RUSSELL ON METAPHYSICS

Selections from the writings of
Bertrand Russell

Edited by Stephen Mumford

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

RUSSELL ON METAPHYSICS

Russell on Metaphysics brings together for the first time a comprehensive selection of Russell's writing on metaphysics in one volume. Russell's major and lasting contribution to metaphysics has been hugely influential and his insights have led to the establishment of analytic philosophy as a dominant stream in philosophy. Stephen Mumford chronicles the metaphysical nature of these insights through accessible introductions to the texts, setting them in context and understanding their continuing importance.

The volume is divided into five parts following the development of Russell's thought. Starting with papers from his pre-analytic period, the volume collects Russell's main realist accounts, his discussions of the problems of universals, and his writing on causation and the laws of nature. The final part covers Russell's thoughts on diverse questions of metaphysics that occupied him later in life, including his classic paper on vagueness.

Russell on Metaphysics is both a valuable introduction to Bertrand Russell as a metaphysician, and to analytic philosophy and its history.

Stephen Mumford is Reader in Metaphysics at the Department of Philosophy, University of Nottingham. He is author of *Dispositions* (1998) and various papers in metaphysics.

RUSSELL ON ...
General editor's introduction
A. C. Grayling

Russell achieved public fame – often enough, notoriety – because of his engagement in social and political debates, becoming known to a wide audience as a philosopher in the popular sense of the term. His chief contributions, the ones that have made a permanent difference to the history of thought, lie in logic and philosophy; and they are such that his influence both on the matter and style of twentieth-century philosophy, principally in its anglophone form, is pervasive. Elsewhere I have described his contribution as constituting the "wall-paper" of analytic philosophy, in the sense that his successors "use techniques and ideas developed from his work without feeling the need – sometimes without recognizing the need – to mention his name; which is influence indeed".

Russell devoted much attention to central technical questions in philosophical logic, epistemology and metaphysics. He also wrote extensively and forcefully about moral, religious and political questions in ways not merely journalistic. Much of his work in all these areas took the form of essays. Some have of course been famously collected, constituting a fundamental part of the canon of twentieth-century analytic philosophy. But there are many more riches in his copious output, their value to some degree lost because they have not hitherto been collected and edited in such a way as to do justice to the development and weight of his thinking about these subjects. This series, in bringing together Russell's chief writings on major subject areas in an editorial frame that locates and interprets them fully aims to remedy that lack and thereby to make a major contribution both to Russell scholarship and to contemporary analytic philosophy.

Russell on Religion
Edited by Louis Greenspan and Stefan Andersson

Russell on Ethics
Edited by Charles R. Pigden

RUSSELL ON METAPHYSICS

Selections from the writings of Bertrand Russell

Edited by Stephen Mumford

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2003
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2003.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Russell, Bertrand, 1872–1970.
Russell on metaphysics: selections from the writings of Bertrand Russell / edited by Stephen Mumford.
p. cm.
Includes bibliographical references and index.
1. Metaphysics. I. Mumford, Stephen. II. Title.
B1649.R91 M86 2003 110–dc21 2002190865

ISBN 0-203-41752-6 Master e-book ISBN

ISBN 0-203-41900-6 (Adobe eReader Format)
ISBN 0–415–27744–2 (hbk)
ISBN 0–415–27745–0 (pbk)

CONTENTS

ACKNOWLEDGEMENTS

I am grateful to the publishers who have given permission for material to be included. The most important source of Russell's papers is *The Collected Papers of Bertrand Russell*, edited at McMaster University, London: Routledge (1983–present). Many volumes are already in print but the project is ongoing. References to these volumes will be give throughout as '*Papers*', followed by the volume number. In the Bibliography, references to Russell's papers will be given for their location in this edition, rather than the original place of publication.

Paper 8 was used with permission of Douglas Lackey and the Bertrand Russell Archive at McMaster University Library.

I am grateful to Charles Pigden for offering his thoughts on this collection and, in particular, for suggesting the inclusion of paper 18.

The papers in this collection were originally published in the following places:

1 Greek Exercises
First appeared in full in *Papers* 1: 3–20 (3–7, 10, 12, 20 only in this collection).

2 On the Distinction Between the Psychological and Metaphysical Points of View
First appeared in *Papers* 1: 195–8.

3 The Free-Will Problem from an Idealist Standpoint
First appeared in *Papers* 1: 229–39.

4 "The Dialectic of the Sciences"
From selections that first appeared in full in *Papers* 2, Part I (5, 11, 12, 14–16, 84, 92–7 only in this collection).

5 Seems Madam? Nay It Is
First appeared in *Why I Am Not A Christian*, P. Edwards (ed.), London: George Allen & Unwin, 1957: Ch. 5; reprinted in *Papers* 1: 106–11.

6 *The Principles of Mathematics*
London, George Allen & Unwin, 1903. Paperback edition, London, Routledge, 1992: 42–52, 449–50.

7 The Existential Import of Propositions
First appeared in *Mind*, 14 (1905): 398–401; reprinted in *Papers* 4: 486–9.
8 Three Letters to Meinong
First appeared in *Russell*, O.S. 9 (1973): 15–18.
9 The Basis of Realism
First appeared in *The Journal of Philosophy, Psychology and Scientific Methods*, 8 (1911): 158–61; reprinted in *Papers* 6: 128–31.
10 Analytic Realism
First appeared as "Le Réalisme analytique" in *Bulletin de la société française de philosophie*, 11 (1911): 282–91; reprinted in *Papers* 6; 133–46 (133–7 only in this collection).
11 The Philosophy of Logical Atomism
First appeared in *The Monist*, 28 (1918): 495–527 and 29 (1919): 32–63, 190–222, 345–80; reprinted in *Papers* 8:160–244 (163–8, 187–90 and 234–44 only in this collection).
12 On the Relations of Universals and Particulars
First appeared in *Proceedings of the Aristotelian Society*, 12 (1911–12): 1–24; reprinted in *Papers* 6: 167–82.
13 The Problem of Universals
First appeared in *Polemic*, 2 (1946): 21–35; reprinted in *Papers* 11: 258–73.
14 On the Notion of Cause
First appeared in *Scientia*, 13 (1913): 317–38; reprinted in *Papers* 6: 193–210.
15 Causal Laws in Physics
First appeared in *An Outline of Philosophy*, London, George Allen & Unwin, 1927: Ch. 11.
16 Physics and Metaphysics
First appeared in *The Saturday Review of Literature*, 4 (1928): 910–11; reprinted *Papers* 10: 272–8.
17 Causal Laws
First appeared in *Human Knowledge: Its Scope and Limits*, London: George Allen & Unwin, 1948: Part IV, Ch. 9.
18 Vagueness
First appeared in *The Australasian Journal of Psychology and Philosophy*, 1 (1923): 84–92; reprinted in *Papers* 9: 147–54.
19 Physics and Neutral Monism
First appeared in *The Analysis of Matter*, London: George Allen & Unwin, 1927: Ch. 37.
20 Language and Metaphysics
First appeared in *An Inquiry Into Meaning and Truth*, London: George Allen & Unwin, 1940: Ch. 25.
21 The Principle of Individuation
First appeared as "Le Principe d'individuation" in *Revue de métaphysique et de morale*, 55 (1950): 1–15; reprinted in *Papers* 11: 294–303.

INTRODUCTION

This book is a collection, along with a commentary, of the most significant writings on metaphysics by Bertrand Russell. Unlike the other books published in this series so far, the heading under which Russell's writings are selected is one that may not be fully understood in advance by some readers. Russell discussed many things, including politics, religion and ethics. He was, however, one of the greatest analytic philosophers of the twentieth century and this book includes some of the writings for which he deserves this status. Some of the ideas Russell discusses here may be difficult, therefore. But Russell thought that in almost all areas of philosophy, clarity and simplicity was possible and that even very difficult ideas could be stripped down to their easily grasped essentials. He successfully demonstrates this in these papers. Thus, someone completely new to metaphysics ought nevertheless be able to understand what is said. Indeed, this collection would provide a fine introduction to the subject or to analytic philosophy in general. Though it is not a student textbook nor a research monograph, it is the work of a prominent and important philosopher engaged in metaphysical study. Perhaps there is no better introduction to metaphysics than such a thing. Further, this book might be an informative introduction to analytic philosophy and its history. Some of the papers contain the metaphysical development that underpinned the transition from British idealism to contemporary British philosophy.

To some who know a little of Russell's philosophy, it might seem strange to speak of him being engaged in metaphysics. He is often depicted as standing squarely in the empiricist tradition that had, on the whole, rejected metaphysics and was concerned primarily with the theory of knowledge or epistemology. If this book has but one aim, it is to relieve its readers of that misconception. Russell was a metaphysician. It might even be thought that his most important and lasting philosophical insights were metaphysical. Though famed for his work on the foundations of mathematics and logic, perhaps his greatest achievement was thinking of the metaphysical basis on

which to build his mathematics and logic. Arguably, the basis remains even though much of the detail of his logical and mathematical work has been rejected or surpassed. Most importantly, Russell was able to see an alternative to the metaphysics of the dominant idealism with which he grew up. In so doing, he has the strongest claim to being the father of analytic philosophy.

When British idealism was rejected, the reputation of metaphysics in general suffered, despite the fact that Russell and his friend G. E. Moore rejected idealism for metaphysical reasons. Subsequent developments saw metaphysics fall into even greater disrepute. The logical positivists, for instance, spoke of metaphysics as nonsense. It was hardly surprising, therefore, that very little metaphysics was being produced. Because of this, during the first half of the twentieth century, Russell might well have been the most productive analytic metaphysician. This accolade has been disguised, though, because Russell was so productive in many other areas as well. Only a very small portion of Russell's works are directly and overtly metaphysical. But even a small portion of Russell's vast corpus, adds up to a large contribution to the subject. It is thus hard to think of anyone who made as big a contribution to metaphysics in the analytic tradition during this period.

In the rest of this introduction, I aim to give a background for the papers collected. I start with the nature of metaphysics so that readers new to the subject might get an understanding of it before embarking on it. I then move on to the history of metaphysics around the time these papers were written, concentrating on the background of British idealism. The final section considers Russell as a metaphysician and assesses his most important claims. In addition to this introduction, however, I have provided an introduction to each of the selections reproduced so as to set out some more detailed and subject-specific background.

WHAT IS METAPHYSICS?

Metaphysics is one of the most abstract areas of philosophy. Its abstract nature can be quite intimidating for those who do not understand its methods and aims. So as not to exclude such people from the enjoyment of Russell's contribution to metaphysics, I will begin by offering an explanation of the subject. The most difficult part of metaphysics is understanding what it is about. Once one gets to grips with that, former difficulties can melt away.

There are a number of definitions of metaphysics on offer and these vary in the degree that they are interesting or helpful. I will outline the definition that I find most helpful for getting into a metaphysical frame of mind.

Metaphysics aims to uncover the fundamental nature of reality beyond appearance. It studies the world, but not anything about it that can be observed. It follows from this that the questions of metaphysics cannot be settled empirically, by looking for observable evidence, but must be solved using philosophical methods of analysis, reason and argument. Russell himself followed this way of characterising metaphysics when he said "By metaphysical entities I mean those things which are supposed to be part of the ultimate constituents of the world, but not to be the kind of thing that is ever empirically given" (from paper 11, reproduced below).

There are occasions when the physicist and metaphysician are considering the same subject. They might both consider, for instance, what it is for something to be a material substance. But they can consider this subject in very different ways. Similarly, if I punch someone and am observed by a moral philosopher, a lawyer and boxing promoter, they may each think very different things about the same action. Likewise with the metaphysician and physicist studying material substance.

One traditional debate in metaphysics is whether substances are merely bundles of qualities or underlying substrata that hold such qualities together. The issue is discussed by Locke (1690: Book II, ch. 23). In the first view, if the properties of being black, furry, four-legged, smelly, of a certain height and appearance, and so on, are collected together, then this is all that is required for there to be a dog at the place where the properties are gathered. The substratum theorist argues that there is something more that is needed. This something extra might be called "particularity". A problem with the bundle theory is that there might be two particulars that have all the same properties. Suppose our dog has a twin. Would we not want to say that there are two particulars with all the same properties? But if they have all the same properties, how are they to be distinguished other than by being two separate particulars, where the distinct particularity is not accountable in terms of some difference in properties?

Resolving this debate between bundle theorists and substratum theorists about substance is not my aim here. The example is intended to serve as an illustration, though Russell's discussions of the subject can be found in papers 15 and 21. However we are to resolve the debate, we cannot do so empirically. We cannot judge the debate on the basis of how the world looks. Indeed, we could say that the world would look the same to us whichever of the bundle theory or substratum theory is true. Because science deals with the observable, it is, therefore, no use at all to us on this subject. A scientist might sometimes refer to things that are not directly observable, such as sub-atomic particles. But they are at the very least committed to the view that there could, in principle, be some observations which confirm or falsify the existence of sub-atomic particles, otherwise the

subject is not scientific. What a scientist says, therefore, is usually of little use in a metaphysical debate. A physicist might say, for instance, that material substances are ultimately made out of electrons and such things. This answer is of no use to the metaphysician because the very same metaphysical question can arise: is an electron a bundle of qualities or does it have an underlying substratum?

If metaphysicians cannot decide between rival theories on the basis of observations, on what basis can they decide? In short, they decide on the basis of argument and analysis. In doing so, the metaphysician is no different from the moral philosopher, the epistemologist, logician, or practitioner of any other branch of philosophy. Metaphysics is no more puzzling than philosophy in general. A moral philosopher, for example, cannot find answers to their questions empirically. They, like the metaphysicians, are dependent upon reason and argument. The moral philosopher might ask whether it is good to procure the greatest happiness of the greatest number. They cannot decide this by observing groups of happy people. The question is abstract and non-empirical. There is still something to decide after all the observed facts are collected. Collecting such facts is often of no use at all to the philosopher.

Thus metaphysics is no more mystifying than any other branch of philosophy. It deals with the nature of reality but not the empirical part of it, just as moral philosophy deals with right actions without investigating them empirically or epistemology deals with the theory of knowledge non-empirically.

RUSSELL AND METAPHYSICS

Russell became interested in philosophy as a juvenile, at a time when metaphysics was a prominent and fashionable discipline. British metaphysics was mainly under the influence of G. W. F. Hegel's system. Little could Russell have known that he would be one of the main figures to destroy Hegel's reputation in Britain. For while Hegel was an influence on British metaphysics and on Marxism, in the twentieth century there would be reactions against it. In continental Europe, that reaction would take the form of existentialism. In Britain, it would take the form of analytic philosophy. Metaphysics would not have too honoured a place in early analytic philosophy because, to an extent, such philosophy was a rejection of metaphysics. Metaphysics was important in British Hegelianism and when Hegelianism was spurned, so too was its beloved metaphysics. It is arguable, however, that there was an underlying metaphysical basis to analytic philosophy which was distinctly anti-Hegelian and it was Russell who best voiced that basis.

There are two excellent book-length studies of this period in philosophy and Russell's role in it (Griffin 1991 and Hylton 1990). It will be useful here, however, to have a short summary, if only to show the confusing background from which Russell's work emerged. Hylton says of Russell's early work that "Idealism, and the need to refute it, forms a crucial context of this period of Russell's work" (Hylton 1990: 8). In Parts I and II of the current book, I will be offering primary source evidence in support of this view. The idealism with which Russell was concerned was mainly that of Bradley, which had come, through T. H. Green, from Hegel. Green offered a Hegelian version of Kantianism that became the standard British Hegelian philosophy. He began by attacking Locke's version of empiricism for denying the reality of relations and yet requiring them for a plausible account of knowledge. Thought is a pre-condition of knowledge in the way Kant set out. All knowledge is knowledge of the phenomenal world, which consists of the way the world appears to us. We impose form and content upon what is given to us – we impose a structure which is responsible for notions of space and time, causation, substance and so on. Such notions are not derived from experience but are a pre-condition of experience. Green put a Hegelian twist on this orthodox Kantianism. He rejected the notion of a world in itself, separate from thought. There is no *given*, prior to experience. But he also saw the phenomenal world as problematic. Whose mind constitutes the phenomenal world? Green says it is the single spiritual being of which all reality is the activity or expression. Our own minds are manifestations of this single self-consciousness. The whole of reality is constituted by thought, therefore.

Bradley was the most prominent philosopher in Britain in the 1890s, when Russell was beginning his philosophical writing. Bradley's most important book, *Appearance and Reality*, had been published in 1893. In "The free-will problem from an idealist standpoint" (paper 3), Russell describes *Appearance and Reality* as an "epoch-making work" and G. E. Moore had a similarly high evaluation for a time (Hylton 1990: 44). But by 1897 both Russell and Moore had come to reject Bradley and the whole tradition of which he was the culmination. Bradley argues that neither immediate experience nor relations are real, and from the latter we can infer that no ordinary phenomena are real and that there is no absolute truth or falsehood. Reality is a single togetherness rather than being many distinct, related things. The things that appear to us as distinct individuals are actually aspects of the comprehensive, concrete individual, which Bradley calls the Absolute.[1]

1 See also Bradley 1914 for possibly a simpler presentation of these ideas.

In rejecting Bradley and idealism, Russell and Moore came to be realists. They accepted as real all the everyday, common sense, things that Bradley had told us were mere illusions. Whereas Bradley wanted to push appearance and reality apart, Russell and Moore sought to bring them together.

Such a desire, alone, was insufficient, however. Bradley had arguments for his position and Russell and Moore needed refutations if they were going to offer a serious alternative to idealism. By good fortune, Russell discovered the work of Gottlob Frege (Frege 1980) and, from there, works of others in the theory of logic. He describes the accident in his autobiography:

> James Ward was always giving me fresh books … and each time I returned them, saying that they were very bad books. … In the end, but after I had become a Fellow, I got from him two small books, neither of which he had read or supposed of any value. They were Georg Cantor's *Mannichfaltigkeitslehre*, and Frege's *Begriffsschrift*. These two books at last gave me the gist of what I wanted, but in the case of Frege I possessed the book for years before I could make out what it meant. Indeed, I did not understand it until I had myself independently discovered most of what it contained.
>
> (Russell 1967: 68)

What the new work meant was that the logic upon which Bradley and others had based their arguments was totally inadequate. Frege had what would be the foundations of a new logic. Bradley's position was based on an erroneous belief, among others, that all arguments must be based on subject-predicate logic, which had little changed since Aristotle. By discovering the new logic, Russell was able to diagnose all the fallacies of Bradley and the idealists that had lead them to so counterintuitive a position.

The new logic had a metaphysical basis, however. It assumed all sorts of things that Bradley had rejected. It assumed certain objects such as real and mind-independent propositions. It assumed objective truth and falsehood, regardless of belief. It assumed the existence of relations with an independence that was external to their relata. It also assumed a plurality of objects. Russell discussed some of these questions of ontology in the papers collected in Part II. An ontology is simply an inventory of what there is. For a metaphysician, this will be a list of the categories of things that exist, such as propositions, properties and relations.

Russell's own book, *The Principles of Mathematics* (1903), was one of the key texts in the foundations of analytic philosophy. British idealism was vanquished. Russell, and others, were able to work on a new programme and the remaining papers of this collection were all created within the new tradition.

RUSSELL AS A METAPHYSICIAN

As well as collecting Russell's main papers in metaphysics, this book tells also the story of Russell's philosophical thinking. Certainly in the first phases of Russell's thought, this also shows how metaphysical concerns were driving the whole of his philosophical development.[2] Hence, the establishment of analytic philosophy in the Anglo-American world – which in no small part is down to Russell – can be understood as springing from a metaphysical insight. This insight can be summed up as *realism*. Perhaps G. E. Moore had the insight initially, and Gottlob Frege had provided the chief analytic tool, but it was Russell who brought these two together and, in so doing, became the most important early exponent of the new approach.

A pre-condition of understanding the importance of Moore's insight to Russell is an awareness of Russell's training in the idealist metaphysics of the day. Part I of this book concerns Russell's move into, and then back out of, the British Hegelian tradition. The dominant metaphysics day was passed to him through his teachers at Cambridge, James Ward, G. F. Stout and John M. E. McTaggart. At times Russell was Kantian and at other times Hegelian. His Hegelianism had a lasting influence to the extent that the "Tiergarten programme" was seen through (see paper 4), long after its Hegelian *raison d'être* was rejected. His second book, *An Essay on the Foundations of Geometry* (Russell 1897), was both Kantian (in its first part) and Hegelian (in its second). In the first part, Russell argues that certain features of space, and thereby geometry, are known a priori and other features are known only by experience. In the second part, he shows the contradictions inherent in our notion of space, and the need to move to the next stage in the dialectic: taking matter to be real.

But before Russell could complete his Hegelian synthesis of all the sciences, he came under the influence of Moore. Part I ends with a paper that is a landmark in Russell's thinking, "Seems Madam? Nay, It Is". Unpublished at the time, this paper has since appeared in a variety of places but it is so important, in representing the major turning point in Russell's thinking, that it must appear in this collection.

Having made this break, Part II shows how Russell developed his new realist ontology. Because he was largely concerned with the technical detail of his project, most notably in *Principia Mathematica*, with A. N. Whitehead (Russell and Whitehead 1910–13), direct discussion of the broader metaphysical underpinning to the project was comparatively rare. The papers in this part were thus stretched out over a number of years. The rarity with

2 Russell gives his own, detailed account of the transition in Russell 1959: chs 4–9.

which Russell discussed such issues does not indicate a relative lack of importance to his work, however. Without the metaphysical picture in place, there would have been no point in Russell working out the detail.

In the papers in Part II, Russell describes some of the foundation stones of analytic philosophy. He takes the world to contain facts, for instance, "which are what they are whatever we may choose to think about them" (paper 11). Hence, Russell had moved from idealism to a mind-independent view of the world. There are also propositions, which are made true by facts, though they do not name facts. The analytic technique is described and justified, where what is complex is to be analysed into that which is simple because "the existence of the complex depends on the existence of the simple, and not vice versa" (paper 10). The most important doctrine, as Russell claims in various places, such as paper 9, is the doctrine of external relations, which allows the same entity to be a part of many different complexes. Without this latter pluralist insight, there would be no standard extensional logic in the form we now know it. Analytic philosophy was developed in these papers, though it was given its paradigm exemplar in "On denoting" (Russell 1905).

The difficulty of having a metaphysical basis to one's philosophy, however, is that if the metaphysical basis is wrong, so too might be the philosophy. Most of the basic tenets that form the basis of analytic philosophy have been challenged, though some only recently. Hence, extensional logic has been attacked for being based on an erroneous Humean metaphysic (Ellis 2001: ch. 8). Logical analysis appears to rest on an unfounded assumption that there are simples, which is a point Russell was to concede (paper 11). Whether there are mind-independent facts is an issue that has provoked a vast literature. Such a re-appraisal of these assumptions of analytic philosophy is in its early stages, however. Even if it is eventually decided against certain metaphysical commitments, the techniques developed in analytic philosophy are likely to endure. It is unimaginable, for instance, that we should judge that Frege and Russell were wrong to see Aristotelian logic as in need of improvement. If we are able to improve on Russell's metaphysics at all, we will be fortunate indeed if we can do so to the same extent that he improved on idealist metaphysics.

Russell put aside the questions of the foundations of analytic philosophy and moved on to some more detailed studies of specific problems in philosophy. Parts III and IV collect papers on two issues he revisited at various points in his career, spread over a distance of years.

Part III collects Russell's work on the problem of universals, which is one of the central topics of metaphysics in Western philosophy, with a history going back at least as far as Plato. Russell's initial "solution" to the problem, in *The Problems of Philosophy* (1912: chs 9 and 10) did not differ too

far from Plato's. Plato's transcendent realism about universals leaves a number of difficulties, however, and Russell returned to some of these in longer studies. We find, for example, discussion of whether there is a fundamental division in reality between universals and particulars and, if so, what it is. Russell argues that there is such a division (paper 12) but, in a much later paper (paper 13), he argues that particular things are nothing more than a collection of *compresent* qualities. His eventual position seems to be, therefore, that everything is constituted solely by universals but that, under certain conditions of compresence, universals can come together to be a spatially located particular thing. Though many have rejected this basic picture, the notion of compresence is one which has remained and is used by contemporary philosophers. Jan Van Cleve invokes the notion, for instance, in his more recent defence of the bundle theory of substance (Van Cleve 1985).

Russell is often placed within an empiricist tradition that runs from Locke through Berkeley, Hume and Mill. The empiricist dictum – that all knowledge comes from experience – has been reflected in the metaphysics of the empiricists, even if its authenticity is at times unconvincing. In the papers in Part IV, which concern causation, substances and laws, there is good evidence in favour of Russell's classification among the empiricists. Two main claims are developed by Russell, each typical of empiricist metaphysics. First, Russell argues for a general deflationism about causation: there is no causation nor any modal truth in the world (the only necessities are in language or relations between propositions). Second, Russell attempts to replace an ontology of permanent, or at least enduring, physical objects with one of succeeding and connected events. This creates a problem for Russell, however. In the first selection, "On the notion of cause" (paper 14), Russell had argued strongly against the notion of cause and pushed the view that causal claims should not just be replaced with a more precise relation but, rather, causal claims were wildly misconceived and should be eliminated altogether. There is no other relation that maps on to the metaphysical relation of causation. Like demonic possession and phlogiston, causation just does not exist. But in the later papers in this section, Russell needs some relation that unites certain classes of events so as to constitute what we commonly think of as an object. Here, a causal relation would seem particularly helpful and Russell indeed says in "Physics and metaphysics" (paper 16) that "[objects] are a series of more or less similar phenomena, connected, not by substantial identity, but by certain causal connections". Russell is obliged, therefore, to offer an explanation of what can be meant if there really is no causation. This leads him to a discussion of laws – the laws of nature – which are the only thing that can be sensibly put in place of causes. But Russell's empiricism also dictates a certain view on laws. They tell us only what does happen or has happened, they cannot tell us what

must happen. It seems after all, then, that there really is nothing that binds together the successive stages of an object. There is no necessity from one, earlier stage of an object to a later stage.

The final selection in Part IV (paper 17) shows a remarkable change of view. Causality, it is now argued, is one of the fundamental postulates of science. It is one of the basic assumptions upon which science rests. This contrasts radically with the earlier view that only metaphysics, not science, employs such a notion. However, thirty-five years had passed between the first and final papers in this part, a period over which a change of mind can hardly be attributed to fickleness. There are also discernible continuities in Russell's thinking on the subject. The only sensible account of causation is still essentially an empiricist one, according to Russell.

Part V collects papers on diverse matters of ontology from the later period of Russell's work. His paper "Vagueness" (paper 18) is a classic by any measure but especially so as it seems to have been the first paper to have taken vagueness as a serious philosophical subject. "Physics and neutral monism" (paper 19) develops Russell's response to the metaphysical question in the philosophy of mind. Mental and physical events will be, he says, compresent. In "Language and metaphysics" (paper 20), Russell considered the metaphysical implications of language and the final paper, "The principle of individuation" (paper 21), is a look at the difficult and abstract issue of "what is a 'particular'?".

Collectively, the papers in this book constitute a fine and important body of work that alone would be sufficient to mark Russell as one of the greatest philosophers of the twentieth century. That he made similarly significant contributions to many other areas of philosophy – notably in the theory of logic and epistemology – can leave no doubt as to his standing.

Part I

IDEALISM

We may use metaphysics, like poetry and music, as a means of producing a mood, of giving us a certain view of the universe, a certain attitude towards life.

