, 385, 389, 574, 660, 715, 799, 800, 801, 804, 805, 806, 807, 809, 831, 912, 984, 994, 995, 996, 999, 1000 Technology Assessment, 74, 151, 381, 389 teenagers, 525, 527, 528, 620, 960 teens, 459, 460, 467, 469, 471, 506, 635, 636 teeth, 445 telecommunications, 505, 575 telephone, xix, 318, 449, 457, 460, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 474, 691, 701, 812 television, 145, 186, 187, 196, 429, 575, 979, 995, 1010 temperature, 613, 657, 896, 986 temporal distribution, 67 temporal lobe, 208, 212, 219, 226, 228, 243, 244 tenants, 952 tension, 70, 73, 266, 267, 283, 288, 419 terminal illness, xvii, xviii, 279, 297, 298, 301, 306 terminally ill, 6, 7, 8, 294, 301 termination rule, 31 territory, 377, 518 test scores, 784 test statistic, 1084 testimony, 298, 492, 500 testosterone, 1132 Texas, 388, 531, 533, 590, 735, 739, 824, 826, 908, 1112 textile, 423 Thai, 393, 396, 401, 403, 404, 407, 419, 426, 430, 454, 455

405, 407, 415, 425, 426, 450, 454, 455 thalamus, 208, 214, 225, 226, 242, 253, 257 thalassemia, xiii, 93, 94, 95, 96, 99, 100, 101, 102, 103, 104, 105, 107, 108, 109, 110, 111 theoretical, xxvi, 268, 519, 752, 753, 825, 826, 827, 829, 834, 838, 1112 theoretical assumptions, 568, 600, 644 Theories, 839 Theory of Planned Behavior, 314 therapeutic, 265, 482, 518, 526 therapeutic relationship, 6, 164 therapists, 47, 181 therapy, xiii, 7, 41, 78, 93, 94, 95, 105, 106, 107, 137, 154, 156, 160, 301, 345, 355, 382, 387, 474, 489, 493, 796, 936 thinking, 11, 18, 127, 133, 169, 170, 172, 177, 185, 195, 204, 219, 223, 230, 306, 315, 348, 355, 382, 408, 409, 412, 418, 428, 437, 445, 458, 542, 566, 635, 805, 861, 869, 876, 882, 884, 885, 986, 987, 988, 989, 990, 995, 996, 1005, 1007, 1008, 1024, 1034, 1035, 1036, 1042, 1043 thinking styles, 566 third party, xxv, 735 Thomson, 66, 73, 114, 155, 472 threat, 167, 179, 182, 200, 202, 208, 230, 263, 266, 325, 329, 330, 340, 367, 386, 477, 505, 527, 576, 727, 728, 729, 730, 731, 830, 1104, 1110 threatened, xvii, xxv, 164, 279, 281, 329, 527, 712, 730, 731, 735 threatening, 94, 204, 207, 208, 349, 351, 356, 363, 477, 481, 580, 727, 862, 1000 threats, 162, 348, 349, 351, 385, 480, 505, 580, 727, 732, 1110 three-dimensional, xi, 9, 12, 998 three-way interaction, 1056, 1123 threshold, 84, 106, 120, 146, 161, 207, 485, 600, 604, 609, 643, 647, 653, 828, 843 thresholds, 84, 233, 477, 605, 649 thrombosis, 138 tides, 1021 time commitment, 880 time constraints, 351, 789 time factors, 71 time frame, 6, 87, 280 time periods, 70, 817, 819, 820, 1066, 1111 time preferences, 66, 71, 686 time pressure, 727 time series, xxv, 751, 752, 763, 765 time variables, 1047 time-frame, 1019