"Seems Madam? Nay, It Is"

1

GREEK EXERCISES

Extracts 1888

After Russell's parents died, against their wishes he was given a religious Victorian upbringing mainly by his formidable grandmother, Lady Russell. He was educated well but, on certain matters, freedom of speech was limited. This was why he decided, aged fifteen, to write down his thoughts in a coded diary. It was in transliterated Greek characters with "Greek Exercises" written as the title. In it, Russell could question convention and religious orthodoxy without fear of intrusion or concern from his people. He mentions a former tutor in §1, Mr Ewen, who was an acquaintance of Marx's daughter Eleanor (see *Papers* 1: xvi). With Ewen's departure, Russell could only consider these questions in soliloquy so that appears to be the event that immediately prompted Russell's earliest surviving philosophy. Russell refers, in the selection, to a work by Argyll (1866). This is ignored by contemporary metaphysicians.

Free will and determinism was a chief concern of the author of "Greek Exercises" and this led him to consider issues of causation and laws of nature. God's existence was accepted at this stage but only because there was a "scientific" argument for it. If matter and force were created, they could only have been so by divine power. Even if they were not created, they are regulated by laws of nature, which entails a divine controller. It followed from this argument that the God of reason could not be expected to love us or answer our prayers. Further, there would be no immortality and no miracles, for both these and free will would entail constraints on God's omnipotence.

Some of the argument is less than rigorous and would not be considered serious metaphysics by any professional contemporary philosopher. But it is still advanced for an adolescent and sets the scene for a number of Russell's papers that follow. Papers 2, 3 and 14 make an interesting comparison with the juvenile reasoning here.

He manages by the later sections, after reading Argyll, to produce the following "solution" to his concerns about free will. God's omnipotence is said (§20) to be the same thing as the reign

of (natural) law and the determination of actions by motives is the form the reign of law takes in man. He is thus doubtful that there is any room for free will. We are biological machines and there is no sharp dividing line between man, protozoa and trees. Why should we have free will if trees do not?

We can see these early efforts as evidence of a fertile mind undergoing development and self-reflection yet concerned with the basic metaphysical nature of the universe and the place of human beings in it.

1

Eighteen eighty-eight. March 3. I shall write about some subjects especially religious ones which now interest me. I have in consequence of a variety of circumstances come to look into the very foundations of the religion in which I have been brought up. On some points my conclusions have been to confirm my former creed, while on others I have been irresistibly led to such conclusions as would not only shock my people, but have given me much pain. I have arrived at certainty in few things but my opinions, even where not convictions are on some things nearly such. I have not the courage to tell my people that I scarcely believe in immortality. I used to speak freely to Mr. Ewen on such matters, but now I cannot let out my thoughts to any one, and this is the only means I have of letting off steam. I intend to discuss some of my puzzles here.

[...]

3

19th. I mean today to put down my grounds for belief in God. I may say to begin with that I do believe in God and that I should call myself a theist if I had to give my creed a name. Now in finding reasons for belief in God I shall only take account of scientific arguments. This is a vow I have made, which costs me much to keep and to reject all sentiment. To find then scientific grounds for a belief in God we must go back to the beginning of all things. We know that if the present laws of nature have always been in force, the exact quantity of matter and energy now in the universe must always have been in existence; but the nebular hypothesis points to no distant date for the time when the whole universe was filled with undifferentiated nebulous matter. Hence it is quite possible that the matter and force now in existence may have had a creation, which clearly could be only by divine power. But even granting that they have always been in existence, yet whence come the laws which regulate the action of force on matter?

I think they are only attributable to a divine controlling power, which I accordingly call God.

4

March 22. In my last exercise I proved the existence of God by the uniformity of nature, and the persistence of certain laws in all her ways. Now let us look into the reasonableness of this reasoning. Let us suppose that the universe we now see has as some suppose grown by mere chance. Should we then expect every atom to act in any given conditions precisely similarly to another atom? I think if atoms be lifeless, there is no reason to expect them to do anything without a controlling power. If on the other hand they be endowed with free will we are forced to the conclusion that all atoms in the universe have combined in the commonwealth and have made laws which none of them ever break. This is clearly an absurd hypothesis, and therefore we are forced to believe in God. But this way of proving his existence at the same time disproves miracles and other supposed manifestations of divine power. It does not however disprove their possibility for of course the maker of laws can also unmake them. We may arrive in another way at a disbelief in miracles. For if God is maker of the laws, surely it would imply an imperfection in the law if it had to be altered occasionally, and such imperfection we can never impute to the divine nature. (As in the Bible, God repented him of the work.)

5

April 2nd. I now come to the subject which personally interests us poor mortals more perhaps than any other. I mean, the question of immortality. This is the one in which I have been most disappointed and pained by thought. There are two ways of looking at it. First, by evolution, and comparing man to animals, second by comparing man with God. The first is the more scientific, for we know all about the animals, but not about God. Well, I hold that, taking free will first to consider, there is no clear dividing line between man and the protozoon. Therefore if we give free will to man, we must give it also to the protozoon. This is rather hard to do. Therefore unless we are willing to give free will to the protozoon, we must not give it to man. This however is possible, but it is difficult to imagine, if, as seems to me probable, protoplasm only came together in the ordinary course of nature, without any special providence from God, then we and all living things are simply kept going by chemical forces and are nothing more wonderful than a tree (which no one pretends has free will) and even if we had a good enough knowledge of the forces acting on anyone at any time, the motives pro and con, the constitution of his brain at any time,

then we could tell exactly what he will do. Again from the religious point of view, free will is a very arrogant thing for us to claim for of course it is an interruption of God's laws, for by his ordinary laws all our actions would be fixed as the stars. I think we must leave to God the primary establishment of laws which are never broken and determine everybody's doings. And not having free will we cannot have immortality.

6

[...]

Monday April 9. I do wish I believed in the life eternal. For it makes me quite miserable to think man is merely a kind of machine endowed unhappily for himself with consciousness. But no other theory is consistent with the complete omnipotence of God, of which science I think gives ample manifestations. Thus I must either be an atheist or disbelieve in immortality. Finding the first impossible, I adopt the second, and let no one know. I think, however disappointing may be this view of man, it does give us a wonderful idea of God's greatness to think that he can in the beginning create laws which, by acting on a mere mass of nebulous matter perhaps merely ether diffused through this part of the universe, will produce creatures like ourselves, conscious not only of our existence but even able to fathom to a certain extent God's mysteries! All this with no more intervention on his part! Now let us think whether this doctrine of want of free will is so absurd. If we talk about it to any one, they kick their legs or something of that sort. But perhaps they cannot help it, for they have something to prove and therefore that supplies a motive to them to do it. Thus in anything we do we always have motives which determine us. Also there is no line of demarcation between Shakespeare or Herbert Spencer and a Papuan. But between them and a Papuan there seems as much difference as between a Papuan and a monkey.

7

April 14th. Yet there are great difficulties in the way of this doctrine that man has not immortality, nor free will, nor a soul, in short, that he is nothing more than a species of ingenious machine endowed with consciousness. For consciousness in itself is a quality quite distinguishing men from dead matter. And if they have one thing different from dead matter, why not another, free will? (By free will I mean, that they do not for example obey the first law of motion, or at least that the direction in which the energies they contain is employed depends not entirely on external circumstances.) Moreover, it seems impossible to imagine that man, the great man, with his reason, his knowledge of the universe; and his ideas of right and wrong,

man, with his emotions, his love and hate, and his religion, that this man, should be a mere perishable chemical compound, whose character, and his influence for good or for evil, depends solely and entirely on the particular motions of the molecules of his brain, and that all the greatest men have been great by reason of some one molecule hitting up against some other a little oftener than in other men! Does not this seem utterly incredible, and must not any one be mad who believes in such an absurdity? But what is the alternative? That (accepting the evolution theory, which is practically proved), apes having gradually increased in intelligence, God suddenly, by a miracle, endowed one with that wonderful reason which it is a mystery how we possess. Then is man, truly called the most glorious work of God, is man destined to perish utterly, after he has been so many ages in evolving? We cannot say; but I prefer that idea to God's having needed a miracle to produce man, and now leaving him free to do as he likes.

[...]

10

April 25th. I have begun reading Argyll's *Reign of Law*, and have read about half his chapter on the supernatural. I am much interested by it, but I think there are lots of fallacies in it. For example he seems to assume that there are lots of laws of nature, of which God chooses the ones necessary for his purpose, and by suitably choosing them performs an apparent miracle. But I think, very likely, that first, there is only one ultimate law of nature (since fresh discoveries tend to diminish their number, as gravitation reduced Kepler's laws to one, and as I hope before long all the inverse squares will be reduced to one law, and as I hope all elements will reduce to one, ether, their differences, on the vortex theory, being caused by different kinds of vortices), which law of nature is really pretty much the same as God (bearing about the same relation as the Logos in the gospel according to John); secondly, God, I should say, lets his laws act for themselves, and choosing them out in that way would be in itself an act of divine intervention, which I should have called a miracle. Many other things in the book struck me as unsound but perhaps I didn't understand them.

[...]

12

[...]

May 3. I am beginning to understand a possibility of the existence of free will. I am reading over again Argyll's chapter on the supernatural. I am

much interested by his idea of miracles. He says they may be produced not by breaking through the laws of nature, but by the use of laws not commonly brought into play, i.e. of peculiar circumstances. Now may not we apply this argument to man and the animals? May we not say, in the ordinary course of nature, when the constituent parts of protoplasm came together, by a law at present quite beyond us, the compound formed was endowed with a germ of consciousness certainly and possibly of free will? This germ, if it existed in the protozoon, may easily be conceived of as developing itself more and more until it has evolved into the marvellous product of nature which we behold in man. Even man may be only a prelude to something grander and more gifted still, which may even now be in process of evolution. I don't believe in the Duke when he says there is an obvious purpose running through all animals, which received its final attainment in man. For is there not just as much an evolution from the Papuan to a philosopher or a Newton, as from the monkey to the Papuan? May not this process of evolution continue until a being is evolved differing greatly from the man of today? A being perfectly reasonable, without superstition? (able possibly to comprehend the infinite), and in many other, perhaps as yet inconceivable ways, superior to the greatest of modern philosophers? This loophole about free will, (which does not at all convince me of its actual existence, but only of its possibility) does not however affect my views on immortality. For there is another very strong argument which I did not insert in its place, namely, that the soul here below seems so inseparably bound up with the body, growing with it, weakened with it, sleeping with it, and affecting the brain and affected in return by anything abnormal in the brain. Wordsworth's "Intimations" are humbug, for it is obvious how the soul grows with the body, not as he says, perfect from the first.

[...]

20

30th July 1888. In numbers fifth, sixth and seventh I have treated of free will and immortality. Now however I should not say what I then said, that without free will, immortality is impossible. Force and matter are most likely wanting in free will, yet they have immortality. So that argument is done for. In fact, since then, the question of free will has only grown clearer and clearer to me, while immortality remains wrapped in uncertainty. I will here make an exposition of my views about free will ...

There are about three different, though comerging, ways of looking at this question of free will, first, from the omnipotence of God, second, from

the reign of law, and third, from the fact that all our actions, if looked into, show themselves as caused by motives. These three ways we see at once to be really identical, for God's omnipotence is the same thing as the reign of law, and the determination of actions by motives is the particular form which the reign of law takes in man. Let us now examine closely each of these ways.

First, from the omnipotence of God. What do we mean, in the first place, by free will? We mean that where several courses are open to us, we can choose any one. But according to this definition, we are not ruled by God, and alone of created things, we are independent of him. That appears unlikely, but is by no means impossible, since his omnipotence is only an inference. Let us then pass on to the

Second, from the reign of law. Of all the things we know, except perhaps the higher animals, it is obvious, that law is completely the master. That man is also under its dominion, appears from a fact such as Grimm's law, and again from the fact that it is possible sometimes to predict human actions. If man, then, be subject to law, does not this mean, that his actions are predetermined, just as much as the motions of a planet or the growth of a plant? The Duke of Argyll indeed speaks of freedom within the bounds of law, but to me that's an unmeaning phrase, for subjection to law must mean a certain consequence always following in given conditions. No doubt different people in the same circumstances act differently, but that is only owing to difference of character, just as two comets in the same position move differently because of differences in their eccentricities.

The third, from the consideration of motives, is about the strongest. For if we examine any action whatsoever, we find always motives, over which we have no more control than matter over the forces acting on it, which produce our actions. The Duke of Argyll says we can present motives to ourselves, but is not that an action, determined by our character, and other unavoidable things.

The argument for free will from the fact that we feel it, is worthless, for we do not feel motives which we find really exist, nor that mind depends on brain, etc. But I am not prepared dogmatically to deny free will, for I have often found that good arguments don't present themselves on one side of a question till they are told one. My nature may incline me to disbelieve free will, and there may be very excellent arguments for free will which either I have never thought of, or else have not had their full weight with me. All my arguments may be answerable, but my present opinion is that free will is a delusion, arising from the imperceptibility of the bonds that hold us. It is however a hard thought, and one which causes one at first much pain, for it reduces man to the level of a conscious steam engine or

electric battery. It is impossible for us to imagine, although we find we must be, that we are

> … only cunning casts in clay,

and we may say with the poet

> Let science prove we are, and then
> What matters science unto men?

We can scarcely continue to believe that "life is real, life is earnest", and it is difficult not to become reckless and commit suicide, which I believe I should do but for my people.

2

ON THE DISTINCTION BETWEEN THE PSYCHOLOGICAL AND METAPHYSICAL POINTS OF VIEW

*c.*1894

We know that Russell attended James Ward's metaphysics course at Cambridge in the academic year 1893/4. This selection appears to have been written for that course though it is unlike Russell's other graduate essays which have survived. They were mainly commentaries on the great philosophers whereas this essays takes a more problem-based approach to the subject. The manuscript also differs from his other essays in that it has none of Ward's comments marked anywhere. Could Russell, therefore, have been inspired to write this essay of his own initiative? With some moments of clarity, the essay contains many puzzling and unusual claims. In the introduction, I have tried to set out the confusing philosophical background, and understanding of metaphysics, against which this essay was written.

Russell distinguishes the viewpoints of physics, psychology and metaphysics. He offers a critique of idealism. The early idealists, such as Berkeley, had taken it that the psychological viewpoint was a given, that could be known immediately without any inference or presupposition. Russell argues that the positions of physics and psychology are on an equal footing. Physics has "transcended the immediate datum" in inferring physical objects beyond the given experience. The mistake is to think that psychology has made no such inference but that it merely reports the given data in presuppositionless statements such as "I see the appearance of a table". Russell notes that, in such cases, I am going beyond the given when I note *that I have a certain perception*. Physics infers matter while psychology infers an experiencing subject. Metaphysics appears, at this point in the argument, as an attempt (somewhat lame, Russell judges) to "restore the original concrete unity" of objective reference and subjective idea. In short, the acceptance of

> both a physical world and of individual minds are metaphysical commitments. This characterisation of metaphysics, and Russell's assessment, might explain away what looks at first sight to be an inconsistency in the essay. In the last paragraph, Russell says that metaphysics sees the objective reference and subjective nature of an idea as known intuitively and immediately. Given that Russell has said earlier that the physical and psychological viewpoints transcend the given, it can only be assumed that Russell thought metaphysics mistaken in this respect.

In the science of Psychology we are concerned with mental states as such; we analyse them into their simplest constituents, we study the laws of their generation and decay; if possible, we find out their physiological antecedents or concomitants (though this strictly speaking belongs to Psychophysics). We accept every mental state, or psychosis, as a fact, just as in Astronomy we accept the motions of the heavenly bodies. We try to discover causal laws as to the succession of such states, and to discover what, in isolation from the outer world, their true nature and composition may be.

But, besides their existence and their nature, our ideas have what we may call *meaning*. This word is used to denote their objective reference, that is, their reference to something beyond themselves, to something which they are not, but with which they are intimately concerned. And this aspect of our ideas is more interesting to the plain man, is indeed altogether more that which presents itself in every-day life, than the psychological aspect. This may be illustrated by pretty nearly every state of mind, though in the case of pure sensation this distinction between subject and object is perhaps hardly as yet explicit, and the reference to anything beyond itself is very vague. If I lie in a field on a hot day with my eyes shut, and feel sleepily the heat of the sun, the buzz of the flies, the slight tickling of a few blades of grass, it is possible to get into a frame of mind which seems to belong to a much earlier stage of evolution; at such times there is only what Bradley would call "a vague mass of the felt"; I do not reflect on the outside causes of the various blurred and indistinct sensations, nor on the fact that I am feeling these sensations. Perhaps we may hope that there are also possible states of mind where we are above the distinction of subject and object, as in pure sensation we are below it; but to pursue such a possibility would be to plunge into mysticism, from which for the present we shall do well to keep free.

Neglecting, then, such extreme cases, and confining ourselves to more normal states of consciousness, whether cognitive or volitional, we see that it is the objective reference which first attracts our attention. If I see a table, the normal reflection is "There is a table", not "I am in a state of mind in which a table appears to my sight." Indeed in each of these I have transcended the

immediate datum, in the one case by judging that there is a table, in the other by judging that I have a certain perception (which, be it observed, is an entirely different frame of mind from that in which I have the perception, and contains knowledge only obtainable by memory and retrospection). It is extremely important to realize that the psychological reflection is a transcending of the given every whit as much as the physical; indeed even more if anything. It was the failure to notice this which led to Berkeley and Subjective Idealism, and made people suppose their knowledge of their own states of mind was more certain and ultimate than that of the outside world. The experience is originally given as one whole, a subjective idea with an objective reference; thought splits the whole into two parts, relegating the one to Physics, the other to Psychology; Metaphysics endeavours, somewhat lamely, it is to be feared, to undo the work of thought and restore the original concrete unity; or, as has been said, to stitch Cassim together again after the robbers have hewn him asunder – The above instance was from a cognitive element of consciousness: similar remarks apply to the volitional. Suppose I desire to eat an apple: there is here (1) a subjective state of mind, springing from certain causes, hunger, gourmandise or what not; this is what interests the psychologist (2) an objective reference, to the apple namely; this latter is what interests the student of Ethics. The state of mind *is* (to the psychologist) a state of tension accompanied or produced by a certain state of the body, having definite causal antecedents and consequents just like any other event, physical or mental; but it *means* (to the student of Ethics) a felt want, a contrast between the ideal and the actual, the craving for a fuller self-realization, the momentary identification of the apple with my Good; and he has to discuss how far such identification is moral, just as the Physicist has to discuss how far our perception is correct. Here again the metaphysician endeavours to combine the two; for him the absolute chasm between the ideal and the real, which Ethics has produced by abstraction, has somehow to be bridged: the desire and the desired have somehow to be brought into harmony – But to pursue this idea would take us too far from our subject.

We have thus seen how a psychosis has three aspects under which it may be viewed: (1) Its existence (2) Its content (3) Its *meaning*. Psychology deals with the second alone, and in doing so it is of course perfectly justified; like every Science, it has to proceed by abstraction. But – and here lies the peculiarity of this Science – every idea in Psychology, when it comes into the mind of the discoverer or the learner, is itself a fresh idea, with all the three sides to it, whereas we are very apt, unless we take care, to regard it as the old idea which forms our matter. Let us take an instance in which this is particularly glaring, the act of belief. Any particular act of belief exists, and it has meaning; the meaning, indeed, is here evidently of overwhelming

importance; but, in accordance with the principle of Psychology, we abstract from these two sides of belief, leaving it merely as a psychical phenomenon. It follows that we can no longer be concerned with the logical grounds for any belief, since these have reference to its meaning, but only to the psychological causes. But every result we obtain as to these causes is itself a proposition demanding our belief, and if we subject it to a similar criticism to that which produced it, we only get a fresh proposition, and so on *ad infinitum*, so that our procedure is circular. Of course it is none the worse for that: philosophical procedure is necessarily circular; but this affords an instance of the slipperiness of psychological reasoning, for until this point has been made clear, it might seem as if such criticism were giving valuable aid in deciding what we should believe and what not, whereas in reality it of necessity takes its stand within the matter it is criticising, and cannot therefore have any such value (directly at any rate). The same criticism applies (though less obviously) to a psychological treatment of desire, for desire is an ever-present element in every psychosis, and but for it we should not go on reading about Desire in our Psychology books. Of course I do not mean that Psychology is not a valuable study, full of important results; I only mean, that, being from the first one-sided, its results throughout are necessarily one-sided; and that its criticism, like every criticism of fundamentals, is of necessity not thorough, and can have no direct bearing on our views as to the correctness of beliefs or the morality of desires, except on purely psychological subjects.

Metaphysics then endeavours to bridge the chasm between physics and psychology. Seeing that the objective reference of an idea is known as intuitively and immediately as its subjective nature (if not more so), it frankly accepts both; it allows a world other than the individual mind, concerning which we have knowledge and desire; its criticism is not of these themselves, but of a world concerning which they are possible: *given* Self and the world in relation, the problem is to make each term and their relation intelligible; and Self and the World are given, because the only alternative is blank and absolute scepticism.

3

THE FREE-WILL PROBLEM FROM AN IDEALIST STANDPOINT

1895

In contrast to the previous two selections, this paper shows signs of the clarity and maturity that were distinctive marks of Russell's philosophy. Though it is not known why Russell wrote the paper, there is evidence that he was deliberately writing it in a popular rather than technical style (see *Papers* 1: 229). As far as is known, the paper remained unpublished.

In the essay, Russell returns to the problem of free will that had concerned him in "Greek Exercises" and applies to it the then fashionable idealism to offer a Kantian solution. He calls this solution "reconciliationist". This corresponds to what philosophers now call "compatibilism" – the view that free will is compatible with determinism – but it does so in a distinctly idealist way.

The crux of Russell's argument for the reconciliationist conclusion is his analysis of the laws of nature. If psychology has laws like physics, those laws cannot control mental phenomena such as our thoughts, desires and decisions. Russell takes laws to be descriptions of uniformity. They cannot bind or control what will happen because they merely report what does happen. Put another way, a law does not regulate its instances, it is nothing more than its instances ("the law is nothing but a compendious description of these various motions"). There is, thus, no ground for saying that people could not have acted otherwise than they in fact did. There is no external constraint on the universe. It has perfect freedom and perfect determinism: self-determinism. All its laws (regularities) are self-imposed laws. But the universe is the sum of all spirits that are in it, as the idealist would put it, hence all contribute to the self-determined laws so participate in the freedom and self-determinism of the whole. One may still wonder what kind of freedom this is. Russell explains that we are free to the extent that we are in harmony with the whole. The ethical thing to do is to seek such harmony.

The understanding of laws, that is crucial to the argument, is controversial. Russell is evidently employing a Humean metaphysic

in which there are no necessary connections between events (Hume 1777: VII). The cause does not determine the effect; neither compels the other. Rather, we infer patterns of regularities from what we have observed so far. Laws are nothing more than such discernible patterns or, perhaps more accurately, a record, summary or systematisation of what events follow others. After being the dominant view for around two centuries, the Humean metaphysics is now under attack (Ellis 2001). But it has also been defended, most notably by David Lewis (1973: 72–7 and 1986: postscripts B and C). The Humean view has been so philosophically entrenched that it might have been thought by Russell to need no argument. Nevertheless, he does offer the argument that laws must be discovered by experience so must be merely perceivable regularities. But this is not satisfactory. The anti-Humean could claim that laws concern the essential properties of things but that those properties can only be discovered empirically. The fact that laws are discovered does not settle the issue either way.

Supposing that Russell's argument goes through, and laws are nothing more than regularities in the events that occur, are such laws "self-imposed" in anything more than a metaphorical sense? Do we want something more than metaphorically self-imposed laws? Further, is it clear that we have freedom if it is only freedom to be in harmony with the General Will of the universe?

As stated above, compatibilism is generally regarded as being the best attitude to take to the free will and determinism issue so Russell's "reconciliationist" position might still be considered good sense. The other options, strong determinism and libertarianism are thought to have obvious drawbacks that leave compatibilism the winner by default. However, Honderich has recently challenged the whole division of compatibilism and incompatibilism as involving a misconception of the notion of freedom (Honderich 1988).

Among the countless readers of Mr. Balfour's *Foundations of Belief*, there must have been many who, after appreciating to the full his very able criticism of Naturalism, have felt a certain disappointment at the brevity and inconclusiveness of his Chapter on Idealism.[1] There is a cardinal weakness, to begin with, in the choice, as the representative of English Idealists, of the late Mr. T. H. Green, who can scarcely, at this date, be regarded as a satisfactory champion. It is a canon of criticism – at any rate if criticism is to be conclusive – to criticise the *best* of a school. But we cannot deny, however

1 Loc. cit., Book. II, ch. II, pp. 137–55.

we may honour Mr. T. H. Green as a pioneer, that the school which he founded has, since his day, made very material advances in several directions. Mr. Balfour complains of the non-appearance of "original work which shall represent the constructive views of the younger school of thinkers". Mr. Bradley's *Appearance and Reality* might, one would think, have answered very accurately to this description; but Mr. Balfour dismisses this epoch-making work with the remark that "he does not know whether it has yet commanded any large measure of assent from the few who are competent to pronounce a verdict upon its merits". His criticism of Idealism cannot, therefore, be regarded as touching the more modern work of those to whose views this name applies.

As a large part of Mr. Balfour's objections, both to Naturalism and to Idealism, are founded on their treatment of the Free-Will problem, and the supposed insufficiency of that treatment as a basis for Ethics, it seemed worthwhile, perhaps, to state popularly the chief points of the accepted Idealist solution of this problem, and to trace, briefly, the effects of that solution on the foundations of Ethics. Such a statement seemed the more desirable, as Mr. Balfour's only allusion to this view, in a footnote, contains, if I am not mistaken, an error as to the nature of law in the psychical world, which a discussion of this problem is calculated to remove.

The question of Free Will is one in which we must all feel some interest, whether we are metaphysicians or "plain men". Many people have an uneasy sense that Science has destroyed the possibility of Freedom, and yet they cling passionately to their belief in it – so passionately, sometimes, as to dread all discussion of it, for fear their faith might be shaken. Some, indeed, who are of a scientific turn of mind, and anxious to be "up to date", reject it scornfully, as a metaphysical mystery, fit only for the Middle Ages. These men will tell you boldly that all your actions are due to processes in the brain, that life is nothing but protoplasm, and matter is the only ultimate reality. But from these the "plain man" shrinks; he cannot face "that moral impoverishment which", as Mr. Balfour says, "determinism is calculated to produce". He falls back, usually, on the supposed self-evidence of freedom; "why, I *know* I can do whatever I choose", he says, "and it is simply absurd to tell me I couldn't have done otherwise if I had wanted to". "No doubt", you might retort, "you can do whatever you choose, but can you *choose* at haphazard? Your choice always has a motive, which is determined by your whole nature, and your nature came to you partly through inheritance, partly through education. At any rate, you have it now, for better or worse, and such as it is, you cannot alter it". But if you made such a retort, you would not convince him; you would only bring down upon you a vehement abuse of Metaphysics. In this abuse, if in nothing else, the "up to date" scientist would be ready enough to join. He would join the more readily,

if you previously applied the same process of questioning to him. You would easily, by suggesting a few metaphysical abstractions, get him to refuse belief to anything which was not evident to his senses, and from this point his ruin would be easy. "How so?" you would rejoin, "you allow nothing to exist except what is evident to the senses? How about matter which no one sees? How about the world before the germ of life appeared?" At this point he would begin to be puzzled, and if you pushed your advantage home, and urged that such a world, by his own definition, was impossible, and therefore mind must be just as ultimate as matter – if you urged all this upon him, he would join with the plain man in abuse of Metaphysics, and say, with mock humility, that, for his part, he preferred to confine himself to the plain teachings of Science.

Thus the "plain man" and the "up to date man" both abuse Metaphysics, and yet there is no other way out of their difficulties. They cannot hope finally to settle the problem – so important to them both – except by Metaphysics, that "unusually obstinate attempt to think clearly", as Mr. Bradley has called it. To make such an attempt, however imperfectly, in connection with the question of Free Will; to point out that, in the main, each of the men we have been cross-questioning is right in what he asserts, but wrong in what he denies; to supply a positive doctrine, in which each of them shall find all that he really values in his view – this is the object of the present Essay. The solution which will be suggested is no new one in philosophy, but philosophy is so apt, in this country at any rate, to be crowded out by science, and has been, on this question, so hardly used by Mr. Balfour, that it may not be fruitless to render popularly the chief points on which the solution depends.

First of all, what is the meaning of our terms? Freedom and Determination, are both ambiguous, have both a stricter and a wider sense – determination, in the narrower sense, means fatalism, and freedom, the freedom of caprice. This ambiguity is a very fruitful source of controversy, the stricter sense being usually the one attacked, while the wider sense is the one defended. Thus the arguments on *both* sides are apt to be irrefutable, and by sufficient care in avoiding definitions, the discussion can be prolonged *ad infinitum*. The fact is, as we shall see, that in their narrower senses both are false, while in then broader senses both are true – the reconciliation of these broader senses, however, belongs to that domain of which the plain man entertains a vague though vivid horror, the domain of "metaphysical subtleties".

To begin with Determinism: What its opponents usually mean by it is the doctrine – more properly called fatalism – that, struggle as you will, you are fatally and irrevocably compelled, from *without*, to a certain type of actions and of thoughts. Your whole soul may rebel against the acts you

commit, and yet, as in some horrid nightmare, you are driven by that vague power, the force of circumstances, into this, that or the other course.

> If we are doomed to die, we are enow
> To do our country loss; and if to live,
> The fewer men, the greater share of honour.

What need of more men, since heaven has ordained our victory or defeat, and our efforts cannot alter the eternal decrees? Of late, this doctrine has acquired a pseudo-scientific sanction, from such works as Ibsen's *Ghosts*; the sins of the fathers are supposed, without any acquiescence of our own, to lead us, blindly and helplessly, along the downward path. But, whatever may be said by crude interpreters of science, the doctrine of Fatalism, though it has led to some of the most heroic episodes in the world's history, is manifestly false; we cannot reasonably doubt, however much our actions may be determined by previous causes, that those causes lie partly within ourselves. We know that, in different characters, similar circumstances lead to different actions, and this alone shews that our character, however acquired, counts for something in determining what our acts shall be. The only sense, therefore, in which Determinism deserves our consideration, is the sense in which it asserts that all our actions have a cause. Our own character and disposition, however, is always a part of that cause and can never be neglected without error. So much as to the two meanings of Determination.

There remain the two meanings of Freedom. The freedom of caprice – or Liberty of Indifference, as it was called – though formerly the only sort of freedom contemplated, has now been almost universally abandoned. It was supposed, by those who held this view, that a man could, by a sheer act of will, in the presence of any motives, act in any way whatever; no knowledge of his previous character, not even the most perfect, and no knowledge of the motives at present soliciting him, would enable the onlooker, according to these people, to decide with certainty as to the course of conduct which the man will pursue. And this inability to predict does not spring from the insufficiency of the data, or from the complexity of the problem, but simply from the fact that Will is exempted from the domain of Causation. This view was upheld by Kant, and formed the basis of his whole Ethic. The pure or moral Will, in his view, was never determined by external causes, was never, in short, an effect, but always itself a new cause breaking in upon the series of events – it was the beginning of a new chain of causes and effects, but never a link in the middle of the chain. Kant admitted that the Will was often determined by previous causes, but he regarded actions so caused as base, and held that it is always possible to prevent the natural effect on the Will from taking place. He imagined a Transcendental Ego sitting aloft, in eternal calm, above the realm of passion and of sense, and able,

despite all the solicitations of the phenomenal world, to determine, by the light of pure reason, what it should do and what it should avoid. Such an Ego has, however, become impossible now-a-days; we know enough about Psychology to be sure that volitions always obey the law of causation, and that, like all other events, they are at once effects of what precedes, and causes of what follows. Further, even if such capricious acts of will *were* possible, we should not be bound to restrict the notion of Freedom to these; indeed, there is much to be said for denying that such volitions would be free at all. Pure caprice, if it could exist, would involve that the agent did not necessarily pursue his own ends, for his ends, since he desires them, are motives, and if he pursues them, his will is *caused* by them. We may even go further, and say, if there are to be real cases of caprice, that the agent *cannot* always pursue his own ends – but this, surely, is not freedom, but the most intolerable bondage! Not only, then, is this Liberty of Indifference, philosophically impossible, but, even if it could exist, it would not really be Freedom.

But if not this, what *do* we mean by Freedom? The true meaning seems to be, absence of *external* compulsion. A man is free, when the laws he follows are self-imposed, when he is "a law unto himself". Such laws – being, in truth, nothing but the resolution consistently to pursue certain ends – will in no way limit our freedom – the power of making them *is*, indeed, that very freedom which we all desire.

But how, in the face of science and its demands, can we maintain that such a freedom is real? On this point, let us hear Mr. Balfour, who has condescended, for this occasion only, to make himself the mouthpiece of science.

> Mankind, it seems, are on this theory free, but their freedom does not exclude determinism, but only that form of determinism which consists in external constraint. Their actions are upon this view strictly prescribed by their antecedents, but these antecedents are nothing other than the characters of the agents themselves.
>
> Now it may seem at first sight plausible to describe that man as free whose behaviour is due to "himself" alone. But without quarrelling over words, it is, I think, plain that, whether it be proper to call him free or not, he at least lacks freedom in the sense in which freedom is necessary in order to constitute responsibility. It is impossible to say of him that he "ought", and therefore he "can". For at any given moment of his life his next action is by hypothesis strictly determined. This is also true of every previous moment, until we get back to that point in his life's history at which he cannot, in any intelligible sense of the term, be said to have a character at all. Antecedently to this, the causes which have produced him are in no special sense connected with his

> individuality, but form part of the general complex of phenomena which make up the world. It is evident, therefore, that every act which he performs may be traced to pre-natal, and possibly to purely material, antecedents, and that, even if it be true that what he does is the outcome of his character, his character itself is the outcome of causes over which he has not, and cannot by any possibility have, the smallest control.[2]

How are we to answer this argument? This question is, as Mr. Balfour rightly points out, the crux of the Free-Will problem, but I believe, contrary to Mr. Balfour's view, that it is a question which Idealism is able to answer.

The apparent cogency of the above argument depends upon an insufficient criticism of the nature of those causes and laws with which science deals. To make this plain, let us consider some science, such as Astronomy, in which the "Reign of Law" has been thoroughly established, and chance and caprice have long ago disappeared. A cause may have, here, in popular language, two meanings, but only one of these, as the following discussion will shew, can be admitted by philosophy.

(1) Cause may denote a more general law, such as the Law of Gravitation, under which the particular phenomenon can be subsumed: the fall of Newton's apple, we say, was *caused* by gravitation. This sense, however, is very loose. It regards the law as something existing independently of the phenomena, just as a human law exists on the statute-book, whether or not there are any cases to which it applies. Hence we get the notion that natural laws *govern* phenomena, as human laws govern states, but this notion rests, at bottom, on no better foundation than the double meaning of the word Law. The law of Gravitation was *discovered* by *observing* the phenomena. Newton saw the apple fall, and he saw the moon go round the earth; he found that the one formula, of the earth's attraction, would describe both these facts. Finally, he discovered that attraction all round, according to the same formula, would also describe the motions of the planets round the sun. Thus the law is nothing but a compendious description of these various motions. Without the motions, it would be nothing; it is a uniformity analysed out of the phenomena, and has no existence apart from the phenomena which exemplify it. Hence it can no more *control* the phenomena than a census controls the population. It is, therefore, only in a very loose and popular sense, that a law can be regarded as the *cause* of the phenomena to which it applies.

(2) But, again, we say that a phenomenon is *caused* by some phenomenon which immediately preceded it – gunpowder explodes because we put

2 Loc. cit., p. 147.

a match to it, a man dies because he has taken arsenic, a billiard-ball moves because we hit it, and so on. In this sense, the notion of cause has still a function in science. In Physics, cause and effect are phenomena of the same kind, and are quantitatively equal. One billiard-ball is in motion, it hits another, and communicates its motion to that other; the total amount of motion is the same after the impact as it was before, and the former motion is the cause of the latter. In this sequence of cause and effect, there are certain uniformities by which, when we know the cause, we can discover the effect, and when we know the effect we can discover the cause. These uniformities are the laws of science. But these laws only tell us *how* phenomena follow each other – they do not shew us *why* phenomena should choose that particular mode of sequence. Nor do they give us any more reason for regarding the future as determined by the past, than for regarding the past as determined by the future. From the present state of the heavenly bodies, we can discover past eclipses and transits, just as easily as those to come. The fact is, that, in the most perfect sciences, any two phenomena have to each other a perfectly determinate relation. We cannot say that the cause determines the effect, without admitting, at the same time, that the effect determines the cause – had the effect been different, we should have inferred a different cause. Neither, in any way that we can see, compels the other – all we can say is, that, as far as our experience goes, the one can be inferred from the other, by certain rules which, in all hitherto observed sequences, have been found to hold.

If, then, Psychology is to have laws like those of Physics, we must not expect those laws to be such as to *control* psychical phenomena – they will only be laws which the phenomena themselves contain. If people, on certain occasions, act in certain ways, that is no ground for saying they *could* not have acted otherwise – we cannot say this even of the heavenly bodies – all we can say is, that people *do* act in these ways. Again, we saw that it was just as reasonable to regard the future as determining the past as it was to take the opposite view, and when we come to examine the nature of will, we shall see how important this way of regarding things may be.

When we start from the later phenomenon, and regard it as determining the earlier one, we call the one which comes last an *End*, and that which went before a *means*. This way of looking at things does not upset the causal relation, but supplements it – and in dealing with will, at any rate, it is essential to any complete account of the facts. My putting the kettle on the fire *causes* it to boil, but it is also the *means* I adopt to the *end* of making it boil. Now in any act of will, we always have some end in view; hence we may regard the end, which is future, as determining the will, which is the means to its attainment. The will, however, regarded as cause, also determines,

as its effect, the occurrence of the end which it has in view; thus neither will nor the end it has in view outweighs its alternative in the relation, but each, as an event in the time-series, would enable an ideally perfect Psychology to infer the other.

After this, I fear, somewhat tedious series of definitions and distinctions, let us see what consequences they will lead to in the metaphysical discussion of Free Will. We have, here, the inestimable advantage of direct introspection – we may hope, therefore, to get a deeper insight into the nature of laws and sequences than was possible where, as in the physical sciences, we had only outward observation to guide us.

The meaning which we found above for Freedom was, it will be remembered, the absence of *external* compulsion. Where our acts follow no laws save those of our desires – where we are able, always, to pursue what seems to us our highest good – there, surely, it would be out of place to complain of want of freedom. There are cases where such a complaint would be justified – cases where the outer world, breaking in upon our hopes, interposes some impassible obstacle between ourselves and our desires. The death of a friend, the deafness of a Beethoven, the defeat of a cause to which all life's energies had been devoted, these are real limits to liberty, in these the world does definitely declare itself hostile to our cherished ends; but where we are freely able to work out the realization of our desires, there, surely, it would be unreasonable to complain because our desires cannot be other than they are. Such a complaint would be the more unreasonable, as this inability to be different is, at bottom, only the law of identity. If they *are* such and such, it would be a contradiction to suppose them anything else; but if we *choose*, we can make of our desires what we will. Whatever appears to us desirable, we desire; what appears otherwise, we surely should not wish to desire.

But to what extent does such freedom fall to our lot? And what are the conditions of our attaining to the fullest possible development of it? To answer these questions, it will be necessary to consider man's relation to his environment, and his share in the life of the whole.[3]

3 I ought to point out that, in what follows, I have departed somewhat from the traditional meaning of freedom. We may distinguish freedom in volition and freedom in attainment. Whenever our volition has, as to some extent it always has, an end in view which is not wholly dependent on ourselves for its possibility, there, though our volition proper may be free – as, indeed, we have just seen that it must be – yet, as beings with ends to attain, we are *not* wholly free. We cannot, that is, attain to any end we may desire – even our most immediate volitions, which have only bodily motions in view, may be impeded by paralysis or bodily infirmity. The will, then, in the strict sense, is always free, but we are not always free to work out the effects which we hope from our volitions. The latter sort of freedom, which I may call freedom in attainment, is that considered throughout the subsequent parts of this article.

Since freedom consists of the absence of external compulsion, it is clear that the Universe, as the sum of all that is, must be free, for there is nothing outside it by which its freedom could be interfered with. Nevertheless, since the universe is amenable to science, since the world, so far as we know it, is an orderly world, this freedom, which the universe enjoys, is not Liberty of Indifference – on the contrary, we must regard the universe as imposing laws upon itself, and these laws must be regarded as flowing from its own nature. We have thus, in the universe as a whole, a combination of perfect freedom with perfect determination. The determination is self-determination – that is, it comes from within, and may be regarded as an expression of the Universal Will – but it is none the less perfect on that account, for no one, now-a-days, would maintain that there is any part of the universe not amenable to law.

But, further, the Universe is an organism – that is, it is a whole, in which all the parts are interrelated, so that a change in any one part involves a change in every other. Even in physical science, this interdependence may be seen – laws such as Gravitation and the Conservation of Energy, which apply to the whole world, make the motion of every single particle dependent on the motions and attractions of all other particles throughout the material universe – so that we cannot give a full account, even here, of any one portion of the whole, without considering, at the same time, its relation to every other portion. And when we come to the human world, this interdependence is infinitely more varied and more intimate – each of us has a thousand points of contact with the world of his fellow-men, and is dependent on them in a thousand ways in every part of his life. How can such dependence, we may ask, be compatible with Freedom?

The laws of the universe, as we saw, are self-imposed laws. Now the universe is only the sum of all the beings in it, so that every being must have some share in determining what the laws shall be. As a participator in the nature of the whole, every finite spirit must, to some extent, participate in the freedom of the whole; we each and all of us add our votes in the construction of the whole, and every one of us, by his will, makes the world, to some extent, a different place from what it would have been without that will. The whole sum of such wills constitutes the whole activity of the universe. We shall be free, then, when the activity we propose to ourselves is one which the universe as a whole allows, when the law we impose upon ourselves is the same as the law of the whole. A limitation to our freedom will only occur when we set ourselves in opposition to the General Will; in such a case, we shall find our activity impeded from without, and our actions will no longer be completely self-determined. Hence, finally, we are free in proportion as we are self-determined, and we are self-determined in proportion as we are in harmony with the whole. To be in harmony with

the Universal Will is, therefore, the full and sufficient condition of our freedom.

We have now arrived at the end of the discussion of Free Will, which is the more special subject of this Essay. We have seen in what sense man has freedom, how far that freedom extends, and how to make its extent as great as possible. But it may not be out of place, before taking leave of the subject, to add a few words as to the ethical bearing of our solution.

Ever since the time of Kant, the dependence of freedom upon virtue has been recognized in some sense or other, but owing to Kant's faulty psychology of volition, his own account of this dependence is no longer completely adequate. For we saw that liberty, to him, meant Liberty of Indifference; the will, according to his view, was only moral when it had no motive outside itself, when, as he expressed it, the action sprang from pure respect for the moral law as such. The moral will, in such an act, was supposed to have no end in view – it was not determined by any prevision of consequences – the pure love of virtue alone determined it, and virtue lay, not in any prevision of desirable consequences, but in the will alone. Such a pure, unmotived will was possible to a believer in Liberty of Indifference, but to us, with our vastly improved psychology, a will which has no purpose in view has become a chimera. The only acts which are not determined by a prevision of the end to be attained are purely instinctive or reflex acts, and these, unfortunately, are precisely the acts which Ethics and common sense agree in regarding as outside the pale of the moral judgment. The acts which we judge good or bad are deliberative acts – choices between two or more courses which lie open to us, each presenting some advantageous features which the other lacks. In these cases, we estimate, as best we can, the consequences of either course, and judge which of the two sets of consequences we most desire. The motives to will, even in what are called acts of self-sacrifice, are always desires – desires, namely, for the consequences of the volition in question. We can desire the happiness of another just as truly, and in just the same sense, as we desire our own happiness – when we sacrifice our own pleasure to another's, we do so because our desire for the other's pleasure is stronger than that for our own. It may be, indeed, that the conquering desire is not, by itself, so strong as that which it conquers, but secures the victory by seeking allies among other desires which point to the same act. Thus we may sacrifice our own pleasure to another's because we desire to cultivate what is called an unselfish disposition, or because we desire the other's esteem; these desires will reinforce that for the other's happiness, and may make it prevail, even though singly it was weaker than the desire for our own happiness. Thus stated, in fact, the doctrine that the strongest desire always conquers, becomes a truism, for, as in the Survival of the Fittest, the test of superior strength lies only in the conquest.

The only part of the doctrine which is not a truism, is the psychological fact that our actions, except when they are purely reflex, always spring from the desire for their consequences.

In all deliberate action, then, what we seek is the *satisfaction of desire*. If this is so, if all that seems good to us appears as the realization of desired ends, the fundamental principle of Ethics must be to obtain as much of this satisfaction as the world can yield. But here it becomes necessary to guard against a probable error. Satisfaction of desire is not to be confounded with pleasure. As I pointed out above, we may desire another's pleasure just as much as our own; we may desire truth, or beauty, or posthumous fame, or a thousand other things besides our own pleasure. If any one of these objects is attained, the desire for it is satisfied, and what seemed to us good is, so far, realized in the world. Every object of desire, it is true, gives us, when attained, a greater or less degree of pleasure, but this pleasure is not itself the object which we desired. The satisfaction of desire is to be understood as meaning, not the mere feeling of attainment, but the concrete realization, in the actual world, of the object of our desire.

We may, therefore, state our fundamental moral principle thus: In any case of a conflict of desires, that desire is to be followed which will bring with it the largest satisfaction – the satisfaction being great in proportion to the possibility of fully realizing the object of the desire. From this point, it is easy to see the bearing, on morals, of our proposed solution of the Free-Will problem.

A man is free, as we saw, when he is unimpeded in his activity, that is, when his desires are such as can be progressively realized by his actions. But we also saw that his activity is unimpeded in proportion as he is in harmony with the universe as a whole. Hence a man will attain to the satisfaction of his desires in proportion as he is in harmony with the world. Our principle of morals, therefore, reduces itself to the precept of harmony with the Universal Will, and such harmony, as we saw, constitutes freedom. The moral law, therefore, has turned out to be the one and only law "whose service is perfect freedom". Virtue consists in harmony with the will of the Whole, and in harmony with the Whole lies the secret of perfect freedom.

In this conclusion the "plain man", would, I hope, be ready to acquiesce, even if he thought the argument by which it was reached unnecessarily tedious. If the edifying nature of my result does anything to reconcile him to the "metaphysical subtleties" of my method, I may hope to have done something, however slight, towards reconciling him to that "obstinate attempt" in which Philosophy consists.

4

THE DIALECTIC OF THE SCIENCES

1896–8

Russell's most Kantian book was *An Essay on the Foundations of Geometry*. It was, however, to be but a small part of a large and ambitious project requiring a critique of every science. Russell had become a Hegelian and this had suggested to him a project of writing that was to last many years.

Russell was later to explain how he came to Hegelianism and why he rejected it:

> At Cambridge, under the influence of McTaggart, I was taught to hope that the philosophy of Hegel might provide a substitute for traditional dogmas. I listened to McTaggart eagerly, and studied Bradley, with whom I felt inclined to agree. But alas the time came when I could no longer put off reading Hegel himself. The effect was shattering. Much of his two *Logics* I could not understand, but what I could understand, especially the parts dealing with the philosophy of mathematics, seemed to me then, as it still does, to be nothing but a tissue of ignorant nonsense. From that moment I ceased to look for any satisfaction in the Hegelian metaphysics.
>
> (Russell 1946: 69)

For some time, however, the general Hegelian approach captivated Russell. In his autobiography, he recounts an incident which occurred in Berlin just before he became twenty-three years old:

> I remember a cold, bright day in early spring when I walked by myself in the Tiergarten, and made projects of future work. I thought that I would write one series of books on the philosophy of the sciences from pure mathematics to physiology, and another series of books on social questions: I hoped that the two series might ultimately meet in a synthesis at once scientific and practical. My scheme was largely inspired by Hegelian ideas.
>
> (Russell 1967: 125)

This is referred to by Russell scholars as the Tiergarten programme. The programme was to start a series on abstract subjects that gradually became more practical and a series on practical subjects that gradually became more abstract. Eventually the two would meet in a Hegelian synthesis with a metaphysics free of contradiction. The first book in each series was published before Russell ceased being a Hegelian. *German Social Democracy* (1896) can be regarded the first practical book and *An Essay on the Foundations of Geometry* (1897) the first abstract book. Although Russell abandoned the Hegelian masterplan in 1898, his writing career showed that he continued to produce books that would fit into each series.

Hegelian metaphysics is driven by dialectical logic. Thinking and reality, which turn out to be one and the same, develop by a dialectical process in which a contradiction in one thought or category results in it being opposed by its antithesis. This opposite also contains contradictions and the two are merged into a synthesis. The synthesis is not satisfactory for long, however, as it constitutes a new thesis which will soon generate an antithesis because of its own inherent contradictions. The end fulfilment of this historical process will appear in both thought and the world beyond. Hegel's view of what constituted the fulfilment of the process seems quite amusing. In thought, its appearance is in Hegel's own realisation of these truths. In the world, it appears in the Prussian state.

Russell's Hegelian project is set out in the first selection, the note "On the logic of the sciences", which was probably written in late 1896 in his notebook of that time. His aim is to show how each currently existing science is incomplete but could, through a dialectical process to be demonstrated by Russell, be formed into a single consistent system. Each science treats its own subject matter as an object of independent study, attempting "to construct a universe out of none but its own ideas". As the Universe is a unified whole, the incompleteness of each science reveals itself in contradictions. This incompleteness is remedied by a dialectical transition to a broader science. The epistemologist's task is to uncover, by a Kantian method, the a priori component of each science. This reveals its fundamental postulates and exposes the contradictions. We can then provide the necessary supplement to these postulates that will overcome the contradictions and lead to the next science. This process ends in metaphysics, the only self-subsistent and independent knowledge.

This note was Russell's working manifesto. The task he now saw for himself was to produce a detailed study of each science and show the contradictions it generated in virtue of its incompleteness. Geometry was the first science to be attacked and this project he completed with his *Essay On The Foundations Of Geometry*. He argued that there were inevitable contradictions in geometry

because its space was relative, infinitely divisible and unbounded. For instance, as geometrical space is relative, spatial figures can be individuated only relative to others. This being so, no sense can be made of one figure being moved from one part of space to another. But the *axiom of free mobility* says that figures can be so moved without loss of identity. The axiom of free mobility is necessary to make judgements of measurements as these are based on judgements of congruence. We can judge congruence only by superimposing one figure on another but this requires that a figure be moved and, we have seen, this cannot be done. These contradictions can be removed only if, as a first step, we restore the notion of matter as that which is localised and interrelated in space. Matter or substance, unlike geometrical space, is neither relative, infinitely divisible nor unbounded. Thus, movement in space will be possible and so will be judgements of congruence and thereby measurement.

Matter would, therefore, have to be the next "science" exposed to examination. Russell began his work on this as we see in the next two selections. *The Philosophy of Matter* was to be Russell's next book in the abstract series. His plan of 1897 survives (reprinted below) and, as further proof that Russell continued with the Tiergarten programme long after abandoning its Hegelian justification, Russell did eventually publish his book on matter (1927). The "Various notes on mathematical philosophy" were written between 1896 and 1898 and could be seen as his attempts to deal with the detail of this next stage in his dialectic of the sciences. These notes experiment with different ideas of matter in an attempt, first, to solve the contradictions of geometry and, second, set up new contradictions concerning matter. He moves from a purely kinematic conception to a dynamic one. But even under the dynamical conception, matter fails to be "self-existent": it must still be defined relationally. This problem is overcome only by an eventual transition from physics to psychology but Russell was unable to demonstrate this transition before he gave up the dialectical project.

When the Hegelian masterplan was abandoned, it seems likely that Russell threw away many of his manuscripts, regarding the work as worthless.[a] One fairly polished paper from the period which survives, and which can be understood even independently of the Hegelian programme, is "Why do we regard time, but not space, as necessarily a plenum?" Russell's answer can be summed up: "Because we're fools."[b] Russell does not answer whether space and time are plenal or punctual but does argue that they must be alike, both plenal or both punctual. The wider debate is that

a The evidence for this is assessed in *Papers* 2: xxiv.

b Letter to Alys Russell, 1 June 1897, excerpted in *Papers* 2: 91.

between monism and pluralism or monadism. In monism, there is just one substance so space and time are plenal, that is, filled at every point. In monadism or pluralism, space and time are punctual, allowing gaps between times and spaces. The debate between monism and pluralism or monadism thus turns on whether space and time are plenal (adjectival) or punctual (relational). The notion of punctual time is sometimes regarded as absurd, permitting blank spaces of empty time, when the issue of space being plenal or punctual is regarded as an empirical matter, where space could be either without absurdity. One reason for this division is that our experience is in time but not in space. We can have experience *of* space but not *in* space. Another dissimilarity is thought to be that, while I can choose which place to be in, I cannot choose which time to be in. Russell rejects both these grounds.

A NOTE ON THE LOGIC OF THE SCIENCES

Every science works with a certain limited number of fundamental ideas, which number is smaller than that of all fundamental ideas. Now every Science may be regarded as an attempt to construct a universe out of none but its own ideas. What we have to do, therefore, in a logic of the sciences, is to construct, with the appropriate set of ideas, a world containing no contradictions but those which unavoidably result from the incompleteness of these ideas. Within any science, all contradictions not thus unavoidable are logically condemnable; from the standpoint of a general theory of knowledge, the whole science, if taken as a metaphysic, i.e. as independent and self-subsistent knowledge, is condemnable. We have, therefore, first to arrange the postulates of the science so as to leave the minimum of contradictions; then to supply, to these postulates or ideas, such supplement as will abolish the special contradictions of the science in question, and thus pass outside to a new science, which may then be similarly treated.

Thus e.g. Number, the fundamental notion of Arithmetic, involves something numerable. Hence Geometry, since space is the only directly measurable element in sensation. Geometry, again, involves something which can be located, and something which can move – for a position, by definition, cannot move. Hence matter and Physics.

I think, however, that two types of dialectical transition will have to be distinguished: the one, like the transition from number to things numerable, from space to matter, merely supplies to an abstract idea its necessary and substantive complement, while leaving, to the abstract science, full validity on its own level. In this case, there is hardly contradiction, but only incompleteness. The other kind of transition, like that from continua to

discreta, or from matter and force to (?), is dialectical in the true Hegelian sense, and shews that the notion of the science in question is fundamentally self-contradictory, and must be throughout replaced by another, in any metaphysical construction of the real.

THE PHILOSOPHY OF MATTER

The philosophy of matter is distinguished from Philosophy in general by a special hypothesis, which, without suicide, it cannot dispense with. This is the hypothesis that matter is real – that, though perhaps not all reality, it yet forms a self-consistent and self-dependent constituent of reality. Not that the Philosophy of matter is compelled, as the final outcome of its speculation, to substantiate this hypothesis, but that, in abandoning it, speculation passes at once beyond the realm which is marked out by our title. At most we may include the disproof of our special hypothesis, and an outline of the manner in which the facts may be reconciled with such disproof. To rationalize the Universe, as far as this may be possible, within our hypothesis, to admit and demonstrate our necessary failure to rationalize completely within this hypothesis, and to suggest some theory containing whatever truth we have found on the road – this must be the task of a philosophy of matter. The first step in this task must be the definition of matter.

[...]