Thailand, vi, xix, 391, 392, 393, 394, 395, 396, 399,

timing, 65, 67, 68, 74, 146, 283, 362, 370, 513, 881, 1015, 1023, 1027, 1029, 1030, 1031, 1067, 1133

tissue, 238, 338, 437, 587 title, 52 T-maze, 1025 tobacco, xx, 466, 467, 470, 473, 474, 504, 507, 513 toddlers, 479 tolerance, 54, 60, 366, 369, 458, 609, 653, 870, 874, 879, 883, 990 tonic, 245, 247, 248, 250, 251, 253, 254 top management, 1103 top-down, 12, 243, 464 tornadoes, 748, 826, 828, 831, 833 torture, 519, 526, 877 total energy, 604, 647 totalitarian, 18, 732 tourism, 869, 885 toxic, 104 toxicity, 31, 32, 43 tracking, xxix, xxx, 375, 474, 1013, 1014, 1015, 1018, 1019, 1045, 1068 trade, xiii, 10, 31, 43, 63, 64, 67, 70, 281, 458, 463, 535, 600, 617, 644, 661, 691, 702, 712, 801, 902, 918, 921, 923, 930, 973, 1020, 1083, 1107 trade-off, xiii, 10, 31, 43, 63, 64, 67, 70, 281, 458, 600, 617, 644, 661, 691, 702, 918, 921, 923, 930, 973, 1020, 1083 trading, 71, 923 tradition, 303, 304, 548, 795, 953 traditional healers, 395, 403, 431, 435, 440 traditional model, 752 traditionalism, 999 traffic, xxvi, 841, 842, 843, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 858, 859, 860, 862, 863, 864, 865, 1078, 1086, 1087, 1090 traffic violations, 859, 860 training, 7, 53, 54, 56, 57, 58, 60, 61, 153, 166, 182, 261, 271, 274, 287, 402, 403, 404, 434, 451, 452, 463, 464, 474, 504, 505, 542, 577, 608, 651, 716, 726, 792, 860, 862, 871, 873, 874, 875, 876, 877, 879, 880, 989, 1005, 1009, 1017, 1024, 1036, 1038 training programs, 56, 57, 860 traits, 50, 51, 114, 141, 142, 672, 673, 726, 998, 999, 1004, 1097 trajectory, 286, 295, 344, 362, 477 trans, 168, 768 transcendence, xxvii, 197, 867, 868 transcript, 122, 318, 319 transcription, 122, 401, 404 transcripts, 122, 319, 320, 405 transfer, xii, 45, 47, 52, 55, 87, 222, 462, 678, 679, 916, 924, 925, 928, 930, 995 transformation, xxii, 179, 528, 557, 569, 897, 1060 transformations, 897, 1057, 1079

transfusion, xiii, xiv, 93, 94, 95, 106 transfusions, xiii, 93, 94, 105 transition, xxix, 280, 394, 396, 402, 458, 461, 466, 492, 505, 558, 573, 576, 624, 633, 773, 999, 1013, 1016, 1018, 1019, 1022, 1026, 1030, 1118 transitions, 259, 471 translation, 168, 402 transmission, 253, 256, 312, 620, 622, 624, 924, 928, 983, 984, 1008 transparent, 10, 12, 22, 70, 555 transplant, xiii, xiv, 93, 94, 95, 97, 98, 99, 105 transplantation, vi, xiii, xiv, 93, 94, 95, 96, 97, 98, 100, 105, 106, 107, 109, 110, 111 transport, 132, 136, 188, 255, 691, 701, 987, 1107 transportation, xx, 457, 460, 461, 464, 465, 466, 484, 842 transpose, 755 traps, 13 trauma, xxi, 169, 396, 443, 478, 486, 517, 518, 519, 523, 524, 525, 526, 527, 528, 842, 876 travel, xx, 138, 146, 360, 503, 505, 606, 607, 608, 609, 610, 617, 649, 650, 651, 653, 660, 868, 870, 872, 879, 880, 883, 1029, 1032, 1073, 1074, 1075, 1076, 1077, 1078, 1079, 1081, 1082, 1087, 1090, 1091, 1092, 1093 travel time, 606, 607, 608, 649, 650, 651, 879, 1029, 1077, 1078, 1079, 1091, 1092, 1093 treatable, 286, 482 treatment, 264, 266 trees, 14, 84, 89, 357, 1018 trend, 176, 762, 918, 957, 1035, 1036 triage, 467, 469 trial, xxv, 1, 29, 30, 31, 41, 42, 111, 148, 150, 151, 156, 195, 196, 214, 227, 243, 470, 472, 474, 547, 550, 604, 606, 608, 611, 647, 650, 651, 654, 666, 667, 668, 709, 767, 768, 772, 775, 776, 778, 785, 786, 787, 788, 790, 791, 794, 795, 796, 797, 895, 915, 916, 966, 987, 988, 989, 1017, 1020, 1028, 1053, 1117, 1118, 1122, 1123, 1125 trial and error, 895, 987, 988 triggers, 119, 208, 1049 trucking, 463 true belief, 548, 550, 552, 553, 554 trust, 6, 58, 73, 165, 186, 194, 213, 222, 231, 285, 331, 401, 403, 406, 417, 440, 488, 490, 713, 805, 806, 807, 809, 827 trusts, 803, 805 trustworthiness, 40 tryptophan, 233, 1133 tubal ligation, 576, 577, 585, 587, 588 Tucker-Lewis Index (TLI), 852, 858 tuition, 594, 638