VARIOUS NOTES ON MATHEMATICAL PHILOSOPHY

I. Note on transition from geometry to dynamics

Matter commonly thought of as defined by one or other of two properties: extension, or force. But if space *purely* relative, as discussion of Geometry suggests, extension cannot be the distinguishing mark of matter, which has to do duty as substance.[1] Therefore only force remains, i.e. atoms are to be

1 This point may be stated thus: Space is perfectly homogeneous, therefore so long as we leave matter out of account, one position is exactly like another, and motion has no meaning. Even if we had one material point, there would be no sense in saying it either moved or stood still. If we had two, there would be no sense in any motion except in the straight line joining them, because their distance would be the only determinable spatial relation. [This proves, by the way, that any force which two particles may exercise on each other, independently of other matter, must be in straight line joining them.] Now if we had a homogeneous mass filling all space, motion would remain unmeaning, for the same reasons. Similarly motion in a perfect fluid, as conceived in hydrodynamics, is nonsense. And if we take our atoms as having finite extension, and homogeneous within that extension, the internal spatial relations of a single atom demand infinite divisibility, and therefore an infinite number of smaller atoms, or else demand, as ultimate material units, point-atoms within the above extended atoms.

regarded as unextended centres of force, not inherently spatial, and localized only by their interactions. Now force can only manifest itself by producing motion: statical conception of equilibrium of forces only deduction from dynamical conception. Hence Geometry involves consideration of matter, and matter must be considered primarily as that which produces motion in other matter. We have, here, a mainly relative conception of matter, which is desirable: a conception, moreover, whose relativity involves contradictions if matter taken as ultimate category. We have first to discuss laws of motion, and then shew that these and their view of matter involve something more, and lead us on to some other science.

Can we identify force with distance? If so, we have two centres of force

$$\overset{A}{\cdot}\xrightarrow{ab} \quad \xleftarrow{ba}\overset{B}{\cdot}$$

with a relation, which is geometrically distance dist. AB or BA, dynamically force $\overrightarrow{ab}$ or $\overrightarrow{ba}$: third law of motion. Law of gravitation in general form: $\overrightarrow{ab} = f(AB)$; since force and distance aspects of same phenomenon. We are now on level of mechanical view of nature, and can use no conceptions but space, and matter or force (which last involves motion and time). Matter, as atom, punctual carrier (*Träger*) of relations which constitute space: but in order that it may appear, in space, as preferable to other points, must have some super-spatial property. This supplied, on mechanical view of nature, by *force*, i.e. by causal relations between atoms. But if atoms constituted by relations (i.e. destitute of adjectives which have not an external reference), and if only intrinsic relation between two atoms be distance (but why should it be?), then force between two atoms, which is an intrinsic relation, must be derivative from distance, i.e. must equal f(distance).

Above keeps to mechanical view of nature. In reality, atoms will have properties derived from temporal continuity, from new complexes of relations allowing organic growth, etc.

Dynamics distinguished from Geometry by *force* or *substance* (which are interchangeable), while geometrical relations statical. But if statical relations of Geometry turn out to involve interactions between material atoms, and atoms can only be conceived as centres of force, then geometrical relations become dynamical, and have inherent tendency to produce changes in themselves. [See <Note III>.]

Above leaves obscure why carriers of relations not perfectly unchanging, why universe not motionless. But carriers must be substance, and substance always conceived as involving action. Why? Presumably to give individuality or self-assertion required for the real. But this seems vague.

† *Would Dynamics be possible if force did not vary as f(distance)?* Some force between particles necessary for mere notion of such a Science. To make force

vary as time introduces absolute time. To make force vary as direction of join introduces relation to other matter, which can, I think, be proved impossible by same argument I used for distance. Therefore *force between two particles intrinsic relation between them*. If force *measurable*, must be function of measurable intrinsic relation, i.e. *distance*. [Discuss magnetism, polarization, etc. at this point.]

Above analytical argument, from conditions which make Dynamics possible. Deductive argument, from implicit principle making above necessary, still in requisition.

[This will have to start from motion as immediate datum in experience, and analyze conditions of possibility of experience of motion.]

Observe. For the dialectical transition from Geometry to Dynamics, that Geometry involves comparison of different parts or figures in space, and that this involves motion; and that motion involves a more than spatial matter, for a position in space, being defined solely *by* its position, cannot move. Hence Geometry impossible without moving matter. This brings us to Kinematics, and hence to Dynamics. For motion involves a moving matter, whose motion is only relative to other matter. The motion must have a cause, and being a reciprocal relation between bits of matter, the interactions of these bits must be the cause. This contains already the laws of motion.

[...]

III. Some definitions of matter

General definition. Matter is that, in the data of the outer sense, which can be regarded, with less contradiction than any other sensational datum, as logical subject, or as *substance*.

1. *Kinematical definition*. Matter is that of which spatial relations are adjectives.

We saw, in Geometry, that the attempt to make space a logical subject breaks down: that those axioms, which alone make a knowledge of space possible, can only be true on condition that space is a mere adjective. It must therefore be an adjective of *something*: and even Geometry, though otherwise indifferent to matter, requires this something, in general, as a condition of its possibility. For Geometry compares different parts of space: therefore its possibility involves the possibility of motion, i.e. of change of position. This does not yet, so far as Geometry is concerned, introduce time, for how the change of place is effected is irrelevant – nor does it introduce any property of matter except that of being susceptible of varying spatial adjectives without loss of identity. But so much is necessary, for motion is necessary, and motion involves something more than space, since positions, being defined solely as such, are immoveable. Space, in short, is immoveable,

and therefore, if Geometry is impossible without motion, we require something which can move in space – Again space, as required for Geometry, is not merely an adjective, but an adjective of relation: therefore the ultimate constituents of this kinematical matter must *contain* no space,[2] but be localized, by their spatial relations, as points. These punctual atoms must, for the axiom of Free Mobility e.g. actually move, i.e. change their spatial relations – but how they move, is here irrelevant. The atoms are only localized by their relations to each other, and these relations alone, in the manifold of their *possible* values, give space. Thus, e.g. if there were only two atoms, space would *be* only the straight line joining them: if three, the plane in which they lie.

II. *Dynamical Definition of Matter*. Matter is not only the moveable, but the mover: two pieces of matter are capable of causally affecting one another in such a way as to change their spatial relations.

2 There seem three kinematical definitions of matter which are prima facie geometrically possible (1) the plenum (2) material points (3) as a compromise, extended atoms with empty space between them.

(1) The *plenum* may be either homogeneous or heterogeneous.

(a) In the first case, it suffers all the incapacity of space for thinghood. Homogeneity is synonymous with complete relativity, and further, there is no sense in motion in a homogeneous plenum, for if two parts are interchanged, there is no change. Some differentiation is necessary for motion, and homogeneous plenum could no more move than space itself, from which it would be indistinguishable.

(b) In the second case, what is to mark its heterogeneity? What is the quantum of which denser portions contain more? The Plenum has nothing from which to form a quantum except *extension*, which makes it necessarily homogeneous. Moreover the search for elements, which in Geometry leads to the absurdity of points, leads here to an equal absurdity, more gross in a matter *defined* as substance, than in space, which we *could* degrade to an adjective.

(2) Is my view.

(3) Extended atoms must, internally, be regarded as on a level with the plenum. Moreover, if they have extension, they have internal relations, and therefore internal complexity, and therefore are not unanalyzable, but consist of fresh atoms. Moreover they must have a centre of mass, a centre of gravity, etc., and the determination of these demands internal elements, since infinitesimal calculus really only possible where we have indivisibles and finite differences. Moreover gravitation is inexplicable on any theory of contact-action, and thus even extended atoms have to be capable of action at a distance, and lose their only advantage, of allowing reduction to contact-action. But again, contact-action, which we think self-evident because we have played billiards, is wholly inexplicable for the atom – the whole of molar impact depends on deformation, but the atom cannot suffer deformation.

For we saw, in the above definition, that matter must actually move, i.e. change its spatial relations to other matter; now such a change is an event, and must, by the Law of Causation, have a cause. Moreover, if we are to be able to construct a Dynamic, i.e. a science of matter in motion, considered apart from other things in the universe, we must be able to find this cause within the conceptions we already have, i.e. within matter and spatial relations.[3] Hence we say the motion of matter is caused by matter: any two pieces of matter have a reciprocal causal relation tending to change their spatial relation, i.e. their distance. This relation is *force*.

Force must be reciprocal (third law), since its effect is change of distance, which is a reciprocal relation: moreover, unless we suppose it capable of a finite effect in an infinitesimal time, which would be absurd, its effect must be to produce a finite change of spatial relation, and therefore a finite velocity, in a finite time, which gives its instantaneous effect as acceleration. (This is equivalent to first law.) Again, in order that a science of force may be possible, the force between two atoms must be a function of their spatial relation, since this alone is measurable. (This necessity may also be deduced from inverse of Law of Concomitant Variations, since spatial relation and force causally connected.) Hence $F = f(\text{distance})$, which is general form of law of Gravitation. Since experience does not immediately confirm this, we invent a new conception, mass, and put $F = mm'f(r)$. (This includes second law of motion.) This assumes mass (equivalent to quantum of matter) constant for same particle at all times and places, which follows from conception of matter as substance – The above gives Gravitation as ultimate law of Dynamics, and astronomical measure of mass as the fundamental one. Hence, for Dynamics, matter consists of interrelated things, whose relations consist of (1) spatial relations (2) causal relations (forces) tending to change the spatial relations, and themselves measured by their effects in changing these, and functionally connected with these. Their measurement, as well as the subsidiary measurement of mass, thus depends on the measurement of space and time, and thus ultimately on that of space.

IV. Dynamics and absolute motion

The only way of defining a position, and hence a motion, is by reference to axes, which axes, to be perceptible, and to be capable of supplying relata for spatial relations, must be material, or rather, must be generated by the

3 That we cannot really construct such a Science independently of higher categories is proved by the antinomy of absolute motion. The cause of what appears as motion of matter must therefore really be something more complex than mere matter or force.

relations of material points. Motion can therefore only be defined by relation to matter. But it is essential to the laws of motion that this matter should have no *dynamical* (i.e. causal) relation to the matter whose motion is considered or, indeed, to *any* matter. If it *has* such a relation, the laws of motion become inapplicable, and our equations become untrue. But the laws of motion lead to Gravitation, and if this be universal, there is *no* matter without any dynamical relation to any given matter. Hence arises an antinomy: For dynamics, it is *geometrically* necessary that our axes should be material, and *dynamically* necessary that they should be *immaterial*.

How solve this antinomy? It is plain it is so fundamental as to render a purely dynamical universe absurd – real things, in short, must have other adjectives than space and force, whose relativity destroys them – For practical purposes, the antinomy does not destroy the utility of Dynamics, for we can always find matter sufficiently unrelated to any matter whose motion we are studying to make our equations practically true. But for theory, we must replace space and force by relations whose relativity does not render them unintelligible. *Perhaps* there may be hope in restoring the preëminence of the *here*, as a source of absolute position; *perhaps* we may replace force by conation, and pass on into psychology.

WHY DO WE REGARD TIME, BUT NOT SPACE, AS NECESSARILY A PLENUM?

It is usually held that events must be absolutely contiguous, that the cause must be temporally in contact with the effect, that changes cannot occur in jerks, like the motions of a clock, and that the conception of blank spaces of empty time, interspersed with (strictly) instantaneous phenomena, is wholly absurd and untenable. As regards space, on the contrary, it is now generally held that the doctrine of the plenum is one to be proved or disproved by empirical science, or, at any rate, that it cannot be established *a priori*. It seems not improbable that an investigation of the grounds for this difference might prove instructive, either as disproving the received position, or as pointing to some fundamental limitation of the analogy between space and time.

Broadly speaking, there are two views – perhaps not wholly incompatible – which, as soon as we have discovered that time and space are not solid things, appear naturally as the only available alternatives. Either space or time, we may say, may be regarded as an adjective, or series of adjectives, of one and the same thing; while either, again, may be regarded as a relation, or series of relations, between different things. It is fairly evident that, regarded exclusively as adjectives, both space and time must be plena, for both, wherever they are, will be adjectives of some thing, which thing, therefore, will be

at that point of space or time. It seems equally evident that, regarded exclusively as relations, space and time must be empty; for the relata cannot themselves, on this view, be spatial or temporal, so that space and time will fall between things and be empty. The above difference in the ordinary views of space and of time shows, therefore, that time is normally regarded as an adjective, while space is regarded as possibly a relation. For this reason, time is regarded as a plenum, while space, it is thought, may or may not be a plenum. I wish to investigate, by way of preliminary to the very difficult metaphysical discussion of these alternatives, some of the reasons which are held to make self-evident this difference between space and time. When these are disposed of, the metaphysical argument may proceed with greater freedom.

The view we have to discuss is roughly as follows. Space, it may be said, is not an adjective of the real, but a relation between spaceless reals: while time, on the contrary, is an adjective of reals which have temporal extension. Thus empty space becomes not only possible, but necessary, while empty time becomes unmeaning. This is a very important view, and one which asserts a fundamental difference between space and time. Let us see on what grounds it rests.

Our own experience, we are told, is of space, but not in it, while it is both of time and in time. We are not forced, therefore, to regard space as an adjective of ourselves: we may take it as involved only in relations with the not-self. The not-self is thus the spatially external, and space is part of our relation to this not-self – so our present view would urge. But however much one might wish to apply a similar argument to time, this cannot be done, for our own experience is in time, and time is therefore not a mere relation to the not-self. Wherever we have no experience, therefore, there is no time, and empty time is impossible. It seems hard, no this view, to account for sleep or trance, but I will pass to more vital arguments.

I am bold enough, in connection with the above contention, to raise a fundamental doubt, namely: Is it obvious that our experience is in time, in any sense in which it is not also in space? I will discuss this question at length.

In what sense is our experience in time? Our thoughts, one may say, take place in different times, and even occupy time in the thinking. Now the first part of this reply is undeniable, but worthless: our thoughts also take place in different places. Now I think in England: six months ago I thought in America. But, I may be told, you cannot think two thoughts in the same instant, whereas you can think two thoughts in the same place. Now if this means that change of position in the time-series occurs necessarily and continuously in one direction, that, I reply, is a mere result of the one-dimensional nature of time, and would be equally true of a one-dimensional space. If it means that I cannot stand still in time, it is a mere confusion; for to stand still is to retain the same position at different times, and

obviously I cannot, at different times, be in the same point of time. There is nothing mystical about this. If it means that there is some peculiar unity about a present thought, it seems either verbal or false. Either my present thought may be *defined* as what is meant by one thought – the statement is then purely verbal. Or my present thought is supposed to possess some unity or simplicity beyond what is possessed by a train of thoughts – in this case the statement seems false. But all these meanings are frivolous, and somewhat superficial. The important point seems to be, that, while I cannot be in two places at once, "like a bird", I can be in the same place at different times. This, if true, is important; but it seems questionable what is meant by being in a place. If my body is meant, the statement is false: for my body, being extended, is in many places. In the body, duration in time and extension in space are on a level. If a single point of the body is meant, we have a perfectly barren statement: the infinite division in space has not been accompanied by an infinite division in time, and hence, naturally, the parallelism of the two has been destroyed. To give a useful meaning, therefore, to the statement that I cannot be in two places at once, it must be assumed that my soul is in a place, and moreover that it is definitely in *one* place, and in a place which it can change. Now this is a doctrine by no means universally received, and having, indeed, somewhat of the air of a paradox. If space, as is usually held, be wholly alien to the soul, it would seem that the soul is everywhere or nowhere, according to the point of view adopted. If this is so, the argument fails; but whether this be so or not, I leave undecided for the present. There is yet another meaning, it may be said. I can choose my position in space, but not in time; I can choose my country, but not my century. This proves, I may be told, that time is more intimately bound up than space with the nature of the soul. But the argument seems false. What I can choose (within limits) is the correlation of my positions in space and in time, not either separately. Being where I am now, I can choose only within certain limits where I shall be an hour hence. I can equally choose, however, within certain limits, at what time I shall be in any given place. The parallelism does not seem, therefore, to be broken by this contention.

Nothing would seem to result, then, as regards the difference between space and time, from the fact that our thoughts occur in different times. Much more important is the contention that our thoughts take time to think. Do we, in judging, actually pass, in time, from subject to predicate? or in syllogizing, from premiss to conclusion? Or when we read a book or watch a play, are our thoughts necessarily a continuous stream, filling all the interval marked by our clocks, or may they, like the ticking of those clocks, be discrete occurrences, and unlike the ticking, timeless occurrences, having position, but not duration, in the time-series? If the latter view is

possible, there seems little left of the supposed difference between space and time. Time, also, becomes a series of relations between different experiences. It is true that the various experiences I call *mine* occur necessarily at different points of time, but not necessarily at different points of space. But to this there are two replies: First, the reply I gave above, that *one* experience is usually defined by its occurring in one time; second, the reply that if space is experienced, I can experience different points of space, as well as different points of time. Yes, but my experience, I shall be told, is not *in* different points of space, but *of* them. This, however, is by no means obvious. There remains the view which I left undiscussed above, the view, namely, that the soul has a position *in* space, and that other positions are experienced as the not-self. There seem thus, both in space and time, to be the same alternatives, neither of which we have as yet discussed. Either space or time, in fact, may be regarded as an adjective of one extended experience – in which case we get the doctrine of the plenum – or as relations between different experiences, in which case we get the doctrine of discrete real points separated by empty space or time, both of which are relations falsely hypostatized. Why this latter doctrine, which is allowed to pass freely in space, should be regarded, when applied to time, as an impossible paradox, I for one am wholly unable to understand.

That our thoughts actually do take time, and are not discrete timeless units, is a doctrine which, like the continuity of matter, cannot be directly proved by experience. That they *appear* as a continuous series, is undoubtedly true; but the appearance of spatial continuity in matter has been no bar to atomic theories, and I fail to see why the temporal appearance should be differently treated. No difficulty arises, to my knowledge, in connection with the perception of time, which does not arise in connection with space also: memory, in fact, can surely render the same services for time as are rendered, for space, by external perception. That the present thought can refer outside itself, is a cardinal necessity for knowledge and volition alike; that it should refer to something separated by empty time seems no more extraordinary than that it should refer to something separated by empty space. I would suggest, therefore, that the doctrine of the plenum is no more self-evident, when applied to time, than when applied to space, and that in either case it is open to us, should good reasons arise, to reject the doctrine. I would suggest, further, that the many difficulties in the way of the reality of time, which, while experience is regarded as temporal, form an almost insuperable obstacle to any satisfactory metaphysic, would, on the relational view of time, be very much less paradoxical, and would demand a considerably smaller degree of transformation of our ordinary experience.

Having now, I hope, cleared the ground of any common-sense preconceptions, let us see whether the metaphysical arguments for and against

these alternative views of space and time allow any distinction, in their application, between the one and the other.

The first point to observe is, that the distinction between a relation and an adjective, if it is to have any real import, demands a monadistic view of the Universe. If the world contains simple and substantially unanalyzable elements, then a distinction can be drawn between their natures, and their references to other monads. Even if all their states refer to other monads, still the manner of such reference will be determined by their intrinsic natures. In such a case, every monad would have an essence *consisting* of adjectives, but *exhibited* only in relations. There thus remains, on a monadistic hypothesis, an irreducible difference between adjectives and relations. On a strictly monistic view, on the contrary, no such distinction can be maintained. Everything is really an adjective of the One, an intrinsic property of the Universe; the Universe is not validly analyzable into simple elements at all, but if we postulate such elements, our error will avenge itself by pointing out that the elements are still analyzable. In this case, whatever appears as a relation between two reals is, properly speaking, *only* an adjective of a whole which embraces both. (It is this *also*, even on a monadistic view, but not this *only*.) These considerations suggest, what I believe to be the truth, that the choice between the relational and the adjectival views of space and time is inseparably bound up with the choice between Monadism and Monism. If so, since these two philosophies have divided the world ever since thought began, it would be idle to hope for a speedy verdict on the spatio-temporal alternative.

But does not monadism result from the relational view of space alone, even if time be left adjectival? Suppose we have simple (but not self-subsistent) reals, each non-spatial, but each connected by spatial relations with all the others. Suppose, further, that each persists through a continuous time, and has therefore, as an adjective, the time of its persistence. We have here an undoubted monadism. But will such a view stand? We must suppose that some metaphysical reason has forced us to assume simple reals, and to deny the possibility of a differentiated whole not analyzable into constituent elements. Applying this reason – whatever it may be – to each of our persistent monads, we arrive at the same need for further elements to explain their diversity. A monad persisting, with changing states, through time, is practically the same problem as a Universe extending through space. If our Universe must consist of non-spatial monads, these monads themselves, surely, must, by the same argument, be analyzable into non-temporal monads. We should thus obtain, as the simple elements of the real, non-spatial and non-temporal monads, localized by their mutual relations at definite points of space and time.

But, the indignant objector will exclaim, an ultimate constituent of Reality must be eternal, and therefore persistent through time. To suppose simple reals, which perish in the very moment of their birth, is surely the height of absurdity and paradox! I must confess that I can see little force in this argument. That a monad must be eternal I of course admit, if by *eternal* is meant *timeless*; but I fail to see any inference from timelessness to persistence through time. Any such inference must depend on first establishing that time is an adjective, and not a relation, which is the very point in dispute. Until such a conclusion is proved, I see no more reason why a monad should persist through all time, than why it should extend through all space.

Let us now see whither we are led by taking the alternative view, the view, namely, that a real persisting and changing in time, is not necessarily analyzable into simpler constituent reals. In this case, we must admit that a whole may be differentiated, without being analyzable, as to its substance, into undifferentiated elements. When once this admission is made, however, the grounds for a spatial monadism vanish. We can, according to the present view, conceive of a spatially differentiated Universe, in which there are no constituents of equal reality, but on the contrary, only parts which a thorough understanding reveals to be not parts, but mere aspects of the one real whole. Thus the ground for spatial monadism is gone – space is left as an adjective of the One, and as therefore necessarily a plenum, in the sense that the One is omnipresent. On either view, therefore, there seems no ground for a distinction between space and time. If space is necessarily relations between non-spatial monads, the same must be true of time; if time is necessarily a plenum, the same argument will prove the same conclusion concerning space. To decide between these alternatives would be to solve the most fundamental question of metaphysics; but to state the alternatives may surely be not unimportant.

5

SEEMS, MADAM? NAY, IT IS

1897

The break with Hegelianism, in 1898, Russell regarded as the most important revolution in his lifetime's thinking (Russell 1959: 11). The conversion to realism was largely inspired by his contact with G. E. Moore. Russell explained:

> I think that although we agreed in our revolt, we had important differences of emphasis. What I think at first chiefly interested Moore was the independence of fact from knowledge and the rejection of the whole Kantian apparatus of a priori intuitions and categories, moulding experience but not the outer world. I agreed enthusiastically with him in this respect, but I was more concerned than he was with certain purely logical matters.
>
> (Russell 1959: 12)

Russell thus plunged himself directly into technical matters of mathematical logic and there was little general metaphysical speculation for some time. He had a new technical programme on which to work and little time to outline the broad basis of realism. This explains the importance of "Seems, madam? Nay, it is".

Russell wrote it in December 1897 to be read to the Cambridge Apostles, the secretary of which was Moore. The paper marks the beginning of Russell's rejection of Hegelianism which was a gradual process stretching well into the next year and after which Russell's philosophy became highly technical. The argument of the paper is that the world of appearance is real, for all practical purposes, whereas the world of the reality behind appearances is a practical irrelevance. Furthermore, the metaphysical reality bears no relation to experience and is an "empty abstraction", existing only in the metaphysics books. A motivation for these claims becomes apparent towards the end of the paper: metaphysics cannot be expected to provide any religious satisfaction. We should pursue it out of intellectual curiosity and be guided by that alone, as opposed to a desire for religious satisfaction.

Immediately following the last quotation above, Russell goes on to explain the main metaphysical commitment that was to drive the next period of his work.

> The most important of these [logical matters], and the one that has dominated all my subsequent philosophy, was what I called "the doctrine of external relations". Monists had maintained that a relation between two terms is always, in reality, composed of properties of the two separate terms and of the whole which they compose, or, in ultimate strictness, only of this last. The view seems to me to make mathematics inexplicable. I came to the conclusion that relatedness does not imply any corresponding complexity in the related terms and is, in general, not equivalent to any property of the whole which they compose. Just after developing this view in my book on *The Philosophy of Leibniz*, I became aware of Peano's work on mathematical logic, which led me to a new technique and a new philosophy of mathematics. Hegel and his disciples had been in the habit of "proving" the impossibility of space and time and matter, and generally everything that an ordinary man would believe in. Having become convinced that the Hegelian arguments against this and that were invalid, I reacted to the opposite extreme and began to believe in the reality of whatever could not be disproved – e.g. points and instants and particles and Platonic universals.
>
> (Russell 1959: 12)

Philosophy, in the days when it was still fat and prosperous, claimed to perform, for its votaries, a variety of the most important services. It offered them comfort in adversity, explanation in intellectual difficulty, and guidance in moral perplexity. No wonder if the Younger Brother, when an instance of its uses was presented to him, exclaimed with the enthusiasm of youth.

> How charming is divine Philosophy!
> Not harsh and crabbed, as dull fools suppose,
> But musical as is Apollo's lute.

But those happy days are past. Philosophy, by the slow victories of its own offspring, has been forced to forego, one by one, its high pretensions. Intellectual difficulties, for the most part, have been acquired by Science – philosophy's anxious claims on the few exceptional questions, which it still endeavours to answer, are regarded by most people as a remnant of the dark ages, and are being transferred, with all speed, to the rigid Science

of Mr. F. W. H. Myers. Moral perplexities – which, until recently, were unhesitatingly assigned by philosophers to their own domain – have been abandoned by McTaggart and Mr. Bradley to the whimsies of statistics and common sense. But the power of giving comfort and consolation is still supposed by McTaggart to belong to philosophy. It is this last possession of which, tonight, I wish to rob the decrepit parent of our modern gods.

It might seem, at first sight, that the question could be settled very briefly. "I know that philosophy can give comfort", McTaggart might say, "because it certainly comforts me". I shall try to prove, however, that those conclusions which give him comfort are conclusions which do not follow from his general position – which, indeed, admittedly do not follow, and are retained, it would seem, only *because* they give him comfort.

As I do not wish to discuss the truth of philosophy, but only its emotional value, I shall assume a metaphysic which rests on the distinction between Appearance and Reality, and regards the latter as timeless and perfect. The principle of any such metaphysic may be put in a nutshell. "God's in his heaven, all's wrong with the world" – that is its last word. But it seems to be supposed that, since he is in his heaven, and always has been there, we may expect him some day to descend to earth – if not to judge the quick and the dead, at least to reward the faith of the philosophers. His long resignation, however, to a purely heavenly existence, would seem to suggest, as regards the affairs of earth, a stoicism on which it would be rash to found our hopes.

But to speak seriously. The emotional value of a doctrine, as a comfort in adversity, appears to depend upon its prediction of the future. The future, emotionally speaking, is more important than the past, or even than the present. "All's well that ends well" is the dictum of unanimous common sense. "Many a dull morning turns out a fine day" is optimism; whereas pessimism says:

> Full many a glorious morning have I seen
> Flatter the mountain tops with sovereign eye
> Kissing with golden face the meadows green
> Gilding pale streams with heavenly alchemy;
> Anon permit the basest clouds to ride
> With ugly rack on his celestial face
> And from the forlorn world his visage hide
> Stealing unseen to west with this disgrace.

And so, emotionally, our view of the universe as good or bad depends on the future, on what it will be; we are concerned always with appearances in

time, and unless we are assured that the future is to be better than the present, it is hard to see where we are to find consolation.

So much, indeed, is the future bound up with optimism, that McTaggart himself, while all his optimism depends upon the denial of time, is compelled to represent the Absolute as a future state of things, as "a harmony which must some day become explicit". It would be unkind to urge this contradiction, as it is mainly McTaggart himself who has made me aware of it. But what I do wish to urge is, that any comfort, which may be derived from the doctrine that Reality is timeless and eternally good, is derived only and exclusively by means of this contradiction. A timeless Reality can have no more intimate connection with the future than with the past: if its perfection has not appeared hitherto, there is no reason to suppose it ever will – there is, indeed, every likelihood that God will stay in his heaven. We might, with equal propriety, speak of a harmony which most once *have been* explicit; it may be that "my grief lies onward and my joy behind" – and it is obvious how little comfort this would afford us.