tumour, 28, 281

tumours, 28 Turkey, 825 turnover, 46, 55, 1100 two-dimensional, 566 two-way, 1094, 1095, 1123, 1125, 1126 typology, 340, 862

# U

ubiquitous, 1101 Ultraviolet, 673 uncertain outcomes, 769, 774 unconditioned, 208, 246 undergraduate, 623, 624, 630, 736, 739, 955, 957, 971 undergraduates, xxii, 557, 561, 562, 1051 underwriters, 901, 936 unemployment, 163, 408, 580, 714, 715, 716, 719 UNESCO, 393, 455, 1010 uniform, 961, 965, 966, 967, 977 Uniform Distribution, 961, 965, 967 uninsured, 938 unions, 580 United Kingdom (UK), xiv, 109, 114, 115, 116, 119, 157, 159, 163, 177, 182, 261, 264, 266, 275, 276, 290, 295, 308, 327, 333, 381, 384, 388, 482, 489, 541, 593, 624, 635, 636, 637, 676, 677, 678, 683, 685, 688, 690, 699, 700, 705, 789, 792, 793, 799, 889, 911, 1042, 1100, 1110 United States (US), xvii, xxii, xxiii, 47, 52, 154, 162, 195, 266, 267, 268, 275, 297, 298, 302, 303, 309, 310, 470, 489, 514, 515, 538, 571, 592, 619, 620, 633, 634, 635, 636, 948 univariate, xiii, 75, 82, 83, 84, 563 universal law, 980 universities, xxix, 981, 1010 university students, xxiii, 567, 619, 625, 626, 635, 684, 710, 981, 1010, 1035, 1036, 1038, 1039, 1040, 1051 unmarried women, 393, 404, 415, 444 unplanned pregnancies, xix, 363, 369, 391, 392, 393, 395, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 410, 411, 413, 416, 418, 419, 420, 422, 425, 428, 429, 430, 432, 437, 440, 444, 445, 446, 448, 449, 450, 451, 452, 453 unpredictability, 281, 330, 519, 610, 654 updating, 26, 28, 802, 807, 1126 upload, 634 upward comparisons, xxix, 951, 971, 977 urban, 533 urban population, 575 urbanization, 409, 574, 576 urinary, 480, 497

urinary tract, 480, 497 urinary tract infection, 481, 497 urine, 323, 413, 417, 418, 432, 441 URL, 515 US dollar, 756 USDA, 13 users, xiv, xv, 114, 157, 158, 159, 163, 169, 172, 173, 175, 176, 177, 261, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 395, 449, 450, 453, 463, 775, 848, 864, 876 uterus, 428, 772, 786, 787 utilitarianism, 886

# V

vaccination, 186, 620

- vacuum, 164, 589
- vagina, 437, 439, 623, 624, 625, 626
- vaginal, 518, 525, 623, 624, 625, 627, 629, 634, 636
- valence, 3, 214, 794, 1070, 1126
- validation, 77, 78, 79, 80, 83, 88, 89, 91, 170, 171, 386, 566, 568, 937, 1001, 1036
- validity, xxx, 2, 69, 78, 79, 80, 88, 186, 196, 227, 274, 320, 401, 402, 473, 490, 492, 562, 566, 567, 568, 612, 628, 656, 704, 719, 753, 791, 853, 854, 855, 907, 929, 996, 998, 1024, 1028, 1046, 1050, 1052, 1053, 1057, 1058, 1060, 1061, 1064, 1068, 1069, 1091
- valine, 1118
- vandalism, 458
- variability, xi, xxix, 5, 31, 33, 35, 114, 147, 148, 242, 258, 300, 477, 604, 626, 648, 660, 682, 683, 687, 688, 689, 702, 717, 762, 763, 764, 871, 874, 918, 1013, 1014, 1015, 1016, 1017, 1018, 1019, 1021, 1022, 1024, 1026, 1027, 1028, 1031, 1111, 1134
- variance, xxviii, 33, 330, 563, 595, 600, 604, 605, 615, 639, 643, 648, 658, 675, 740, 746, 752, 754, 758, 759, 762, 816, 817, 846, 852, 858, 951, 957, 989, 993, 998, 999, 1066, 1067, 1068, 1085, 1089
- variation, 14, 33, 34, 42, 72, 115, 154, 175, 374, 492, 605, 614, 648, 658, 664, 672, 674, 678, 758, 828, 916, 1017
- vasectomy, 576
- vector, 756, 816, 939, 1081, 1082, 1088
- vegetables, 472

vegetative state, 191

- vehicles, 177, 185, 687, 698, 844, 918, 1087
- vein, 199, 698
- velocity, 980
- ventilation, 187, 188, 189, 192, 301
- Ventral tegmental area, 256
- Venus, 51, 61

verbal abuse, 46 verbal justifications, xxi, 541 versatility, 363, 370 vertebrates, 674 Vice President, 55 victimization, 476, 480, 496, 497, 499, 521 victims, xx, 163, 475, 476, 479, 480, 486, 487, 488, 492, 495, 500, 573, 876 Victoria, 771, 794, 1099 Videoconferencing, 996 videotape, 491 vigilante, 525 village, 128, 308 violence, 77, 78, 80, 81, 88, 90, 91, 163, 267, 412, 477, 478, 481, 482, 483, 486, 489, 506, 519, 529, 586,859 violent, xxv, 78, 81, 82, 83, 85, 90, 489, 525, 586, 735, 826, 844 violent behavior, 78, 82 violent crime, 81, 85 violent crimes, 81, 85 violent offenders, 85 violent recidivism, 83 virus, 96, 634 visible, 164, 262, 484, 525, 543 vision, 381, 673 visual, 263, 521 visual area, 208, 242 visual images, 186, 188, 189, 193 visual perception, 238 visual stimuli, 223 vocational, xx, 396, 413, 418, 421, 423, 427, 432, 434, 435, 437, 441, 442, 446, 457, 463, 466, 468, 535, 558, 563, 564, 565, 567, 568 vocational identity, 558, 567 vocational training, 434, 463 voice, 52, 169, 177, 188, 189, 800 voiding, 350, 370 volatility, 688, 754, 757, 758 Volunteers, 739 vomiting, 99, 101, 103, 104, 411 voting, 13 vulnerability, xxi, 133, 271, 350, 353, 394, 480, 498, 517, 528, 621 vulnerable people, 271