All our experience is bound up with time, nor is it possible to imagine a timeless experience. But even if it were possible, we could not, without contradiction, suppose that we ever *shall* have such an experience. All experience, therefore, for aught that philosophy can show, is likely to resemble the experience we know – if this seems bad to us, no doctrine of a Reality distinguished from Appearances can give us hope of anything better. We fall, indeed, into a hopeless dualism. On the one side we have the world we know, with its events, pleasant and unpleasant, its deaths and failures and disasters – on the other hand an imaginary world, which we christen the world of Reality, atoning, by the largeness of the R, for the absence of every other sign that there really is such a world. Now our only ground for this world of Reality is, that this is what Reality would have to be if we could understand it. But if the result of our purely ideal construction turns out so very different from the world we know – from the real world, in fact – if, moreover, it follows from this very construction that we never shall experience the so-called world of Reality, except in a sense in which already we experience nothing else – then I cannot see what, as concerns comfort for present ills, we have gained by all our metaphysicizing. Take, for example, such a question as immortality. People have desired immortality either as a redress for the injustices of this world, or, which is the more respectable motive, as affording a possibility of meeting again after death those whom they have loved. The latter desire is one which we all feel, and for whose satisfaction, if philosophy could satisfy it, we should be immeasurably grateful. But philosophy, at best, can only assure us that the soul is a timeless reality. At what points of time, if any, it may happen to appear, is thus wholly irrelevant to it, and there is no legitimate

inference from such a doctrine to existence after death. Keats may still regret

> That I shall never look upon thee more,
> Never have relish in the fairy power
> Of unreflecting love

and it cannot much console him to be told that "fair creature of an hour" is not a metaphysically accurate phrase. It is still true that "Time will come and take my love away", and that "This thought is as a death which cannot choose But weep to have that which it fears to lose". And so with every part of the doctrine of a timelessly perfect Reality. Whatever now seems evil – and it is the lamentable prerogative of evil that to seem so is to be so – whatever evil now appears may remain, for aught we know, throughout all time, to torment our latest descendants. And in such a doctrine there is, to my mind, no vestige of comfort or consolation.

It is true that Christianity, and all previous optimisms, have represented the world as eternally ruled by a beneficent Providence, and thus metaphysically good. But this has been, at bottom, only a device by which to prove the future excellence of the world – to prove, for example, that good men would be happy after death. It has always been this deduction – illegitimately made of course – which has given comfort. "He's a good fellow, and *'twill* all be well."

It may be said, indeed, that there is comfort in the mere abstract doctrine that Reality is good. I do not myself accept the proof of this doctrine, but even if true, I cannot see why it should be comforting. For the essence of my contention is, that Reality, as constructed by metaphysics, bears no sort of relation to the world of experience. It is an empty abstraction, from which no single inference can be validly made as to the world of appearance, in which world, nevertheless, all our interests lie. Even the pure intellectual interest, from which metaphysics springs, is an interest in explaining the world of appearance. But instead of really explaining this actual palpable sensible world, metaphysics constructs another fundamentally different world, so different, so unconnected with actual experience, that the world of daily life remains wholly unaffected by it, and goes on its way just as if there were no world of Reality at all. If even one were allowed to regard the world of Reality as an "other world", as a heavenly city existing somewhere in the skies, there might no doubt be comfort in the thought that others have a perfect experience which we lack. But to be told that our experience, as we know it, is that perfect experience, must leave us cold, since it cannot prove our experience to be better than it is. On the other hand, to say that our actual experience is not that perfect experience constructed by philosophy, is to cut off the only sort of existence which philosophical reality can

have – since God in his heaven cannot be maintained as a separate person. Either, then, our existing experience is perfect – which is an empty phrase, leaving it no better than before – or there is no perfect experience, and our world of Reality, being experienced by no one, exists only in the metaphysics books. In either case, it seems to me, we cannot find in philosophy the consolations of religion.

There are, of course, several senses in which it would be absurd to deny that philosophy may give us comfort. We may find philosophizing a pleasant way of passing our mornings – in this sense, the comfort derived may even, in extreme cases, be comparable to that of drinking as a way of passing our evenings. We may, again, take philosophy aesthetically, as probably most of us take Spinoza. We may use metaphysics, like poetry and music, as a means of producing a mood, of giving us a certain view of the universe, a certain attitude towards life – the resulting state of mind being valued on account of, and in proportion to, the degree of poetic emotion aroused, not in proportion to the truth of the beliefs entertained. Our satisfaction, indeed, seems to be, in these moods, the exact opposite of the metaphysician's professions. It is the satisfaction of forgetting the real world and its evils, and persuading ourselves, for the moment, of the reality of a world we have ourselves created. This seems to be one of the grounds on which Bradley justifies metaphysics. "When poetry, art and religion", he says, "have ceased wholly to interest, or when they show no longer any tendency to struggle with ultimate problems and come to an understanding with them; when the sense of mystery and enchantment no longer draws the mind to wander aimlessly and love it knows not what; when, in short, twilight has no charm – then metaphysics will be worthless." What metaphysics does for us in this way is essentially what, say, *The Tempest* does for us – but its value on this view, is quite independent of its truth. It is not because Prospero's magic makes us acquainted with the world of spirits that we value the *Tempest*; it is not, aesthetically, because we are informed of a world of spirit that we value metaphysics. And this brings out the essential difference between the aesthetic satisfaction, which I allow, and the religious comfort, which I deny to philosophy. For aesthetic satisfaction, intellectual conviction is unnecessary, and we may therefore choose, when we seek it, the metaphysic which gives us the most of it. For religious comfort, on the other hand, belief is essential, and I am contending that we do not get religious comfort from the metaphysic which we believe.

It is possible, however, to introduce a refinement into the argument, by adopting a more or less mystical theory of the aesthetic emotion. It may be contended that, although we can never wholly experience Reality as it really is, yet some experiences approach it more nearly than others, and such experiences, it may be said, are given by art and philosophy. And under the

influence of the experiences which art and philosophy sometimes give us, it seems easy to adopt this view. For those who have the metaphysical passion, there is probably no emotion so rich and beautiful, so wholly desirable, as that mystic sense, which philosophy sometimes gives, of a world transformed by the beatific vision. As Bradley again says: "Some in one way, some in another, we seem to touch and have communion with what is beyond the visible world. In various manners we find something higher, which both supports and humbles, both chastens and supports us. And, with certain persons, the intellectual effort to understand the Universe is a principal way of thus experiencing the Deity. ... And this appears", he continues, "to be another reason for some persons pursuing the study of ultimate truth".

But is it not equally a reason for hoping that these persons will not find ultimate truth? If indeed ultimate truth bear any resemblance to the doctrines set forth in *Appearance and Reality*. I do not deny the value of the emotion, but I do deny that, strictly speaking, it is in any peculiar sense a beatific vision, or an experience of the Deity. In one sense, of course, all experience is experience of the Deity, but in another, since all experience equally is in time, and the Deity is timeless, no experience is experience of the Deity – "as such" pedantry would bid me add. The gulf fixed between Appearance and Reality is so profound, that we have no grounds, so far as I can see, for regarding some experiences as nearer than others to the perfect experience of Reality. The value of the experiences in question must, therefore, be based wholly on their emotional quality, and not, as Bradley would seem to suggest, on any superior degree of truth which may attach to them. But if so, they are at best the consolations of philosophizing, not of philosophy. They constitute a reason for the pursuit of ultimate truth, since they are flowers to be gathered by the way; but they do not constitute a reward for its attainment, since, by all that appears, the flowers grow only at the beginning of the road, and disappear long before we have reached our journey's end.

The view which I have advocated is, no doubt, not an inspiriting one, nor yet one which, if generally accepted, would be likely to promote the study of philosophy. I might justify my paper, if I wished to do so, on the maxim that, "where all is rotten, it is a man's work to cry stinking fish". But I prefer to suggest that metaphysics, when it seeks to supply the place of religion, has really mistaken its function. That it can supply this place, I admit; but it supplies it, I maintain, at the expense of being bad metaphysics. Why not admit that metaphysics, like science, is justified by intellectual curiosity, and ought to be guided by intellectual curiosity alone? The desire to find comfort in metaphysics has, we must all admit, produced a great deal of fallacious reasoning and intellectual dishonesty.

From this, at any rate, the abandonment of religion would deliver us. And since intellectual curiosity exists in some people, it is probable that some attempts would still be made to understand the world, and it is possible that they would be freed from certain hitherto persistent fallacies. "The man", to quote Bradley once more, "whose nature is such that by one path alone his chief desire will reach consummation, will try to find it on that path, whatever it may be, and whatever the world thinks of it; and if he does not, he is contemptible".

Part II

ONTOLOGY (I)

The fundamental doctrine in the realistic position, as I understand it, is the doctrine that relations are "external".

"The Basis of Realism"

6

THE PRINCIPLES OF MATHEMATICS

Extracts 1903

In the first of the selections from *The Principles of Mathematics*, Chapter 4, Russell outlines some of the basic ontological commitments behind his early philosophy of mathematical logic. As he says at the end, the subject might seem remote from the principles of mathematics but is necessary for his treatment of some of the problems to follow. That Russell saw these questions of ontology as basic and fundamental to his project can be seen from their placing in Part I of the work, on the *indefinables* in mathematics.

The ontology advocated is one guided by the logic of language. This is not a view advocated by Russell in some of his later work. One of his final philosophical campaigns was a vehement attack on ordinary language philosophy which claimed, disingenuously, to read philosophical truths from such ordinary language nonsense as "I ain't never done no harm to no one" (Russell 1957: 633). However, his work leading up to the "big book of 1903" was written more under Frege's influence where logic and ontology were seen as closely connected. Russell first read Frege seriously in 1901 and it had a profound influence on his approach (see his account in Russell 1955: 107). However, even here there are reservations about language *dictating* ontology. Language is seen only as *prima facie* evidence, where grammar "though not our master, will yet be taken as our guide".

What Russell finds from the guide of grammar is that a distinction is made between substantive and adjective that is more or less equivalent to that between substance and attribute. Every object of thought is a *term* and may also be called a unit, individual, entity, whatever may be counted as *a one*. Some terms can be used both as predicates in assertion, as in "Socrates is *human*", and as subjects of assertion, as in "*humanity* belongs to Socrates". Terms capable of this twofold use are *concepts*. Terms that cannot be used as predicates are *things*. This may seem to be a conflation of language and ontology in that such entities, which are clearly

meant to be individuals in the world, are also called subjects or substantives, while properties are also called concepts. However, when contrasting this division with Bradley's, Russell makes it clear that he allows no psychological element to his propositions or terms. The term *a man* is a concrete particular, in §51, a certain featherless biped that exists whether or not we use a word to name him or a proposition to talk about him. In Fregean style, terms and propositions have an almost Platonic existence, independent of any mind that apprehends them. In this, Russell is asserting a major claim of the realist metaphysic. He is rejecting the whole tradition of British Hegelianism in favour of the new ontology he had found employed by Frege.

§55 contains a key component of Russell's realism: the doctrine of external relations (though it is not yet given that name). Where *A* differs numerically from *B*, and *C* differs numerically from *D*, it is the same difference in both cases. In the idealist tradition, internal relations are employed, where the difference between *A* and *B* is "unique and unanalyzable and not shared by any other pair of terms". Hence, a relation between *C* and *D* can never be the same relation as one which holds between *A* and *B*. This issue will be discussed in more depth in "The basis of realism", later in this part.

A further theme to be found in this selection is the pluralism, avowed in two places. In §47 he admits many terms which, he says, destroys monism. In §50 he admits conceptual diversity, which implies numerical diversity where numerical diversity is the source of plurality. In taking time to make these points, which now seem obvious, Russell is showing that the monism of idealism was still a position requiring rebuttal. Acceptance of pluralism seemed to be one of the main motivations for Russell giving up idealism. If he were to convert idealists to his cause, no opportunity should be passed up for a disproof of monism.

The second selection from *The Principles of Mathematics* is an early attempt to address the ideas of Alexius Meinong. This is a philosopher that Russell's readers know chiefly from the discussion in "On denoting" (Russell 1905) but those who know Meinong only through that work will think he was "something of a madman" (Lackey, in Russell 1973: 17). Other works by Russell show that he had taken Meinong far more seriously than is conveyed in "On denoting", however, and in these selections we see that respect for Meinong displayed (see also Russell 1904). In §47, for instance, Russell allows that anything that may be an object of thought – a man, a number, or a chimera – is a term. But a term, we have seen, is to be thought of as a mind-independent, really existing individual and not just a word. This view is, thus, a good representation of Meinong's view and it was not until Russell discovered his theory of descriptions that he was able to reject it.

In §427, Russell goes into these matters again. He affirms a division between *Being* and *Existence* and allows that every conceivable term has the former but not the latter, otherwise we could form no coherent proposition about non-existent objects such as Homeric gods and golden mountains. "*A* is not" implies, we are startled to read, that there is a term *A* whose being is denied, and hence that *A is*. We shall see in the next selection that this confusing Meinongian solution is rejected but it is clear that, as late as 1903, Russell saw it as the best available response to the problem.

PROPER NAMES, ADJECTIVES, AND VERBS

46. In the present chapter, certain questions are to be discussed belonging to what may be called philosophical grammar. The study of grammar, in my opinion, is capable of throwing far more light on philosophical questions than is commonly supposed by philosophers. Although a grammatical distinction cannot be uncritically assumed to correspond to a genuine philosophical difference, yet the one is *prima facie* evidence of the other, and may often be most usefully employed as a source of discovery. Moreover, it must be admitted, I think, that every word occurring in a sentence must have *some* meaning: a perfectly meaningless sound could not be employed in the more or less fixed way in which language employs words. The correctness of our philosophical analysis of a proposition may therefore be usefully checked by the exercise of assigning the meaning of each word in the sentence expressing the proposition. On the whole, grammar seems to me to bring us much nearer to a correct logic than the current opinions of philosophers; and in what follows, grammar, though not our master, will yet be taken as our guide.[1]

Of the parts of speech, three are specially important: substantives, adjectives, and verbs. Among substantives, some are derived from adjectives or verbs, as humanity from human, or sequence from *follows*. (I am not speaking of an etymological derivation, but of a logical one.) Others, such as proper names, or space, time, and matter, are not derivative, but appear primarily as substantives. What we wish to obtain is a classification, not of words, but of ideas; I shall therefore call adjectives or predicates all notions which are capable of being such, even in a form in which grammar would call them substantives. The fact is, as we shall see, that *human* and *humanity*

1 The excellence of grammar as a guide is proportional to the paucity of inflexion, i.e. to the degree of analysis effected by the language considered.

denote precisely the same concept, these words being employed respectively according to the kind of relation in which this concept stands to the other constituents of a proposition in which it occurs. The distinction which we require is not identical with the grammatical distinction between substantive and adjective, since one single concept may, according to circumstances, be either substantive or adjective: it is the distinction between proper and general names that we require, or rather between the objects indicated by such names. In every proposition, as we saw in Chapter III [not included], we may make an analysis into something asserted and something about which the assertion is made. A proper name, when it occurs in a proposition, is always, at least according to one of the possible ways of analysis (where there are several), the subject that the proposition or some subordinate constituent proposition is about, and not what is said about the subject. Adjectives and verbs, on the other hand, are capable of occurring in propositions in which they cannot be regarded as subject, but only as parts of the assertion. Adjectives are distinguished by capacity for *denoting* – a term which I intend to use in a technical sense to be discussed in Chapter V [not included]. Verbs are distinguished by a special kind of connection, exceedingly hard to define, with truth and falsehood, in virtue of which they distinguish an asserted proposition from an unasserted one, e.g. "Caesar died" from "the death of Caesar". These distinctions must now be amplified, and I shall begin with the distinction between general and proper names.

47. Philosophy is familiar with a certain set of distinctions, all more or less equivalent: I mean, the distinctions of subject and predicate, substance and attribute, substantive and adjective, *this* and *what*.[2] I wish now to point out briefly what appears to me to be the truth concerning these cognate distinctions. The subject is important, since the issues between monism and monadism, between idealism and empiricism, and between those who maintain and those who deny that all truth is concerned with what exists, all depend, in whole or in part, upon the theory we adopt in regard to the present question. But the subject is treated here only because it is essential to any doctrine of number or of the nature of the variable. Its bearings on general philosophy, important as they are, will be left wholly out of account.

Whatever may be an object of thought, or may occur in any true or false proposition, or can be counted as *one*, I call a *term*. This, then, is the widest word in the philosophical vocabulary. I shall use as synonymous with it the words unit, individual, and entity. The first two emphasize the fact that every term is *one*, while the third is derived from the fact that every term has being, i.e. *is* in some sense. A man, a moment, a number, a class, a relation,

2 This last pair of terms is due to Mr. Bradley.

a chimaera, or anything else that can be mentioned, is sure to be a term; and to deny that such and such a thing is a term must always be false.

It might perhaps be thought that a word of such extreme generality could not be of any great use. Such a view, however, owing to certain widespread philosophical doctrines, would be erroneous. A term is, in fact, possessed of all the properties commonly assigned to substances or substantives. Every term, to begin with, is a logical subject: it is, for example, the subject of the proposition that itself is one. Again every term is immutable and indestructible. What a term is, it is, and no change can be conceived in it which would not destroy its identity and make it another term.[3] Another mark which belongs to terms is numerical identity with themselves and numerical diversity from all other terms.[4] Numerical identity and diversity are the source of unity and plurality; and thus the admission of many terms destroys monism. And it seems undeniable that every constituent of every proposition can be counted as one, and that no proposition contains less than two constituents. *Term* is, therefore, a useful word, since it marks dissent from various philosophies, as well as because, in many statements, we wish to speak of *any* term or *some* term.

48. Among terms, it is possible to distinguish two kinds, which I shall call respectively *things* and *concepts*. The former are the terms indicated by proper names, the latter those indicated by all other words. Here proper names are to be understood in a somewhat wider sense than is usual, and things also are to be understood as embracing all particular points and instants, and many other entities not commonly called things. Among concepts, again, two kinds at least must be distinguished, namely those indicated by adjectives and those indicated by verbs. The former kind will often be called predicates or class-concepts; the latter are always or almost always relations. (In intransitive verbs, the notion expressed by the verb is complex, and usually asserts a definite relation to an indefinite relatum, as in "Smith breathes".)

In a large class of propositions, we agreed, it is possible, in one or more ways, to distinguish a subject and an assertion about the subject. The assertion must always contain a verb, but except in this respect, assertions appear to have no universal properties. In a relational proposition, say "*A* is greater than *B*", we may regard *A* as the subject, and "is greater than *B*" as the assertion, or *B* as the subject and "*A* is greater than" as the assertion. There

3 The notion of a term here set forth is a modification of Mr G. E. Moore's notion of a *concept* in his article "On the Nature of Judgment", *Mind, NS* No. 30, from which notion, however, it differs in some important respects.

4 On identity, see Mr G. E. Moore's article in the *Proceedings of the Aristotelian Society*, 1900–1.

are thus, in the case proposed, two ways of analyzing the proposition into subject and assertion. Where a relation has more than two terms, as in "*A* is here now",[5] there will be more than two ways of making the analysis. But in some propositions, there is only a single way: these are the subject-predicate propositions, such as "Socrates is human". The proposition "humanity belongs to Socrates", which is equivalent to "Socrates is human", is an assertion about humanity; but it is a distinct proposition. In "Socrates is human", the notion expressed by *human* occurs in a different way from that in which it occurs when it is called *humanity*, the difference being that in the latter case, but not in the former, the proposition is *about* this notion. This indicates that humanity is a concept, not a thing. I shall speak of the *terms* of a proposition as those terms, however numerous, which occur in a proposition and may be regarded as subjects about which the proposition is. It is a characteristic of the terms of a proposition that any one of them may be replaced by any other entity without our ceasing to have a proposition. Thus we shall say that "Socrates is human" is a proposition having only one term; of the remaining components of the proposition, one is the verb, the other is a *predicate*. With the sense which *is* has in this proposition, we no longer have a proposition at all if we replace *human* by something other than a predicate. Predicates, then, are concepts, other than verbs, which occur in propositions having only one term or subject. Socrates is a thing, because Socrates can never occur otherwise than as term in a proposition: Socrates is not capable of that curious twofold use which is involved in *human* and *humanity*. Points, instants, bits of matter, particular states of mind, and particular existents generally, are things in the above sense, and so are many terms which do not exist, for example, the points in a non-Euclidean space and the pseudo-existents of a novel. All classes, it would seem, as numbers, men, spaces, etc., when taken as single terms, are things; but this is a point for Chapter VI [not included].

Predicates are distinguished from other terms by a number of very interesting properties, chief among which is their connection with what I shall call *denoting*. One predicate always gives rise to a host of cognate notions: thus in addition to *human* and *humanity*, which only differ grammatically, we have *man, a man, some man, any man, every man, all men*,[6] all of which appear to be genuinely distinct one from another. The study of these

5 This proposition means "*A* is in this place at this time". It will be shown in Part VII [not included] that the relation expressed is not reducible to a two-term relation.

6 I use *all men* as collective, i.e. as nearly synonymous with *the human race*, but differing therefrom by being many and not one. I shall always use *all* collectively, confining myself to *every* for the distributive sense. Thus I shall say "every man is mortal", not "all men are mortal".

various notions is absolutely vital to any philosophy of mathematics; and it is on account of them that the theory of predicates is important.

49. It might be thought that a distinction ought to be made between a concept as such and a concept used as a term, between, e.g. such pairs as *is* and *being, human* and *humanity, one* in such a proposition as "this is one" and 1 in "1 is a number". But inextricable difficulties will envelop us if we allow such a view. There is, of course, a grammatical difference, and this corresponds to a difference as regards relations. In the first case, the concept in question is used as a concept, that is, it is actually predicated of a term or asserted to relate two or more terms; while in the second case, the concept is itself said to have a predicate or a relation. There is, therefore, no difficulty in accounting for the grammatical difference. But what I wish to urge is, that the difference lies solely in external relations, and not in the intrinsic nature of the terms. For suppose that *one* as adjective differed from 1 as term. In this statement, *one* as adjective has been made into a term; hence either it has become 1, in which case the supposition is self-contradictory; or there is some other difference between *one* and 1 in addition to the fact that the first denotes a concept not a term while the second denotes a concept which is a term. But in this latter hypothesis, there must be propositions concerning *one* as term, and we shall still have to maintain propositions concerning *one* as adjective as opposed to *one* as term; yet all such propositions must be false, since a proposition about *one* as adjective makes *one* the subject, and is therefore really about *one* as term. In short, if there were any adjectives which could not be made into substantives without change of meaning, all propositions concerning such adjectives (since they would necessarily turn them into substantives) would be false, and so would the proposition that all such propositions are false, since this itself turns the adjectives into substantives. But this state of things is self-contradictory.

The above argument proves that we were right in saying that terms embrace everything that can occur in a proposition, with the possible exception of complexes of terms of the kind denoted by *any* and cognate words. For if A occurs in a proposition, then, in this statement, A is the subject; and we have just seen that, if A is ever not the subject, it is exactly and numerically the same A which is not subject in one proposition and is subject in another. Thus the theory that there are adjectives or attributes or ideal things, or whatever they may be called, which are in some way less substantial, less self-subsistent, less self-identical, than true substantives, appears to be wholly erroneous, and to be easily reduced to a contradiction. Terms which are concepts differ from those which are not, not in respect of self-subsistence, but in virtue of the fact that, in certain true or false propositions, they occur in a manner which is different in an indefinable way from the manner in which subjects or terms of relations occur.

50. Two concepts have, in addition to the numerical diversity which belongs to them as terms, another special kind of diversity which may be called conceptual. This may be characterised by the fact that two propositions in which the concepts occur otherwise than as terms, even if, in all other respects, the two propositions are identical, yet differ in virtue of the fact that the concepts which occur in them are conceptually diverse. Conceptual diversity implies numerical diversity, but the converse implication does not hold, since not all terms are concepts. Numerical diversity, as its name implies, is the source of plurality, and conceptual diversity is less important to mathematics. But the whole possibility of making different assertions about a given term or set of terms depends upon conceptual diversity, which is therefore fundamental in general logic.

51. It is interesting and not unimportant to examine very briefly the connection of the above doctrine of adjectives with certain traditional views on the nature of propositions. It is customary to regard all propositions as having a subject and a predicate, i.e. as having an immediate *this*, and a general concept attached to it by way of description. This is, of course, an account of the theory in question which will strike its adherents as extremely crude; but it will serve for a general indication of the view to be discussed. This doctrine develops by internal logical necessity into the theory of Mr. Bradley's Logic, that all words stand for ideas having what he calls *meaning*, and that in every judgment there is a something, the true subject of the judgment, which is not an idea and does not have meaning. To have meaning, it seems to me, is a notion confusedly compounded of logical and psychological elements. *Words* all have meaning, in the simple sense that they are symbols which stand for something other than themselves. But a proposition, unless it happens to be linguistic, does not itself contain words: it contains the entities indicated by words. Thus meaning, in the sense in which words have meaning, is irrelevant to logic. But such concepts as *a man* have meaning in another sense: they are, so to speak, symbolic in their own logical nature, because they have the property which I call *denoting*. That is to say, when *a man* occurs in a proposition (e.g. "I met a man in the street"), the proposition is not about the concept *a man*, but about something quite different, some actual biped denoted by the concept. Thus concepts of this kind have meaning in a non-psychological sense. And in this sense, when we say "this is a man", we are making a proposition in which a concept is in some sense attached to what is not a concept. But when meaning is thus understood, the entity indicated by *John* does not have meaning, as Mr Bradley contends[7]; and even among concepts, it is

7 *Logic*, Book I, ch. 1, §§ 17, 18 (pp. 58–60).

only those that denote that have meaning. The confusion is largely due, I believe, to the notion that *words* occur in propositions, which in turn is due to the notion that propositions are essentially mental and are to be identified with cognitions. But these topics of general philosophy must be pursued no further in this work.

52. It remains to discuss the verb, and to find marks by which it is distinguished from the adjective. In regard to verbs also, there is a twofold grammatical form corresponding to a difference in merely external relations. There is the verb in the form which it has as verb (the various inflexions of this form may be left out of account), and there is the verbal noun, indicated by the infinitive or (in English) the present participle. The distinction is that between "Felton killed Buckingham" and "Killing no murder". By analyzing this difference, the nature and function of the verb will appear.

It is plain, to begin with, that the concept which occurs in the verbal noun is the very same as that which occurs as verb. This results from the previous argument, that every constituent of every proposition must, on pain of self-contradiction, be capable of being made a logical subject. If we say "*kills* does not mean the same as to *kill*", we have already made *kills* a subject, and we cannot say that the concept expressed by the word *kills* cannot be made a subject. Thus the very verb which occurs as verb can occur also as subject. The question is: What logical difference is expressed by the difference of grammatical form? And it is plain that the difference must be one in external relations. But in regard to verbs, there is a further point. By transforming the verb, as it occurs in a proposition, into a verbal noun, the whole proposition can be turned into a single logical subject, no longer asserted, and no longer containing in itself truth or falsehood. But here too, there seems to be no possibility of maintaining that the logical subject which results is a different entity from the proposition. "Caesar died" and "the death of Caesar" will illustrate this point. If we ask: What is asserted in the proposition "Caesar died"? the answer must be "the death of Caesar is asserted". In that case, it would seem, it is the death of Caesar which is true or false; and yet neither truth nor falsity belongs to a mere logical subject. The answer here seems to be that the death of Caesar has an external relation to truth or falsehood (as the case may be), whereas "Caesar died" in some way or other contains its own truth or falsehood as an element. But if this is the correct analysis, it is difficult to see how "Caesar died" differs from "the truth of Caesar's death" in the case where it is true, or "the falsehood of Caesar's death" in the other case. Yet it is quite plain that the latter, at any rate, is never equivalent to "Caesar died". There appears to be an ultimate notion of assertion, given by the verb, which is lost as soon as we substitute a verbal noun, and is lost when the proposition in question is

made the subject of some other proposition. This does not depend upon grammatical form; for if I say "*Caesar died* is a proposition", I do not assert that Caesar did die, and an element which is present in "Caesar died" has disappeared. Thus the contradiction which was to have been avoided, of an entity which cannot be made a logical subject, appears to have here become inevitable. This difficulty, which seems to be inherent in the very nature of truth and falsehood, is one with which I do not know how to deal satisfactorily. The most obvious course would be to say that the difference between an asserted and an unasserted proposition is not logical, but psychological. In the sense in which false propositions may be asserted, this is doubtless true. But there is another sense of assertion, very difficult to bring clearly before the mind, and yet quite undeniable, in which only true propositions are asserted. True and false propositions alike are in some sense entities, and are in some sense capable of being logical subjects; but when a proposition happens to be true, it has a further quality, over and above that which it shares with false propositions, and it is this further quality which is what I mean by assertion in a logical as opposed to a psychological sense. The nature of truth, however, belongs no more to the principles of mathematics than to the principles of everything else. I therefore leave this question to the logicians with the above brief indication of a difficulty.

53. It may be asked whether everything that, in the logical sense we are concerned with, is a verb, expresses a relation or not. It seems plain that, if we were right in holding that "Socrates is human" is a proposition having only one term, the *is* in this proposition cannot express a relation in the ordinary sense. In fact, subject-predicate propositions are distinguished by just this non-relational character. Nevertheless, a relation between Socrates and humanity is certainly *implied*, and it is very difficult to conceive the proposition as expressing no relation at all. We may perhaps say that it is a relation, although it is distinguished from other relations in that it does not permit itself to be regarded as an assertion concerning either of its terms indifferently, but only as an assertion concerning the referent. A similar remark may apply to the proposition "*A* is", which holds of every term without exception. The *is* here is quite different from the *is* in "Socrates is human"; it may be regarded as complex, and as really predicating Being of *A*. In this way, the true logical verb in a proposition may be always regarded as asserting a relation. But it is so hard to know exactly what is meant by *relation* that the whole question is in danger of becoming purely verbal.

54. The twofold nature of the verb, as actual verb and as verbal noun, may be expressed, if all verbs are held to be relations, as the difference between a relation in itself and a relation actually relating. Consider, for example, the proposition "*A* differs from *B*". The constituents of this proposition, if we analyze it, appear to be only *A*, difference, *B*. Yet these constituents, thus

placed side by side, do not reconstitute the proposition. The difference which occurs in the proposition actually relates *A* and *B*, whereas the difference after analysis is a notion which has no connection with *A* and *B*. It may be said that we ought, in the analysis, to mention the relations which difference has to *A* and *B*, relations which are expressed by *is* and *from* when we say "*A* is different from *B*". These relations consist in the fact that *A* is referent and *B* relatum with respect to difference. But "*A*, referent, difference, relatum, *B*" is still merely a list of terms, not a proposition. A proposition, in fact, is essentially a unity, and when analysis has destroyed the unity, no enumeration of constituents will restore the proposition. The verb, when used as a verb, embodies the unity of the proposition, and is thus distinguishable from the verb considered as a term, though I do not know how to give a clear account of the precise nature of the distinction.

55. It may be doubted whether the general concept *difference* occurs at all in the proposition "*A* differs from *B*", or whether there is not rather a specific difference of *A* and *B*, and another specific difference of *C* and *D*, which are respectively affirmed in "*A* differs from *B*" and "*C* differs from *D*". In this way, *difference* becomes a class-concept of which there are as many instances as there are pairs of different terms; and the instances may be said, in Platonic phrase, to partake of the nature of difference. As this point is quite vital in the theory of relations, it may be well to dwell upon it. And first of all, I must point out that in "*A* differs from *B*" I intend to consider the bare numerical difference in virtue of which they are two, not difference in this or that respect.

Let us first try the hypothesis that *a* difference is a complex notion, compounded of difference together with some special quality distinguishing a particular difference from every other particular difference. So far as the relation of difference itself is concerned, we are to suppose that no distinction can be made between different cases; but there are to be different associated qualities in different cases. But since cases are distinguished by their terms, the quality must be primarily associated with the terms, not with difference. If the quality be not a relation, it can have no special connection with the difference of *A* and *B*, which it was to render distinguishable from bare difference, and if it fails in this it becomes irrelevant. On the other hand, if it be a new relation between *A* and *B*, over and above difference, we shall have to hold that any two terms have two relations, difference and a specific difference, the latter not holding between any other pair of terms. This view is a combination of two others, of which the first holds that the abstract general relation of difference itself holds between *A* and *B*, while the second holds that when two terms differ they have, corresponding to this fact, a specific relation of difference, unique and unanalyzable and not shared by any other pair of terms. Either of these views may be held with

either the denial or the affirmation of the other. Let us see what is to be said for and against them.

Against the notion of specific differences, it may be urged that, if differences differ, their differences from each other must also differ, and thus we are led into an endless process. Those who object to endless processes will see in this a proof that differences do not differ. But in the present work, it will be maintained that there are no contradictions peculiar to the notion of infinity, and that an endless process is not to be objected to unless it arises in the analysis of the actual meaning of a proposition. In the present case, the process is one of implications, not one of analysis; it must therefore be regarded as harmless.

Against the notion that the abstract relation of difference holds between *A* and *B*, we have the argument derived from the analysis of "*A* differs from *B*", which gave rise to the present discussion. It is to be observed that the hypothesis which combines the general and the specific difference must suppose that there are two distinct propositions, the one affirming the general, the other the specific difference. Thus if there cannot be a general difference between *A* and *B*, this mediating hypothesis is also impossible. And we saw that the attempt to avoid the failure of analysis by including in the *meaning* of "*A* differs from *B*" the relations of difference to *A* and *B* was vain. This attempt, in fact, leads to an endless process of the inadmissible kind; for we shall have to include the relations of the said relations to *A* and *B* and difference, and so on, and in this continually increasing complexity we are supposed to be only analyzing the *meaning* of our original proposition. This argument establishes a point of very great importance, namely, that when a relation holds between two terms, the relations of the relation to the terms, and of these relations to the relation and the terms, and so on *ad infinitum*, though all implied by the proposition affirming the original relation, form no part of the *meaning* of this proposition.

But the above argument does not suffice to prove that the relation of *A* to *B* cannot be abstract difference: it remains tenable that, as was suggested to begin with, the true solution lies in regarding every proposition as having a kind of unity which analysis cannot preserve, and which is lost even though it be mentioned by analysis as an element in the proposition. This view has doubtless its own difficulties, but the view that no two pairs of terms can have the same relation both contains difficulties of its own and fails to solve the difficulty for the sake of which it was invented. For, even if the difference of *A* and *B* be absolutely peculiar to *A* and *B*, still the three terms *A*, *B*, difference of *A* from *B*, do not reconstitute the proposition "*A* differs from *B*", any more than *A* and *B* and difference did. And it seems plain that, even if differences did differ, they would still have to have something in common. But the most general way in which two terms can have

something in common is by both having a given relation to a given term. Hence if no two pairs of terms can have the same relation, it follows that no two terms can have anything in common, and hence different differences will not be in any definable sense *instances* of difference.[8] I conclude, then, that the relation affirmed between *A* and *B* in the proposition "*A* differs from *B*" is the general relation of difference, and is precisely and numerically the same as the relation affirmed between *C* and *D* in "*C* differs from *D*". And this doctrine must be held, for the same reasons, to be true of all other relations; relations do not have instances, but are strictly the same in all propositions in which they occur.

We may now sum up the main points elicited in our discussion of the verb. The verb, we saw, is a concept which, like the adjective, may occur in a proposition without being one of the terms of the proposition, though it may also be made into a logical subject. One verb, and one only, must occur as verb in every proposition; but every proposition, by turning its verb into a verbal noun, can be changed into a single logical subject, of a kind which I shall call in future a propositional concept. Every verb, in the logical sense of the word, may be regarded as a relation; when it occurs as verb, it actually relates, but when it occurs as verbal noun it is the bare relation considered independently of the terms which it relates. Verbs do not, like adjectives, have instances, but are identical in all the cases of their occurrence. Owing to the way in which the verb actually relates the terms of a proposition, every proposition has a unity which renders it distinct from the sum of its constituents. All these points lead to logical problems, which, in a treatise on logic, would deserve to be fully and thoroughly discussed.

[...]

427. I come now to the second of Lotze's objections to empty space. This is again of a somewhat abstract logical character, but it is far easier to dispose of, since it depends upon a view more or less peculiar to Lotze. There are, it says, three and only three kinds of being, no one of which belongs to space. These are (α) the being of things, which consists in activity or the power to produce effects; (β) the validity of a truth; (γ) the being which belongs to the contents of our presentations.

The answer to this is, that there is only one kind of being, namely, being *simpliciter*, and only one kind of existence, namely, existence *simpliciter*. Both

8 The above argument appears to prove that Mr Moore's theory of universals with numerically diverse instances in his paper on Identity (*Proceedings of the Aristotelian Society*, 1900–1) must not be applied to all concepts. The relation of an instance to its universal, at any rate, must be actually and numerically the same in all cases where it occurs.

being and existence, I believe, belong to empty space; but being alone is relevant to the refutation of the relational theory – existence belongs to the question which Lotze confounds with the above, namely, as to the reality or subjectivity of space. It may be well first to explain the distinction of being and existence, and then to return to Lotze's three kinds of being.

Being is that which belongs to every conceivable term, to every possible object of thought – in short to everything that can possibly occur in any proposition, true or false, and to all such propositions themselves. Being belongs to whatever can be counted. If *A* be any term that can be counted as one, it is plain that *A* is something, and therefore that *A* is. "*A* is not" must always be either false or meaningless. For if *A* were nothing, it could not be said not to be; "*A* is not" implies that there is a term *A* whose being is denied, and hence that *A* is. Thus unless "*A* is not" be an empty sound, it must be false – whatever *A* may be, it certainly is. Numbers, the Homeric gods, relations, chimeras and four-dimensional spaces all have being, for if they were not entities of a kind, we could make no propositions about them. Thus being is a general attribute of everything, and to mention anything is to show that it is.

Existence, on the contrary, is the prerogative of some only amongst beings. To exist is to have a specific relation to existence – a relation, by the way, which existence itself does not have. This shows, incidentally, the weakness of the existential theory of judgment – the theory, that is, that every proposition is concerned with something that exists. For if this theory were true, it would still be true that existence itself is an entity, and it must be admitted that existence does not exist. Thus the consideration of existence itself leads to non-existential propositions, and so contradicts the theory. The theory seems, in fact, to have arisen from neglect of the distinction between existence and being. Yet this distinction is essential, if we are ever to deny the existence of anything. For what does not exist must be something, or it would be meaningless to deny its existence; and hence we need the concept of being, as that which belongs even to the non-existent.

7

THE EXISTENTIAL IMPORT OF PROPOSITIONS

1905

The famous paper "On denoting" offers the fullest expression of Russell's theory of descriptions but it was not the first place in which the ideas were presented. Nor was it the simplest or clearest expression as it contains a notoriously convoluted account of Frege. It is also anthologised elsewhere so, as representative of the metaphysical basis of the theory which was so crucial to Russell's philosophical development, the following paper is included, which is shorter and in many ways a clearer summary.

Russell discovered his theory of descriptions in the two months after moving, on 24 April 1905, to his house at Bagley Wood. "On denoting" was written at the end of July and published in the October issue of *Mind*. "On the existential import of propositions" was published in the July issue, however, and from the evidence of his correspondence with MacColl (see *Papers* 4: 480–1), was composed in April and May. The paper is an absolutely crucial one in the development of Russell's theory of descriptions, therefore.

Russell was replying to the claims made by MacColl in two papers in *Mind* (MacColl 1905 and 1905a). MacColl had said that the null-class is the class of all non-existent things. Russell responds that the null-class is the class with no members, rather than the class of non-existent members, hence developing his critique of Meinong. To say, for example, that no chimeras exist is to say that the class of chimeras has no members, rather than, as MacColl, that chimeras exist as non-existences. Likewise, "The present King of France" does not denote an unreal individual; rather, it fails to denote.

Mr. MacColl's interesting paper in the January number of *Mind*, together with his note in the April number, raises certain points which call for an answer from those who (like myself) adhere to the usual stand-point of symbolic logicians on the subject of the existential import of propositions.

The first point in regard to which clearness is essential concerns the meaning of the word "existence". There are two meanings of this word, as

distinct as stocks in a flower-garden and stocks on the Stock Exchange, which yet are continually being confused, or at least supposed somehow connected. Of these meanings, only one occurs in philosophy or in common parlance, and only the other occurs in mathematics or in symbolic logic. Until it is realized that they have absolutely nothing to do with each other, it is quite impossible to have clear ideas on our present topic.

(*a*) The meaning of *existence* which occurs in philosophy and in daily life is the meaning which can be predicated of an individual, the meaning in which we inquire whether God exists, in which we affirm that Socrates existed, and deny that Hamlet existed. The entities dealt with in mathematics do not exist in this sense: the number. 2, or the principle of the syllogism, or multiplication, are objects which mathematics considers, but which certainly form no part of the world of existent things. This sense of existence lies wholly outside Symbolic Logic, which does not care a pin whether its entities exist in this sense or not.

(*b*) The sense in which existence is used in symbolic logic is a definable and purely technical sense, namely this: To say that *A* exists means that *A* is a class which has at least one member. Thus whatever is not a class (e.g. Socrates) does not exist in this sense; and among classes there is just one which does not exist, namely, the class having no members, which is called the null-class. In this sense, the class of numbers (e.g.) exists, because 1, 2, 3, etc., are members of it; but in sense (*a*) the class and its members alike do not exist: they do not stand out in a part of space and time, nor do they have that kind of super-sensible existence which is attributed to the Deity.

It may be asked: How come two such diverse notions to be confounded? It is easy to see how the confusion arises, by considering classes which, if they have members at all, must have members that exist in sense (*a*). Suppose we say: "No chimeras exist." We may mean that the class of chimeras has no members, i.e., does not exist in sense (*b*), or that nothing that exists in sense (*a*) is a chimera. These two are equivalent in the present instance, because if there were chimeras, they would be entities of the kind that exist in sense (*a*). But if we say "no numbers exist", our statement is true in sense (*a*) and false in sense (*b*). It is true that nothing that exists in sense (*a*) is a number; it is false that the class of numbers has no members. Thus the confusion arises from undue preoccupation with the things that exist in sense (*a*), which is a bad habit engendered by practical interests.

Mr. MacColl assumes (p. 74) two universes, the one composed of existences, the other of non-existences. It will be seen that, if the above discrimination is accepted, these two universes are not to be distinguished in symbolic logic. All entities, whether they exist or whether they do not

(in sense (*a*)), are alike real to symbolic logic and mathematics. In sense (*b*), which is alone relevant, there is among classes not a multitude of non-existences, but just one, namely, the null-class. All the members of every class are among realities,[1] in the only sense in which symbolic logic is concerned with realities.

But it is natural to inquire what we are going to say about Mr. MacColl's classes of unrealities, centaurs, round squares, etc. Concerning all these we shall say simply that they are classes which have no members, so that each of them is identical with the null-class. There are no Centaurs; "*x* is a Centaur" is false whatever value we give to *x*, even when we include values which do not exist in sense (*a*), such as numbers, propositions, etc. Similarly, there are no round squares. The case of nectar and ambrosia is more difficult, since these seem to be individuals, not classes. But here we must presuppose definitions of nectar and ambrosia: they are substances having such and such properties, which, as a matter of fact, no substances do have. We have thus merely a defining concept for each, without any entity to which the concept applies. In this case, the concept is an entity, but it does not denote anything. To take a simpler case: "The present King of England" is a complex concept denoting an individual; "the present King of France" is a similar complex concept denoting nothing. The phrase intends to point out an individual, but fails to do so: it does not point out an unreal individual, but no individual at all. The same explanation applies to mythical personages, Apollo, Priam, etc. These words have a *meaning*, which can be found by looking them up in a classical dictionary; but they have not a *denotation*: there is no entity, real or imaginary, which they point out.

It will now be plain, I hope, that the ordinary view of symbolic logicians as to existential import does not require Mr. MacColl's modifications. This view is, that *A* and *E* do not imply the existence, in sense (*b*), of their subjects, but that *I* and *O* do imply the existence, in sense (*b*), of their subjects. No one of the four implies the existence, in sense (*a*), either of its subject or of any of the members of its subject. We have, adopting Peano's interpretation:

A All *S* is *P* = For all values of *x*, "*x* is an *S*" implies "*x* is a *P*".
E No *S* is *P* = For all values of *x*, "*x* is an *S*" implies "*x* is not a *P*".
I Some *S* is *P* = For at least one value of *x*, "*x* is an *S*" and "*x* is a *P*" are both true.
O Some *S* is not *P* = For at least one value of *x*, "*x* is an *S*" and "*x* is not a *P*" are both true.

1 This holds even of the null-class. Of all the members of the null-class, *every* statement holds, since the null-class has no members of which it does not hold. See below, on the interpretation of the universal affirmative *A*.

Thus *I* and *O* require that there should be at least one value of x for which x is an S, i.e. that S should exist in sense (*b*). *I* also requires that P should exist, and *O* requires that not-P should exist. But *A* and *E* do not require the existence of either S or P; for a hypothetical is true whenever its hypothesis is false,[2] so that if "x is an S" is always false, "All S is P" and "No S is P" will both be true whatever P may be.

The above remarks serve to answer the objection raised by Mr. MacColl in the April number of *Mind* (p. 295) to the equation $0A = 0$. To begin with, 0 does not represent the class of non-existences, but the non-existent class, i.e. the class which has no members. Thus, if "$XA = X$" means "every X is an A",[3] then "$0A = 0$" means "every member of the class which has no members is an A", or "for every value of x, 'x is a member of the class which has no members' implies 'x is an A'". This hypothetical is true for all values of x, because its hypothesis is false for all values of x, and a hypothetical with a false hypothesis is true. Thus Mr. MacColl's objection rests upon his taking 0 to be the class of non-existences, presumably in sense (*a*), since only so would 0 be a class with many members, all of them unreal, as he supposes it to be. The true interpretation of 0, as the non-existent class, in sense (*b*), at once disposes of the difficulty.

The same principles solve Lewis Carroll's paradox, noticed by "W" in the April number of *Mind* (p. 293). I cannot agree with "W" in regarding the paradox as merely verbal; on the contrary, I consider it a good illustration of the principle that a false proposition implies every proposition. Putting p for "Carr is out", q for "Allen is out", and r for "Brown is out", Lewis Carroll's two hypothetical are:

(1) q implies r.
(2) p implies that q implies not-r.

Lewis Carroll supposes that "q implies r" and "q implies not-r" are inconsistent, and hence infers that p must be false. But as a matter of fact, "q implies r" and "q implies not-r" must both be true if q is false, and are by no means inconsistent. The contradictory of "q implies r" is "q does not imply r", which is not a consequence of "q implies not-r". Thus the only inference from Lewis Carroll's premisses (1) and (2) is that if p is true, q is false, i.e. if Carr is out, Allen is in. This is the complete solution of the paradox.

2 See my *Principles of Mathematics*, vol. 1, p. 18.

3 Not "every X is A", as Mr. MacColl says, and as most logicians say.

8

THREE LETTERS TO MEINONG

1904–7

Despite the increasing gulf between Russell and Meinong over the issues of *being* and *existence*, and despite the ridicule poured over the German in "On denoting", relations between the two philosophers remained civilised. An intermittent exchange was sustained over a couple of years and Russell's contributions are reproduced below. Russell wrote these in German and they were published in a collection of Meinong's correspondence (Kindinger 1965). The versions here are Douglas Lackey's translations which first appeared in 1973.

In the first letter, of 1904, we find Russell stressing what is common to both philosophers. An unreal object, such as the golden mountain, still has *being*, Russell admits, but this *being* belongs only to its sense and the term has no reference. This suggests that Russell had almost arrived at his theory of denoting. He is utilising here the famous Fregean distinction between *sinn* (sense) and *bedeutung* (reference) (Frege 1892) but is not quite rebutting the idea of non-existent existences.

Two years later, Russell was still claiming that the differences between Meinong and himself were "unimportant", though this is starting to sound slightly hollow. Russell accepts, as in no way detrimental to his analysis, that impossible and merely contingently non-existent objects are to be put into the same category, and he now rejects the view that anything about the existence of an object follows from it being the subject of a sentence. The method of "On denoting" shows how these sentences can be analysed without any existence claim.

In the final letter, though still maintaining a show of cordiality, Russell has even less sympathy. He maintains, sensibly, that mathematics (and metaphysics in Letter II), being a priori, is existence-free. That is, it assumes certain hypotheses or premises and shows what follows. The nature of space itself must be decided empirically, not a priori by geometry. This view applied to metaphysics accords with what contemporary metaphysician David

Armstrong calls his Irish principle (Armstrong 1978: 11 and 1983: 84). In the current context, this would be something along the lines of *if the existence of some property, relation or thing can be demonstrated a priori, that is good reason to think that there is no such property, relation or thing*.

THREE LETTERS TO MEINONG: A TRANSLATION

[...]

Letter I

Ivy Lodge, Tilford, Farnham.
15 . XII . 1904.

My dear sir,

Many thanks for your friendly letter, and for the treatise "On the Theory of Objects". I have read this treatise, as well as Parts 2–7 of Dr. Ameseder's, with the greatest interest. I find myself in complete sympathy with its general standpoint, and the problems which it treats are such as seem to me important. I am accustomed to using the word "logic" for what you call "the theory of objects"; and the reasons which you present against this use, on pp. 20 ff., seem to me hardly decisive. But this is unimportant, and I concede that a new standpoint should be given a new name.

I have always believed until now that every object must in some sense have *being*, and I find it difficult to admit unreal objects. In such a case as that of the golden mountain or the round square one must distinguish between *sense* and *reference* (to use Frege's terms): the sense is an object, and has being; the reference, however, is not an object. The difference between sense and reference is best illustrated by mathematical examples: "the square root of 4" is a complex sense, the reference of which is the number 2.

The opinion that mathematics is a theory of objects I agree with completely; it is one of the principal theses of my *Principles* [*of Mathematics*]. If you don't have this book, I'll be glad it to send it to your address. Its entire first part is explicitly concerned with questions relating to the theory of objects. Of course there are many discussions whose purpose is purely formal, i.e. they serve only to lead up to the technical mathematical treatment; but the non-technical questions are the most important that are dealt with in it.

In connection with what you say on p. 40 about metaphysics, although I am inclined towards your general view, I feel a certain difficulty: concerning *everything* that exists, empiricism cannot instruct us; consequently, if there is metaphysics, it must be *a priori*.

I hope that your philosophical views will be rapidly circulated, and it will be a pleasure for me to contribute to this as much as possible.

Yours respectfully,
Bertrand Russell.

Letter II

Bagley Wood, Oxford.
5 . XI . 1906.

My dear sir,

Many thanks for your friendly letter, and for your interesting article "On the Place of the Theory of Objects in the System of Sciences". I too am of the opinion that the differences between us are unimportant. In general I find myself sharing the same standpoint as you do. I agree especially when you assert that mathematics is an "existence-free science" and belongs properly to the theory of objects.

With regard to impossible objects, I am in no way deterred by the consequence that, according to my view, the golden mountain should be put into the same class as the round square. Accordingly I have in my article "On Denoting" used the King of France as an example. As you know, there is for me no fundamental concept of necessity: consequently I cannot distinguish between Impossible and non-existent objects. Moreover, I cannot see how one can distinguish between "to exist" and "to be existent". That sentences, true [as well as] false, can be formed in which impossible objects appear as grammatical subjects, of course I do not deny; but I believe that these sentences must be interpreted in the manner I have indicated in my article "On Denoting".

What you say about Frege on p. 51 pleases me greatly. Because of his extraordinary difficulty he is very little read, but he is, in my opinion, worthy of the highest recognition.

Yours most respectfully,
Bertrand Russell.

Letter III

Bagley Wood, Oxford.
5 . II . 1907.

My dear sir,

Many thanks for sending your second article "On the Place of the Theory of Objects in the System of Sciences", which I find very interesting.

I have read carefully what you have written about the concept of necessity, and I believe that the difference of opinion between us is not so great as it first seemed. I fully acknowledge an *epistemic* distinction between *a priori* and empirical knowledge; but it seems to me that the related distinction of the appropriate Objectives consists fully in this, that *a priori* knowledge is always existence-free, while empirical knowledge is always existential. The word *necessary* is ambiguous in ordinary speech, and only a rather long discussion could elucidate all the possible senses of this word.

With what you write concerning non-Euclidean Geometry I am unfortunately not in sympathy. My own opinions I have often defended in the *Revue de Métaphysique et de Morale* against Poincaré; also in *Principles of Mathematics*, Part VI, and briefly in *Mind*, July 1905, pp. 414–5. Non-Euclidean Geometry does not assert that two parallels can intersect; it questions whether there are parallels. I am also of the opinion that Geometry is an existence-free science, insofar as Geometry is pure mathematics. As pure mathematics all geometries are equally true; they merely assert what follows from certain premisses – they are all equally hypothetical. But there is also *one* space that exists, or at least so pertains to existence, that can be called the space of the real physical world. Whether this space is an example of Euclidean or non-Euclidean Geometry can, in my opinion, only be decided empirically. That two parallels cannot intersect is indubitable; but it has to be asked whether the real world admits of parallels or not.

In order to be able to assert the possibility of empirical knowledge of spatial relations, it must of course be admitted that real relations can be given empirically. Then it must be asked: Are the perceived spatial relations (or those derived from perception) Euclidean or non-Euclidean? Mathematics shows that any class that through relations generates a Euclidean space, at the same time, through other subsisting relations, also generates all non-Euclidean spaces. Out of all these systems of relations there is, however, in the actual world, only one system of which it can in a certain sense be said that the relations of which it consists *exist*. In your discussion I find nothing, so far as I can see, which refutes this view.

By and large I concur with your writings, and so it is useful to discuss details. With friendly wishes I remain,

Yours sincerely,
Bertrand Russell.

9

THE BASIS OF REALISM

1911

In 1911, six unremarkable American philosophers published "The program and first platform of six realists" in which, after a joint statement, each gave their views on what they thought realism was in philosophy. Given that Russell thought of himself as a realist, he was inspired to write down his views which were published immediately after theirs. It is clear, however, that Russell's account has a widely different emphasis to that of the Americans. Given that Russell was a central figure in the realist revolt against neo-Hegelianism, however, it is to be expected that we see him highlighting the differences from Bradley more than what is common with the Americans.

Realism in philosophy has various meanings. After the work of Michael Dummett, it has come to be associated with the rejection of his position of anti-realism (Dummett 1978). C. B. Martin, for instance, supports realist intuitions that there are truths wholly inaccessible to empirical detection (Martin 1994). Anti-realists reject claims of verification transcendent truth. Hence the statement "a dinosaur coughed, here, 200 million years ago" is neither true nor false, if there remains no evidence or no way of obtaining evidence for its truth or falsity.

It is thus curious to see Russell opening his statement of realism with the claim that "The fundamental doctrine in the realistic position, as I understand it, is the doctrine that relations are 'external' ". Put in historical context, however, this can be understood as Russell taking realism to get its meaning by contrast with idealism – chiefly idealism of the Bradleyan variety.

That relations are "external" is explained by Russell as meaning that:

(1) relatedness does not imply any corresponding complexity in the *relata*,
(2) any given entity is a constituent of many different complexes.