# W

wages, 411, 589, 713, 714, 715, 716, 717, 719 waiting times, 605, 609, 649, 652 waking, 241 Wales, 150, 275, 716, 722, 789, 795, 1099 walking, 115, 119, 128, 129, 131, 133, 134, 138, 145, 146, 313, 424, 441, 442, 472, 579, 868, 879, 942, 1078 war, 163, 266, 382 war hero, 163 warfare, 163 Warsaw, 908 Washington, 496, 497, 500, 514, 515, 538, 539, 634, 636, 839 Washington Post, 636 waste, 827, 828 watches, 814, 882 water, 666, 667, 668, 691, 701, 818, 870, 872, 875, 877, 881, 882, 986, 987, 1021 Watson, 825, 839 weakness, 288, 728, 731, 1027, 1078 wealth, xxiii, xxxi, 16, 37, 168, 198, 593, 594, 595, 599, 637, 638, 639, 642, 643, 712, 713, 737, 753, 836, 891, 895, 896, 938, 939, 1085, 1099, 1100, 1103, 1110, 1111, 1112 weapons, 460 wear, 140, 142, 250, 843, 876, 882 weathering, 823 web, 182 web-based, 182 weight control, 466 weight loss, 170 welfare, 22, 163, 304, 478, 489, 505, 620, 699, 895, 1134 welfare reform, 620 welfare system, 163 wellbeing, 200, 350 well-being, xiii, 3, 18, 93, 95, 169, 179, 200, 227, 228, 231, 303, 304, 321, 329, 331, 350, 408, 482, 489, 505, 514, 580, 582, 678, 733, 952, 980 western countries, 286, 350 western culture, 163, 170, 559 Western Europe, 264 wheat, 987, 988 wheat germ, 987 wheelchair, 128, 135, 188 wholesale, 346 Wikipedia, 628, 629, 636 wild animals, 1022 wilderness, 871, 877 Williamsburg, 663 wind, xxvi, 748, 811, 813, 815, 816, 818, 819, 820, 821, 822, 823, 833, 1014, 1020 windstorms, xxv, 735 wine, 179, 535 winning, xxiv, xxix, 16, 160, 595, 598, 639, 641, 681, 707, 712, 727, 729, 731, 736, 893, 894, 896,

898, 900, 919, 1033, 1034, 1037, 1040, 1116, 1118, 1119, 1123 wisdom, 22, 107 withdrawal, 134, 208, 217, 250, 287, 295, 411, 437, 523, 576, 584, 728, 731 witness, 495, 500 witnesses, 490 wives, 120, 129, 132, 412, 435 wood, 434, 574 word of mouth, 421, 424, 425, 427, 429 words, 269, 493, 495 work, xvi, 261, 262, 263, 265, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 486, 492, 495 workers, 51, 52, 72, 169, 261, 264, 265, 266, 269, 270, 271, 272, 274, 299, 336, 393, 400, 403, 415, 449, 488, 492, 575, 581, 584, 589, 713, 714, 715, 716, 717 workforce, 60, 578, 982 working class, 583, 584 working memory, 17, 205, 206, 212, 215, 217, 218, 219, 220, 224, 226, 228, 230, 243, 922, 1023, 1035, 1038, 1041, 1130, 1131 working population, 678 workplace, 60, 162, 303, 381, 423, 1091, 1092 World Health Organization (WHO), 115, 394, 455, 476, 488, 506, 771, 863 worldview, 158, 989 worm, 1023 worms, 1023

worry, 16, 67, 129, 164, 285, 329, 365, 371, 409, 410, 422, 437, 442, 543, 544 wrists, 882 writing, 291, 302, 331, 522, 880, 989, 992, 996, 1052 WTP, 718, 826, 827, 828, 832, 834, 835, 836

Х

X-linked, 348, 380, 384 x-rays, 126

yes/no, 550, 829

yield, 33, 492, 558, 607, 651, 738, 986, 1020, 1025, 1052

Υ

young adults, xxiii, 38, 43, 467, 472, 500, 558, 559, 619, 623, 624, 627, 628, 960

young men, 267, 580, 861, 863

young women, xix, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 409, 410, 411, 412, 413, 414, 416, 418, 420, 421, 422, 423, 424, 427, 428, 429, 430, 432, 434, 440, 441, 442, 444, 445, 448, 449, 450, 452, 453, 454, 513, 583

younger children, 524, 1040



.