These are both anti-Bradleyan claims. The Bradleyan says that, for any relation such as *a* being the father of *b*, the relation holds

in virtue of some element in *a*:*a* having the nature of being the father of *b* (and some element in *b*:*b* having the nature of being fathered by *a*). If the reader finds any plausibility in this example, perhaps they will not with the following. Suppose *a* is nearer the North Pole than *b*. Are we to suppose that *a* has, in its nature, an element of being nearer the North Pole than *b*? Instead, Russell came to see that we need to posit *a* and *b* and a relation between them that was not contained in some element of *a* or *b* but was separate or "external" to them. Hence, *a* and *b* have a certain relation external to, but between, them that is not contained in either one: the relation of one (*a*) being nearer the North Pole than the other (*b*). This relation can change without any change in the internal nature of either. Hence *b* can be moved so that it is nearer the North Pole than *a* but without any other change in the non-relational or "intrinsic" natures of either. Such change-without-intrinsic-change is just what Russell's account says it is: purely relational, that is, purely in the relation between *a* and *b* and not in *a* or *b* themselves.

Second, the Bradleyan has a position which does not allow the same thing to be involved in more than one relation at a time. Where *a* is the father of *b* but the son of *c*, it is not, according to Russell's opponent, the same *a* that is father of *b* and son of *c*. Rather it is *a qua father of b* that is father of *b* and *a qua son of c* that is son of *c*; and *a qua father of b* and *a qua son of c* are always two different entities. On Russell's view, a single entity may be involved in many different relations simultaneously so *a* may be father of *b*, son of *c*, taller than *d*, nearer the North Pole than *e*, and so on, and it is that same entity *a* in each case that stands in these different relations to these different things. That this account of the nature of relational complexes seems so plausible to us shows what a debt we owe to Russell and others who were able to overcome the confused and confusing received view of their time.

However, we may not have heard the last of internal relations. While the example of … *being nearer the North Pole than* … might seem a plausible external relation, it is not a proof that *all* relations are external. Some metaphysicians now argue that there are properties that have an internal relation to manifestations of a certain kind (Ellis 2001; Lombardo 2002). While this is not a return to Bradley, it is a rejection of a Humean metaphysic that has the basic elements of reality as causally inert, discrete atoms, freely recombinable with all other elements. Russell's metaphysics thus falls into the category now under attack by the anti-Humeans.

Russell goes on, towards the end of the paper, to discuss issues far closer to contemporary realism. He affirms the current realist doctrine that there are unknown truths. That Russell took up the opportunity to express such views shows that he thought the

> battle against the fallacies of his "anti-realist" opponents had not finally been won and, despite much time for us to reflect, it seems that the battle is still not won.

"The Program and First Platform of Six Realists", which appeared in this *Journal* last July,[1] gives expression to a growing movement in philosophy in a way which deserves the gratitude of all who are in sympathy with that movement. As I find myself in almost complete agreement with the "six realists", I have thought it might be desirable to attempt on my own account a somewhat similar statement of my philosophical opinions.

The fundamental doctrine in the realistic position, as I understand it, is the doctrine that relations are "external". This doctrine is not correctly expressed by saying that two terms which have a certain relation might have not had that relation. Such a statement introduces the notion of possibility and thus raises irrelevant difficulties. The doctrine may be expressed by saying that (1) relatedness does not imply any corresponding complexity in the *relata*; (2) any given entity is a constituent of many different complexes. Each of these propositions requires some expansion.

1. The view which is intended to be denied by this proposition is the view that, whenever a term a has a certain relation R to a certain other term b, that implies some element in a in virtue of which it has the relation R to b. It is usual in this connection to speak rather of the "nature" of a than of a, and it is not clear whether the "nature" of a is or is not identical with a; but in any case, this "nature", according to the view we are denying, is complex, and contains a constituent which expresses or accounts for a's relation to b. Writers who advocate the view in question do not state what they mean by the "nature" of a. Three views may be suggested: (α) We may suppose that the "nature" of a is identical with a. In this case, we affirm, on the basis of a *reductio ad absurdum*, that the contention that a's "nature" must contain a constituent which expresses or accounts for a's relation to b, is demonstrably false. (β) We may suppose that a's "nature" is all the propositions that are true of a, or all the complexes of which a is a constituent. In this case, the view in question becomes a truism, but fails to yield any of the consequences commonly deduced from it. (γ) We may define the "nature" of a as its predicates or attributes, as opposed to its relations. This view requires a word of explanation. We are accustomed to dyadic, triadic, tetradic ... relations. A dyadic relation may be defined as one which *can* occur in propositions containing only two other terms, i.e. as one such that the simplest propositions in which it occurs contain only two other terms.

1 Vol. VII, no. 15, pp. 393–401.

Similar definitions apply to triadic, tetradic ... relations. Now there may also be – I do not say there are – what we may call *monadic* concepts, i.e. concepts which can occur in propositions having only one other term. Such concepts may be called predicates or attributes. It is of course the case that, whenever a subject has a predicate, there is a dyadic relation of subject and predicate, but it does not follow that there is not also a proposition in which the predicate is merely predicated. The analogy of dyadic relations will make this clearer. Whenever *a* has the relation *R* to *b*, there is a triadic relation of *a* and *R* and *b*, but in this relation *R* occurs as a term of the relation, not as the relating relation of the proposition. Similarly, if there are monadic concepts, the propositions in which they are said to have the relation of predication to their subjects will not be identical with the propositions in which they are actually predicated. Assuming, then, that there are monadic concepts, the "nature" of *a* may consist of all those monadic concepts which are predicable of *a*. Then the view that *a*'s "nature" contains a constituent which expresses or accounts for *a*'s relation to *b* will be the view that there is a monadic concept predicable of whatever has the said relation to *b* and of nothing else. More generally, the view in question may be stated as follows: "Every propositional function of one variable is formally equivalent to some monadic propositional function", where a monadic propositional function is one which attributes a monadic concept to a variable subject. It is no concern of the doctrine of external relations either to affirm or deny the above view, though we must contend that there is no reason to suppose the above view to be true. We must, however, deny that a term is composed of all the monadic concepts which are predicable of it, i.e. that a particular subject is identical with the sum of its predicates.

I do not for a moment wish to suggest that any one of the above three views (α), (β), (γ), is held by any of the opponents of external relations They, I believe, hold a confused mixture of all three, and would not continue their opposition if the confusion were removed. My purpose in setting forth the above three views was to make it clear what it is that I affirm, and what it is that I deny. The view I advocate is, that a term *a* may have a relation to a term *b* without there being any constituent of *a* corresponding to this relation. If this were false, simple terms could have no relations, and therefore could not enter into complexes; hence every term would have to be strictly infinitely complex.

2. Authors who deny external relations hold, in addition to the doctrine discussed above, that it is impossible for precisely the same term to be a constituent of two different complexes, or to occur in two different propositions. They say that if *A* is the father of *B* and the son of *C*, it is not strictly the same entity which is father and son, but that it is "*A quâ* father of *B*" who is the father of *B*, and "*A quâ* son of *C*" who is the son of *C*.

This doctrine also is denied by those who advocate external relations. They would argue that "*A quâ* father of *B*" is a complex containing the constituent *A*, and "*A quâ* son of *C*" is also a complex containing the constituent *A*. Thus the attempt to avoid an identical constituent in two complexes breaks down. The two parts of the doctrine of external relations together constitute the justification of analysis, and the denial of the view that analysis is falsification.

The importance of the question as to the value of relations lies in the fact that current arguments against realism and pluralism almost all depend upon the doctrine of internal relations. When this doctrine is rejected, the question as to the number of things that exist becomes purely empirical, but no empirical fact is more certain, if a priori refutations fail, than that many things exist. I should therefore prefer to call the philosophy which I advocate "pluralism" rather than "realism", because realism, in most of its accepted uses, involves long and difficult arguments which might be rejected without contravening anything that was said above as to the nature of relations. Nevertheless, if any case is to be made out against this or that form of realism on the basis of external relations, it will have to be a new case, based upon quite different arguments from those hitherto employed by idealists. I do not myself believe that any such case can be made out; at the same time, all the questions involved seem to me to demand fresh discussion, and what seems to me so far firmly established is a logic and a method, rather than any positive metaphysical results. What is plain is that all arguments based on the contention that knowing makes a difference to what is known, or implies a community or interaction between knower and known, rest upon the internal view of relations, and therefore fail when this view is rejected. It is true there is another argument for idealism, namely the argument which Professor Perry calls the "egocentric predicament". This argument is, in brief, that everything must be known, because we cannot know of anything else. This is a foolish fallacy, which would equally prove that I must be acquainted with everybody whose name is Smith. Sometimes, more plausibly, it is urged merely that we cannot know that there are things we do not know; but this view rests upon a wrong analysis of general propositions, in fact upon the same wrong analysis which led Mill to regard Barbara as a *petitio principii*. When we know a general proposition, that does not require that we should know all or any of the instances of it. "All the multiplication-sums that never have been and never will be thought of by any human being deal with numbers over 1,000" is obviously a true proposition, although no instance of such a sum can ever be given. It is therefore perfectly possible to know that there are propositions we do not know, in spite of the fact that we can give no instance of such a proposition. Thus it is fallacious not merely to argue that everything must be known,

but even to argue that we cannot know that there are things which we do not know. This instance, like many others, illustrates the elementary blunders which philosophers have made owing to their neglect of logic.

To sum up: The primary philosophic effect of the logical doctrines which I share with the "six realists" seems to me mainly negative; it shows that most current philosophical argument is fallacious, and that many questions which have been supposed amenable to a priori treatment must be dealt with empirically, since logic leaves the alternatives undecided. At the same time, in all those matters in which philosophy has been led to contravene science and common sense, there is a presumption, if the arguments of philosophers have been unsound, that their conclusions have been false; hence the logic in question naturally associates itself with pluralism and realism rather than with monism and idealism. Moreover, by the rejection of a priori constructions the way is opened for philosophy to become inductive, and to begin the patient cooperative accumulation of results by which the triumphs of science have been achieved.

10

ANALYTIC REALISM

Extract 1911

"Analytic realism" is another paper with a manifesto-like quality. It was read in Paris on 23 March 1911 and the ensuing discussion, of which Russell's main contribution is included, was recorded verbatim. The analytic element, that Russell introduces here, gives a foretaste of his logical atomism. It is that "the existence of the complex depends on the existence of the simple, and not vice versa". There are two kinds of simples, or atoms: universals and particulars. Russell claims that complexes presuppose simples whereas simples do not presuppose complexes.

There are good grounds on which to dispute this, for reasons which later were to cast the logical atomist programme into doubt. The problem can be approached from an epistemological viewpoint or a metaphysical one.

The epistemological problem of atomism is that, while we can know something to be complex, we can never know that something is simple. Thus, I could see that some phenomenon has parts, and thereby know it to be complex, but to know that something is simple is to know that it does not have parts and this, it seems, cannot be known. Something may appear simple to me, but there will always be the possibility of it having unknown parts. Hence, Russell's assertion that "one must discover the simple elements of which complexes are composed" is an imperative it is impossible to fulfil with any satisfactory degree of certainty.

The metaphysical problem is that there seems no a priori reason why there should be simples. There may be simples but, also, there may be infinite complexity. Consider the history of investigation into the structure of matter. Atoms were first posited as simples until it was discovered that they had parts. These parts were then supposed to be the simples until further parts were discovered to some of them. Currently we think hadrons and leptons are without structure but can we be sure that they do not in turn have parts? Could every such entity have parts? Could there be

infinite complexity, all the way down?[a] Philosophers have not produced a good argument why not so it seems that if atomism is true, it is only contingently so and, prior to a conclusive argument, we should not have a philosophical commitment either way. This leaves us with the disconcerting result that atomism can neither be a priori nor a posteriori proved, but nor can it be a priori or a posteriori disproved.

Those in favour of Russell's general analytic approach can preserve a weakened form of atomism that avoids both the epistemological and metaphysical objections. We may not be able to say certainly that there are absolute, structureless atoms but we can say that there are relative atoms and relative atoms are all we need to say that something is complex. Russell's claim that "complexes presuppose simples" is wrong, if such simples are to be understood literally as simples, but it will be right if they are understood only as relative simples. Hence an atom is a simple relative to the complex molecule even if the atom is in turn complex relative to something else. This kind of relative atomism is developed in Armstrong's factualism, where facts are non-linguistic and equated with states of affairs. Some such states of affairs are relatively simple (Armstrong 1997: 122–3).

In addition to the atomism developed in this paper, Russell discusses some more general metaphysical matters. He argues, for example, that there are both universals and particulars. Universals are covered in more depth in Part III, below. Particulars he allows only the logical properties of traditional substances; that is, they are those things that are the subjects of predicates or the terms in relations. They are not given the metaphysical features of substances, as traditionally conceived, such as being independently existing underlying substrata. Particulars are also considered by Russell in Part IV and in paper 21.

The Philosophy which seems to me closest to the truth can be called "analytic realism". It is realist, because it claims that there are non-mental entities and that cognitive relations are external relations, which establish a direct link between the subject and a possibly non-mental object. It is analytic, because it claims that the existence of the complex depends on the existence of the simple, and not vice versa, and that the constituent of a complex, taken as a constituent, is absolutely identical with itself as it is when we do not consider its relations. This philosophy is therefore an atomic philosophy. The atoms are of two kinds: universals, which are similar to Platonic ideas, and particulars, which, in logic, are similar to substances, since they can never appear as predicates or relations in propositions. (In metaphysics, however, particulars hardly resemble substances, since they have no need for

a Russell admits this possibility in an answer to H. Wildon Carr (Russell 1918: 180).

permanent existence.) Known universals are called *concepts*; known particulars are called *sense-data* (including those of inner sense). Neither concepts nor sense-data are *in* the mind which knows them. There is no general reason to suppose that what is known depends in any way on knowledge. There are, however, reasons of detail for believing that sense-data exist only when they are perceived. The physical world studied by science is therefore in no way identical with the sensible world, and we believe there is a physical world only for inductive reasons, which are more or less probable. Universals, on the other hand, do not exist in the same sense as particulars do; one should rather say that they *subsist*. Their subsistence depends in no way on the minds which know them; the abstract sciences have, therefore, as their object something completely independent of any mental element.

This philosophy recommends itself partly for reasons of a general kind, and partly for reasons derived from the principles of mathematics. Most philosophies arrive at the conclusion that mathematical propositions cannot be completely true, and that they are more or less tainted with contradictions or inaccuracies. The philosophy I call analytic realism, on the other hand, yields the conclusion that there is no reason to doubt the absolute truth of mathematical propositions. It resolves all the contradictions, both ancient and modern, which have been found in mathematics, and it shows that the exactitude of mathematics depends on its being concerned with an abstract world, a world of essences or Platonic ideas, which is not subject to the inaccuracies of sensory knowledge, and which is not constrained by the subjective variations one would have to fear in accepting a psychological theory such as that of Kant. The transition from the world of abstractions to the world of particulars is the same as the transition from pure mathematics to applied mathematics; it is made by finding, in the actual world, something which verifies an hypothesis of pure mathematics.

Besides the completely general problems that analytic realism claims to solve, I should add that (to my mind) it provides the most satisfactory analysis of the infinite, the continuum, space, time and motion. The problem of relating the sensible flux to the abstract world of mathematical physics is, if I am not mistaken, completely solved by a mixture of modern mathematics and the philosophy to which it has led me. As with all philosophies, however, this one recommends itself not so much by its detailed and precise arguments, as by providing a harmonious picture of the universe and the sciences.

DISCUSSION

Mr. B. Russell – The philosophy I espouse is:

Realist, because it considers existence as independent of knowledge. There is an infinite regress in epistemological idealism, since knowledge

can exist only if it is known, etc. Therefore existence does not imply being known: there could be a completely material world, and in the actual world there could be some things of which we have no knowledge. Knowledge of an object is a direct relation between the mind and the object: there is no corresponding change in the mind, but only a relation; in other words, knowing is not having an idea of what is known. Moreover knowing does not imply that the mind and what it knows have a common nature. In this sense, the philosophy in question is realist. I am not saying, however, that sensible objects, i.e. the immediate sense-data, exist independently of us in a causal sense, but I reject the *a priori* arguments which purport to show that nothing can exist independently of the mind, and I believe that what is known cannot be nothing, and must, therefore, subsist, whereas what subsists may very well not be known.

Secondly, the philosophy I espouse is *analytic*, because it claims that one must discover the simple elements of which complexes are composed, and that complexes presuppose simples, whereas simples do not presuppose complexes. Many philosophers believe that the constituent of a complex, as such, is not exactly the same as what it is in itself, but is changed in becoming a constituent. This seems to me to rest on a confusion between the practical identity of people and things in everyday life, and logical identity. For example, a man is changed when he marries a certain woman. Does this mean that the spouses are not two people, whom we can consider without error (however incompletely) in reciprocal isolation? No; it only means that marriage causes changes such that the husband is not exactly the same man as the bachelor. But the analysis of a complex is not its biography, and hence there is no logical difficulty.

I believe there are simple beings in the universe, and that these beings have relations in virtue of which complex beings are composed. Any time a bears the relation R to b there is a complex "a in relation R to b". The change consists in the fact that there are complexes belonging to a time t, and that it might happen that, if we substitute another moment t' for t, there would no longer be a complex. Take motion, for example: One can have the complex "a at point p at time t" without there being "a at point p at time t'". If we have "a at point p' at time t'" we can say that a has moved between t and t'.

You will note that this philosophy is the philosophy of logical atomism. Every simple entity is an atom. One must not suppose that atoms need persist in time, or that they need occupy space: these atoms are purely logical.

You will also observe that in every complex there are two kinds of constituents: there are terms and the relation which relates them: or there might be (perhaps) a term qualified by a predicate. Note that the terms of a complex can themselves be relations, as, for example, in the statement that priority implies diversity. But there are some terms which appear only as

terms and can never appear as predicates or relations. These terms are what I call *particulars*. The other terms found in a complex, those which can appear as predicates or relations, I call *universals*. Terms like diversity, causality, father, white, etc., are *universals*. They are in effect Platonic ideas. Only one should not suppose that universals *exist* in the same sense that particulars exist; it is better to say that they *subsist. Particulars* have the purely logical properties of substances, but do not have their metaphysical properties. That is to say, particulars can only be either the subjects of predicates or the terms of relations. Particulars and only particulars *exist* in the narrow sense of the word. But they need not *persist*: they may exist only for an instant. Neither does their existence need to be causally independent of other existences. (From the logical point of view, any simple existence is independent of any other, and the only dependence is that of the complex on the simple.) The old idea of substance, as one can clearly see in Leibniz for example, contained two elements: (1) a logical element, (2) a metaphysical element, namely permanent and independent existence. I accept the first element, but I do not accept the second.

Some universals are known and some are not. The same can be said of particulars. Universals which are known are called *concepts*; particulars which are known are called *sense-data* (taking this expression in its widest possible interpretation). Concepts and sense-data are equally *objects* to the mind: that is to say, they are entities with some cognitive relationship to the mind. Neither the one nor the other is *in* the mind – except, however, for the data of inner sense.

There is no *general* reason to reject naive realism, i.e. the realism which claims that sense-data are identical with physical objects and that they subsist unchanged when not perceived. There are, however, special reasons – the traditional reasons – for believing that sense-data do not depend (causally) only on the object, but also on the subject. In other words, one is forced to believe that sense-data are entities which exist only when there is a certain relation between the subject and the physical object – a relation partly spatial and partly psychological. The physical object is a "thing in itself" which cannot be known directly. Even its existence is doubtful, since it depends on a rather precarious induction.

In saying that a physical object – i.e. that which subsists independently of the observer – is a "thing in itself", I do not mean to say that the categories do not apply to the physical object, nor that it is unknowable in an absolute sense. It can be described; the nature of its relations with other things can be known – in a word, one can know all that physics has to say on this subject. That it is not known directly, is, one might say, purely an accident; it may even be possible to know it directly, in some cases, where it is given. For the specific reasons which show the "subjectivity" of sense-data

are not demonstrable in all cases. However, it is easier and more plausible to suppose that the physical world and the sensible world have no point in common, and that there is merely a more or less exact correspondence between the two. In the case of dreams, for example, the correspondence is lacking. The particulars which we know are subjective in so far as their existence depends partly on us – from the causal point of view, of course, not from the logical point of view.

Universals, on the contrary, do not depend on us in any way. In the case of particulars, we have a causal dependence, but there could not be a causal dependence in the case of universals, since they do not exist in time. A logical dependence is equally impossible, since simple things do not logically depend on anything, and complex things logically depend only on their constituents. Therefore, universals are completely independent of the mind, as is everything else which exists, in the narrow sense. The laws of logic, for example, while they are customarily called "laws of thought", are just as objective, and depend as little on the mind, as the law of gravity. Abstract truths express relations which hold between universals; the mind can recognize these relations, but it cannot create them.

It is obviously impossible, on this occasion, to give even a general survey of the reasons which led me to adopt analytic realism. I will, however, say a few words on the contrary faults to be found in idealism, on the one hand, and in empiricism, on the other. Idealism, taken seriously, leads to the conclusion that what is called "*a priori* truth" is nothing but an illusion; one cannot help but believe it, but it is not what is the case. On the other hand, empiricism, by which we mean the theory that all evidence rests on sense-experience, leads to the conclusion that one cannot know anything except sense-data. Every law, every prediction, is not only rash, but completely devoid of probability. For to say: "this data makes this other event probable", is to employ a principle of probability which cannot be proved by evidence derived from the senses: one will always have to resort to a *general* principle. Therefore, neither idealism nor empiricism provides a theory of knowledge consistent with the facts. Moreover, both theories refute themselves: idealism, because it leads to a vicious regress, since, if existence depends on knowledge, the existence of knowledge depends on the knowledge of knowledge and so on; empiricism, because its principle that nothing is known except through the senses, cannot itself be known by the senses. It is therefore imperative to find a theory of knowledge which is neither idealist nor empiricist.

11

THE PHILOSOPHY OF LOGICAL ATOMISM

Extracts 1918

Russell's essay on the philosophy of logical atomism is well known, having being anthologised previously and appearing as the basis of an edited book (Pears 1972). Like the last selection, this is a verbatim account of Russell's lecturing. The series was presented at Gordon Square in London when Russell was in much need of money. As a whole, the essay is the largest record of Russell's lecturing. This selection includes the sections on facts, negative facts and the whole final part in which Russell considers some of the implications of his philosophy for metaphysics.

The basis of the philosophy of logical atomism, we have already seen. The difference between the previous selections in this Part, and the current, is that in the meantime Russell had met and been much influenced by Wittgenstein. Russell acknowledged, in a preamble, this essay's debt to his former pupil. Details of their meeting and relationship have been well chronicled elsewhere (e.g. Russell 1968: 98–101) so will receive no comment here. It appears that Russell had made a factualist interpretation of the early work of Wittgenstein. Whether this is the correct interpretation of Wittgenstein is always likely to be contentious. More recent philosophers have followed the Russell-line of interpretation (e.g. Skyrms 1981), so if Russell was misinterpreting Wittgenstein, he deserves credit for originating a philosophy that has proved enduring.

Earlier in the essay, Russell advocates an atomistic logic in which there are many separate things. The atoms are logical, by which Russell means that they are the last residue in analysis. In other words, when one analyses a complex, the atoms are that which are unanalysable and thereby at the end of process of analysis. They are logical atoms rather than physical atoms. Russell thinks that there are logical atoms, which are simple and unanalysable,

though, as he makes clear, this might not be the case and there could be infinite complexity.[a]

The world contains facts, as described in the first selection, below, which exist whether or not we think of them. We have beliefs, which are true or false by reference to facts. A fact is what makes a proposition true or false. There are different kinds of facts: particular (this is a man); general (all men are mortal), and so on. A proposition is the vehicle of truth and falsehood. It asserts something. It is, however, only a symbol, and properties of the symbol are not always properties of the thing. A symbol is something that "means" a fact.

A fact is evidently complex and is expressed by several words. The simplest imaginable fact is the possession of a quality by a particular thing, for example, "this is white". An atomic proposition contains a single verb. An atomic fact is the smallest thing that is, that exists, though it is a complex of particular and universal. David Armstrong has clarified the issue of how the simplest thing can also be complex (Armstrong 1997). The universals and particulars themselves have no independent existence in that there cannot be a particular that instantiates no universal and there cannot be a universal that is not instantiated in particulars. Universals and particulars do not, therefore, qualify as independent existences so the smallest units of existence must be the atomic facts.

You can understand a proposition through understanding the component words but you cannot understand the simplest component words themselves through analysis. The basic components are simples so their meaning cannot be given by any other terms. "Red", for instance, is a simple symbol; "this is red" is a complex one. The simple symbol can only be understood through acquaintance with the object; for example, by experiencing redness.

The next simplest fact is a relation between two facts, for example, "this is to the left of that". Next, there are tetradic relations, and so on. All facts assert some relation (including monadic relations) for some particular, where a particular is defined as the term of relations in atomic facts.

Molecular propositions are built out of propositions by using words such as "if", "and" and "or". Hence the proposition "today is Wednesday or today is Thursday" is a "molecule" built out of

a The importance of this claim is emphasised by Talmont-Kaminski (1998–9) against Pears' interpretation that the existence of simples is essential to logical atomism (Pears 1985: 1–2).

the two "atoms" "today is Wednesday" and "today is Thursday" by connecting them with the disjunction " … or … ". Though there are disjunctive propositions, there are no disjunctive facts, and there are similarly no molecular facts corresponding to other molecular propositions. Molecular facts are simply not needed so, by application of Occam's razor, a principle Russell frequently invokes, they are not included in the list of what exists. We can say that the truth of molecular propositions is a function of the truth of the constituent propositions. Hence, for any two propositions, *p* and *q*, if *p* and *q* are both false, the disjunction "*p* or *q*" will be false whatever the content of *p* and *q*. If at least one of *p* or *q* is true, their disjunction is true.

However, a complication arises in that Russell has to concede that there are, or there may be, negative facts. An opponent, Demos, had said that in asserting "not-*p*" one is asserting that another proposition *q* is true and *q* is incompatible with *p* (Demos 1917). Russell rightly objects to Demos's account on the grounds that: (*a*) it makes incompatibility fundamental and an objective fact, which is not so much simpler than allowing negative facts, (*b*) the incompatibility between two facts would be a molecular fact and the aim of the analysis was to find atomic facts, and (*c*) incompatibility is not between facts but between propositions.

Russell goes on to consider propositions and facts with more than one verb such as wishing, believing, willing. These do require a new form of fact. Where there are two or more verbs, as in "I believe Socrates is mortal", there are correspondingly two facts or constituents (for lack of a suitable term, Russell calls these constituents "verbs" also). The object of a belief must be a proposition, not a fact, because of the possibility of error. However, propositions are "nothing"; they do not inhabit the real world.

We then get accounts of general propositions and existence, singular existence and classes, which take up themes Russell has already developed in earlier work. On classes, Russell iterates the point that statements about classes can be translated into a form that does not mention classes at all. They are thus logical fictions that can be analysed into their logical atoms.

This takes us to the final lecture, "Excursus into metaphysics", which is reproduced in full below. The application of Occam's razor is one of the key themes of the lecture. We should attempt to invoke as few entities as possible, in our metaphysics, as this reduces the possibility of error. The reliability of Occam's principle is controversial, however. Is simplicity a reliable guide to how things really are? Could something that appears simple to us actually be complex in reality? It might be thought also that there is an anti-realist tendency in Russell's philosophy when he applies the razor, as evinced in the claim he makes that, like classes,

physical objects are logical fictions.[b] However, this lecture offers a key to understanding Russell's attitudes to metaphysics, realism and anti-realism. Although logical fictions can be substituted for supposed metaphysical entities, the metaphysical entity is neither asserted nor denied. Rather, Russell refuses to believe in anything for which there is no empirical evidence and metaphysics is not part of empirical knowledge. Epistemological commitments drive this metaphysical quietism. As Russell describes, his key epistemological commitment is that the immediate objects of acquaintance are our sense-data, otherwise known as sensations and qualia. That we see only representations or appearances of objects, rather than objects themselves, continues to be an attractive view for many thinkers. Since Russell's day, however, the view has been subject to a number of attacks (e.g. Austin 1962; Dennett 1991). If one accepts the sense-data theory, however, knowledge is limited to our experiences and what can legitimately be built out of them (Russell attacked this view himself, in the last paragraph of the previous paper, above). A distinction can be made between what is real and illusory, however, even within the sense-data theory. To say a thing is a real thing as opposed to a phantom is to say that it is linked up with other appearances: it has a whole series of correlations to my earlier and later "appearances" and to the appearances to others. The same applies to other persons, however: they are correlated systems of appearances to me. This suggests, finally, a neutral monism where the distinction between mental and physical is entirely a matter of arrangement and context.[c]

The most authoritative source of Wittgenstein's relevant views is his *Tractatus*, which was completed by 1918, but unknown to Russell at the time of his logical atomism lectures. Russell was influenced by Wittgenstein, at this point, only through his discussions with him before the First World War. The *Tractatus* was eventually published in 1921, largely through Russell's efforts (Wittgenstein 1921). There has been a recent renewed interest in Russell's logical atomism, due mainly to work by David Armstrong (Armstrong 1997). Armstrong arrived at his factualist ontology via Brian Skyrms's "Tractarian nominalism" (Skyrms 1981), which, as the title suggests, develops an interpretation of the *Tractatus*. Armstrong argues for a factualist ontology where a fact is a non-linguistic entity, as in Russell. He also calls such facts

b On the development of Russell's fictionalism or *logical constructionist* view of matter, see (Miah 1987).

c On Russell's neutral monism, see paper 19, below, and (Tully 1988; 1993–4).

states of affairs. Armstrong shows how far the factualist ontology can be developed in metaphysics, building out of it accounts of modality, causation, laws of nature, numbers and classes.

Significantly, Armstrong dissents from Russell's view that there are negative facts. Negative facts can be avoided if we allow instead second-order states of affairs. The negative truth "it is not that *a* is G" is made true not by a negative state of affairs but by the second-order state of affairs that the set S is all the states of affairs there are, and S does not include "that *a* is G". However, it might be thought that this is little more economical than Russell's solution, as Armstrong admits second-order states of affairs where Russell admitted only first-order facts, but including first-order negative facts. Further, it has been argued that the putatively objectionable features of Russell's negative facts are not so objectionable after all (Patterson 1996). *a being not-red* can be understood as a negative fact grasped by acquaintance with the property *not-red*.

Traditional views are that one can be acquainted with no such thing but it is arguable that one can be acquainted with the property *red* only if one can be acquainted with *not-red*. One understands what the properties are either both together or not at all. George Molnar, even more recently, has cast doubt on all the accounts of truthmakers for negative truths that are currently available, including Russell's solution and Armstrong's (Molnar 1999). Molnar argues that this issue has yet to be resolved satisfactorily, which is a major embarrassment of metaphysics.

ON FACTS

The first truism to which I wish to draw your attention – and I hope you will agree with me that these things that I call truisms are so obvious that it is almost laughable to mention them – is that the world contains facts, which are what they are whatever we may choose to think about them, and that there are also beliefs, which have reference to facts, and by reference to facts are either true or false. I will try first of all to give you a preliminary explanation of what I mean by a "fact". When I speak of a fact – I do not propose to attempt an exact definition, but an explanation, so that you will know what I am talking about – I mean the kind of thing that makes a proposition true or false. If I say "It is raining", what I say is true in a certain condition of weather and is false in other conditions of weather. The condition of weather that makes my statement true (or false as the case may be), is what I should call a "fact". If I say "Socrates is dead", my statement will be true owing to a certain physiological occurrence which

happened in Athens long ago. If I say, "Gravitation varies inversely as the square of the distance", my statement is rendered true by astronomical fact. If I say, "Two and two are four", it is arithmetical fact that makes my statement true. On the other hand, if I say "Socrates is alive", or "Gravitation varies directly as the distance", or "Two and two are five", the very same facts which made my previous statements true show that these new statements are false.

I want you to realize that when I speak of a fact I do not mean a particular existing thing, such as Socrates or the rain or the sun. Socrates himself does not render any statement true or false. You might be inclined to suppose that all by himself he would give truth to the statement "Socrates existed", but as a matter of fact that is a mistake. It is due to a confusion which I shall try to explain in the sixth lecture of this course [not included], when I come to deal with the notion of existence. Socrates[1] himself, or any particular thing just by itself, does not make any proposition true or false. "Socrates is dead" and "Socrates is alive" are both of them statements about Socrates. One is true and the other false. What I call a fact is the sort of thing that is expressed by a whole sentence, not by a single name like "Socrates". When a single word does come to express a fact, like "fire" or "wolf", it is always due to an unexpressed context, and the full expression of a fact will always involve a sentence. We express a fact, for example, when we say that a certain thing has a certain property, or that it has a certain relation to another thing; but the thing which has the property or the relation is not what I call a "fact".

It is important to observe that facts belong to the objective world. They are not created by our thoughts or beliefs except in special cases. That is one of the sort of things which I should set up as an obvious truism, but, of course, one is aware, the moment one has read any philosophy at all, how very much there is to be said before such a statement as that can become the kind of position that you want. The first thing I want to emphasise is that the outer world – the world, so to speak, which knowledge is aiming at knowing – is not completely described by a lot of "particulars", but that you must also take account of these things that I call facts, which are the sort of things that you express by a sentence, and that these, just as much as particular chairs and tables, are part of the real world. Except in psychology, most of our statements are not intended merely to express our condition of mind, though that is often all that they succeed in doing. They are intended to express facts, which (except when they are psychological facts)

1 I am here for the moment treating Socrates as a "particular". But we shall see shortly that this view requires modification.

will be about the outer world. There are such facts involved, equally when we speak truly and when we speak falsely. When we speak falsely it is an objective fact that makes what we say false, and it is an objective fact which makes what we say true when we speak truly.

There are a great many different kinds of facts, and we shall be concerned in later lectures with a certain amount of classification of facts. I will just point out a few kinds of facts to begin with, so that you may not imagine that facts are all very much alike. There are particular facts, such as "This is white"; then there are general facts, such as "All men are mortal". Of course, the distinction between particular and general facts is one of the most important. There again it would be a very great mistake to suppose that you could describe the world completely by means of particular facts alone. Suppose that you had succeeded in chronicling every single particular fact throughout the universe, and that there did not exist a single particular fact of any sort anywhere that you had not chronicled, you still would not have got a complete description of the universe unless you also added: "These that I have chronicled are all the particular facts there are." So you cannot hope to describe the world completely without having general facts as well as particular facts. Another distinction, which is perhaps a little more difficult to make, is between positive facts and negative facts, such as "Socrates was alive" – a positive fact – and "Socrates is not alive" – you might say a negative fact.[2] But the distinction is difficult to make precise. Then there are facts concerning particular things or particular qualities or relations, and, apart from them, the completely general facts of the sort that you have in logic, where there is no mention of any constituent whatever of the actual world, no mention of any particular thing or particular quality or particular relation, indeed strictly you may say no mention of anything. That is one of the characteristics of logical propositions, that they mention nothing. Such a proposition is: "If one class is part of another, a term which is a member of the one is also a member of the other." All those words that come in the statement of a pure logical proposition are words really belonging to syntax. They are words merely expressing form or connection, not mentioning any particular constituent of the proposition in which they occur. This is, of course, a thing that wants to be proved; I am not laying it down as self-evident. Then there are facts about the properties of single things; and facts about the relations between two things, three things, and so on; and any number of different classifications of some of the facts in the world, which are important for different purposes.

2 Negative facts are further discussed in a later lecture.

It is obvious that there is not a dualism of true and false facts; there are only just facts. It would be a mistake, of course, to say that all facts are true. That would be a mistake because true and false are correlatives, and you would only say of a thing that it was true if it was the sort of thing that *might* be false. A fact cannot be either true or false. That brings us on to the question of statements or propositions or judgments, all those things that do have the duality of truth and falsehood. For the purposes of logic, though not, I think, for the purposes of theory of knowledge, it is natural to concentrate upon the proposition as the thing which is going to be our typical vehicle on the duality of truth and falsehood. A proposition, one may say, is a sentence in the indicative, a sentence asserting something, not questioning or commanding or wishing. It may also be a sentence of that sort preceded by the word "that". For example, "That Socrates is alive", "That two and two are four", "That two and two are five", anything of that sort will be a proposition.

A proposition is just a symbol. It is a complex symbol in the sense that it has parts which are also symbols: a symbol may be defined as complex when it has parts that are symbols. In a sentence containing several words, the several words are each symbols, and the sentence composing them is therefore a complex symbol in that sense. There is a good deal of importance to philosophy in the theory of symbolism, a good deal more than at one time I thought. I think the importance is almost entirely negative, i.e. the importance lies in the fact that unless you are fairly self-conscious about symbols, unless you are fairly aware of the relation of the symbol to what it symbolizes, you will find yourself attributing to the thing properties which only belong to the symbol. That, of course, is especially likely in very abstract studies such as philosophical logic, because the subject matter that you are supposed to be thinking of is so exceedingly difficult and elusive that any person who has ever tried to think about it knows you do not think about it except perhaps once in six months for half a minute. The rest of the time you think about the symbols, because they are tangible, but the thing you are supposed to be thinking about is fearfully difficult and one does not often manage to think about it. The really good philosopher is the one who does once in six months think about it for a minute. Bad philosophers never do. That is why the theory of symbolism has a certain importance, because otherwise you are so certain to mistake the properties of the symbolism for the properties of the thing. It has other interesting sides to it too. There are different kinds of symbols, different kinds of relation between symbol and what is symbolized, and very important fallacies arise from not realizing this. The sort of contradictions about which I shall be speaking in connection with types in a later lecture all arise from mistakes in symbolism, from putting one sort of symbol in the place where another sort of symbol ought

to be. Some of the notions that have been thought absolutely fundamental in philosophy have arisen, I believe, entirely through mistakes as to symbolism – e.g. the notion of existence, or, if you like, reality. Those two words stand for a great deal that has been discussed in philosophy. There has been the theory about every proposition being really a description of reality as a whole and so on, and altogether these notions of reality and existence have played a very prominent part in philosophy. Now my own belief is that as they have occurred in philosophy, they have been entirely the outcome of a muddle about symbolism, and that when you have cleared up that muddle, you find that practically everything that has been said about existence is sheer and simple mistake, and that is all you can say about it. I shall go into that in a later lecture, but it is an example of the way in which symbolism is important.

Perhaps I ought to say a word or two about what I am understanding by symbolism, because I think some people think you only mean mathematical symbols when you talk about symbolism. I am using it in a sense to include all language of every sort and kind, so that every word is a symbol, and every sentence, and so forth. When I speak of a symbol I simply mean something that "means" something else, and as to what I mean by "meaning" I am not prepared to tell you. I will in the course of time enumerate a strictly infinite number of different things that "meaning" may mean, but I shall not consider that I have exhausted the discussion by doing that. I think that the notion of meaning is always more or less psychological, and that it is not possible to get a pure logical theory of meaning, nor therefore of symbolism. I think that it is of the very essence of the explanation of what you mean by a symbol to take account of such things as knowing, of cognitive relations, and probably also of association. At any rate I am pretty clear that the theory of symbolism and the use of symbolism is not a thing that can be explained in pure logic without taking account of the various cognitive relations that you may have to things.

As to what one means by "meaning", I will give a few illustrations. For instance, the word "Socrates", you will say, means a certain man; the word "mortal" means a certain quality; and the sentence "Socrates is mortal" means a certain fact. But these three sorts of meaning are entirely distinct, and you will get into the most hopeless contradictions if you think the word "meaning" has the same meaning in each of these three cases. It is very important not to suppose that there is just one thing which is meant by "meaning", and that therefore there is just one sort of relation of the symbol to what is symbolized. A name would be a proper symbol to use for a person; a sentence (or a proposition) is the proper symbol for a fact.

A belief or a statement has duality of truth and falsehood, which the fact does not have. A belief or a statement always involves a proposition. You say

that a man believes that so and so is the case. A man believes that Socrates is dead. What he believes is a proposition on the face of it, and for formal purposes it is convenient to take the proposition as the essential thing having the duality of truth and falsehood. It is very important to realize such things, for instance, as that propositions are not names for facts. It is quite obvious as soon as it is pointed out to you, but as a matter of fact I never had realized it until it was pointed out to me by a former pupil of mine, Wittgenstein. It is perfectly evident as soon as you think of it, that a proposition is not a name for a fact, from the mere circumstance that there are *two* propositions corresponding to each fact. Suppose it is a fact that Socrates is dead. You have two propositions: "Socrates is dead" and "Socrates is not dead". And those two propositions corresponding to the same fact, there is one fact in the world which makes one true and one false. That is not accidental, and illustrates how the relation of proposition to fact is a totally different one from the relation of name to the thing named. For each fact there are two propositions, one true and one false, and there is nothing in the nature of the symbol to show us which is the true one and which is the false one. If there were, you could ascertain the truth about the world by examining propositions without looking round you.

There are two different relations, as you see, that a proposition may have to a fact: the one the relation that you may call being true to the fact, and the other being false to the fact. Both are equally essentially logical relations which may subsist between the two, whereas in the case of a name, there is only one relation that it can have to what it names. A name can just name a particular, or, if it does not, it is not a name at all, it is a noise. It cannot be a name without having just that one particular relation of naming a certain thing, whereas a proposition does not cease to be a proposition if it is false. It has these two ways, of being true and being false, which together correspond to the property of being a name. Just as a word may be a name or be not a name but just a meaningless noise, so a phrase which is apparently a proposition may be either true or false, or may be meaningless, but the true and false belong together as against the meaningless. That shows, of course, that the formal logical characteristics of propositions are quite different from those of names, and that the relations they have to facts are quite different, and therefore propositions are not names for facts. You must not run away with the idea that you can name facts in any other way; you cannot. You cannot name them at all. You cannot properly name a fact. The only thing you can do is to assert it, or deny it, or desire it, or will it, or wish it, or question it, but all those are things involving the whole proposition. You can never put the sort of thing that makes a proposition to be true or false in the position of a logical subject. You can only have it there

as something to be asserted or denied or something of that sort, but not something to be named.

[...]

ON NEGATIVE FACTS

I do not see any reason to suppose that there is a complexity in the facts corresponding to these molecular propositions, because, as I was saying, the correspondence of a molecular proposition with facts is of a different sort from the correspondence of an atomic proposition with a fact. There is one special point that has to be gone into in connection with this, that is the question: Are there negative facts? Are there such facts as you might call the fact that "Socrates is not alive"? I have assumed in all that I have said hitherto that there are negative facts, that for example if you say "Socrates is alive", there is corresponding to that proposition in the real world the fact that Socrates is not alive. One has a certain repugnance to negative facts, the same sort of feeling that makes you wish not to have a fact "*p* or *q*" going about the world. You have a feeling that there are only positive facts, and that negative propositions have somehow or other got to be expressions of positive facts. When I was lecturing on this subject at Harvard I argued that there were negative facts, and it nearly produced a riot: the class would not hear of there being negative facts at all. I am still inclined to think that there are. However, one of the men to whom I was lecturing at Harvard, Mr. Demos, subsequently wrote an article in *Mind* to explain why there are no negative facts. It is in *Mind* for April 1917. I think he makes as good a case as can be made for the view that there are no negative facts. It is a difficult question. I really only ask that you should not dogmatize. I do not say positively that there are, but there may be.

There are certain things you can notice about negative propositions. Mr. Demos points out, first of all, that a negative proposition is not in any way dependent on a cognitive subject for its definition. To this I agree. Suppose you say, when I say "Socrates is not alive", I am merely expressing disbelief in the proposition that Socrates is alive. You have got to find something or other in the real world to make this disbelief true, and the only question is what. That is his first point.

His second is that a negative proposition must not be taken at its face value. You cannot, he says, regard the statement "Socrates is not alive" as being an expression of a fact in the same sort of direct way in which "Socrates is human" would be an expression of a fact. His argument for that is solely that he cannot believe that there are negative facts in the world.

He maintains that there cannot be in the real world such facts as "Socrates is not alive", taken, i.e. as simple facts, and that therefore you have got to find some explanation of negative propositions, some interpretation, and that they cannot be just as simple as positive propositions. I shall come back to that point, but on this I do not feel inclined to agree.

His third point I do not entirely agree with: that when the word "not" occurs, it cannot be taken as a qualification of the predicate. For instance, if you say that "This is not red", you might attempt to say that "not-red" is a predicate, but that of course won't do; in the first place because a great many propositions are not expressions of predicates; in the second place because the word "not" applies to the whole proposition. The proper expression would be "not: this is red"; the "not" applies to the whole proposition "this is red", and of course in many cases you can see that quite clearly. If you take a case I took in discussing descriptions: "The present king of France is not bald", then, if you take "not-bald" as a predicate, that would have to be judged false on the ground that there is not a present king of France. But it is clear that the proposition "The present king of France is bald" is a false proposition, and therefore the negative of that will have to be a true proposition, and that could not be the case if you take "not-bald" as a predicate, so that in all cases where "not" comes in, the "not" has to be taken to apply to the whole proposition. "Not-p" is the proper formula.

We have come now to the question, how are we really to interpret "not p", and the suggestion offered by Mr. Demos is that when we assert "not-p" we are really asserting that there is some proposition q which is true and is incompatible with p ("an opposite of p" is his phrase, but I think the meaning is the same). That is his suggested definition:

> "not-p" means "There is a proposition q which is true and is incompatible with p".

As, e.g., if I say "This chalk is not red", I shall be meaning to assert that there is some proposition, which in this case would be the proposition "This chalk is white", which is inconsistent with the proposition "It is red", and that you use these general negative forms because you do not happen to know what the actual proposition is that is true and is incompatible with p. Or, of course, you may possibly know what the actual proposition is, but you may be more interested in the fact that p is false than you are in the particular example which makes it false. As, for instance, you might be anxious to prove that some one is a liar, and you might be very much interested in the falsehood of some proposition which he had asserted. You might also be more interested in the general proposition than in the particular case,

so that if some one had asserted that that chalk was red, you might be more interested in the fact that it was not red than in the fact that it was white.

I find it very difficult to believe that theory of falsehood. You will observe that in the first place there is this objection, that it makes incompatibility fundamental and an objective fact, which is not so very much simpler than allowing negative facts. You have got to have here "That *p* is incompatible with *q*" in order to reduce "not" to incompatibility, because this has got to be the corresponding fact. It is perfectly clear, whatever may be the interpretation of "not", that there is *some* interpretation which will give you a fact. If I say "There is not a hippopotamus in this room", it is quite clear there is some way of interpreting that statement according to which there is a corresponding fact, and the fact cannot be merely that every part of this room is filled up with something that is not a hippopotamus. You would come back to the necessity for some kind or other of fact of the sort that we have been trying to avoid. We have been trying to avoid both negative facts and molecular facts, and all that this succeeds in doing is to substitute molecular facts for negative facts, and I do not consider that that is very successful as a means of avoiding paradox, especially when you consider this, that even if incompatibility is to be taken as a sort of fundamental expression of fact, incompatibility is not between facts but between propositions. If I say "*p* is incompatible with *q*", one at least of *p* and *q* has got to be false. It is clear that no two *facts* are incompatible. The incompatibility holds between the propositions, between the *p* and the *q*, and therefore if you are going to take incompatibility as a fundamental fact, you have got, in explaining negatives, to take as your fundamental fact something involving propositions as opposed to facts. It is quite clear that propositions are not what you might call "real". If you were making an inventory of the world, propositions would not come in. Facts would, beliefs, wishes, wills would, but propositions would not. They do not have being independently, so that this incompatibility of propositions taken as an ultimate fact of the real world will want a great deal of treatment, a lot of dressing up before it will do. Therefore as a simplification to avoid negative facts, I do not think it really is very successful. I think you will find that it is simpler to take negative facts as facts, to assume that "Socrates is not alive" is really an objective fact in the same sense in which "Socrates is human" is a fact. This theory of Mr. Demos's that I have been setting forth here is a development of the one one hits upon at once when one tries to get round negative facts, but for the reasons that I have given, I do not think it really answers to take things that way, and I think you will find that it is better to take negative facts as ultimate. Otherwise you will find it so difficult to say what it is that corresponds to a proposition. When, e.g. you have a false positive proposition, say "Socrates is alive", it is false because of a fact in the real world.

A thing cannot be false except because of a fact, so that you find it extremely difficult to say what exactly happens when you make a positive assertion that is false, unless you are going to admit negative facts. I think all those questions are difficult and there are arguments always to be adduced both ways, but on the whole I do incline to believe that there are negative facts and that there are not disjunctive facts. But the denial of disjunctive facts leads to certain difficulties which we shall have to consider in connection with general propositions in a later lecture.

[...]

VIII. EXCURSUS INTO METAPHYSICS: WHAT THERE IS

I come now to the last lecture of this course, and I propose briefly to point to a few of the morals that are to be gathered from what has gone before, in the way of suggesting the bearing of the doctrines that I have been advocating upon various problems of metaphysics. I have dwelt hitherto upon what one may call philosophical grammar, and I am afraid I have had to take you through a good many very dry and dusty regions in the course of that investigation, but I think the importance of philosophical grammar is very much greater than it is generally thought to be. I think that practically all traditional metaphysics is filled with mistakes due to bad grammar, and that almost all the traditional problems of metaphysics and traditional results – supposed results – of metaphysics are due to a failure to make the kind of distinctions in what we may call philosophical grammar with which we have been concerned in these previous lectures.

Take, as a very simple example, the philosophy of arithmetic. If you think that 1, 2, 3, and 4, and the rest of the numbers, are in any sense entities, if you think that there are objects, having those names, in the realm of being, you have at once a very considerable apparatus for your metaphysics to deal with, and you have offered to you a certain kind of analysis of arithmetical propositions. When you say, e.g. that 2 and 2 are 4, you suppose in that case that you are making a proposition of which the number 2 and the number 4 are constituents, and that has all sorts of consequences, all sorts of bearings upon your general metaphysical outlook. If there has been any truth in the doctrines that we have been considering, all numbers are what I call logical fictions. Numbers are classes of classes, and classes are logical fictions, so that numbers are, as it were, fictions at two removes, fictions of fictions. Therefore you do not have, as part of the ultimate constituents of your world, these queer entities that you are inclined to call numbers. The same applies in many other directions.

One purpose that has run through all that I have said, has been the justification of analysis, i.e. the justification of logical atomism, of the view that you can get down in theory, if not in practice, to ultimate simples, out of which the world is built, and that those simples have a kind of reality not belonging to anything else. Simples, as I tried to explain, are of an infinite number of sorts. There are particulars and qualities and relations of various orders, a whole hierarchy of different sorts of simples, but all of them, if we were right, have in their various ways some kind of reality that does not belong to anything else. The only other sort of object you come across in the world is what we call *facts*, and facts are the sort of things that are asserted or denied by propositions, and are not properly entities at all in the same sense in which their constituents are. That is shown in the fact that you cannot name them. You can only deny, or assert, or consider them, but you cannot name them because they are not there to be named, although in another sense it is true that you cannot know the world unless you know the facts that make up the truths of the world; but the knowing of facts is a different sort of thing from the knowing of simples.

Another purpose which runs through all that I have been saying is the purpose embodied in the maxim called Occam's Razor. That maxim comes in, in practice, in this way: take some science, say physics. You have there a given body of doctrine, a set of propositions expressed in symbols – I am including words among symbols – and you think that you have reason to believe that on the whole those propositions, rightly interpreted, are fairly true, but you do not know what is the actual meaning of the symbols that you are using. The meaning they have *in use* would have to be explained in some pragmatic way: they have a certain kind of practical or emotional significance to you which is a datum, but the logical significance is not a datum, but a thing to be sought, and you go through, if you are analyzing a science like physics, these propositions with a view to finding out what is the smallest empirical apparatus – or the smallest apparatus, not necessarily wholly empirical – out of which you can build up these propositions. What is the smallest number of simple undefined things at the start, and the smallest number of undemonstrated premisses, out of which you can define the things that need to be defined and prove the things that need to be proved? That problem, in any case that you like to take, is by no means a simple one, but on the contrary an extremely difficult one. It is one which requires a very great amount of logical technique; and the sort of thing that I have been talking about in these lectures is the preliminaries and first steps in that logical technique. You cannot possibly get at the solution of such a problem as I am talking about if you go at it in a straightforward fashion with just the ordinary acumen that one accumulates in the course of reading or in the study of traditional philosophy. You do need this apparatus

of symbolical logic that I have been talking about. (The description of the subject as symbolical logic is an inadequate one. I should like to describe it simply as logic, on the ground that nothing else really is logic, but that would sound so arrogant that I hesitate to do so.)

Let us consider further the example of physics for a moment. You find, if you read the works of physicists, that they reduce matter down to certain elements – atoms, ions, corpuscles, or what not. But in any case the sort of thing that you are aiming at in the physical analysis of matter is to get down to very little bits of matter that still are just like matter in the fact that they persist through time, and that they travel about in space. They have in fact all the ordinary everyday properties of physical matter, not the matter that one has in ordinary life – they do not taste or smell or appear to the naked eye – but they have the properties that you very soon get to when you travel toward physics from ordinary life. Things of that sort, I say, are not the ultimate constituents of matter in any metaphysical sense. Those things are all of them, as I think a very little reflection shows, logical fictions in the sense that I was speaking of. At least, when I say they are, I speak somewhat too dogmatically. It is possible that there may be all these things that the physicist talks about in actual reality, but it is impossible that we should ever have any reason whatsoever for supposing that there are. That is the situation that you arrive at generally in such analyses. You find that a certain thing which has been set up as a metaphysical entity can either be assumed dogmatically to be real, and then you will have no possible argument either for its reality or against its reality; or, instead of doing that, you can construct a logical fiction having the same formal properties, or rather having formally analogous formal properties to those of the supposed metaphysical entity and itself composed of empirically given things, and that logical fiction can be substituted for your supposed metaphysical entity and will fulfil all the scientific purposes that anybody can desire. With atoms and the rest it is so, with all the metaphysical entities whether of science or of metaphysics. By metaphysical entities I mean those things which are supposed to be part of the ultimate constituents of the world, but not to be the kind of thing that is ever empirically given – I do not say merely not being itself empirically given, but not being the *kind* of thing that is empirically given. In the case of matter, you can start from what is empirically given, what one sees and hears and smells and so forth, all the ordinary data of sense, or you can start with some definite ordinary object, say this desk, and you can ask yourselves, "What do I mean by saying that this desk that I am looking at now is the same as the one I was looking at a week ago?" The first simple ordinary answer would be that it *is* the same desk, it is actually identical, there is a perfect identity of substance, or whatever you like to call it. But when that apparently simple answer is

suggested, it is important to observe that you cannot have an empirical reason for such a view as that, and if you hold it, you hold it simply because you like it and for no other reason whatever. All that you really know is such facts as that what you see now, when you look at the desk, bears a very close similarity to what you saw a week ago when you looked at it. Rather more than that one fact of similarity I admit you know, or you may know. You might have paid some one to watch the desk continuously throughout the week, and might then have discovered that it was presenting appearances of the same sort all through that period, assuming that the light was kept on all through the night. In that way you could have established continuity. You have not in fact done so. You do not in fact know that that desk has gone on looking the same all the time, but we will assume that. Now the essential point is this: What is the empirical reason that makes you call a number of appearances, appearances of the same desk? What makes you say on successive occasions, I am seeing the same desk? The first thing to notice is this, that it does not matter what is the answer, so long as you have realized that the answer consists in something empirical and not in a recognized metaphysical identity of substance. There is something given in experience which makes you call it the same desk, and having once grasped that fact, you can go on and say, it is that something (whatever it is) that makes you call it the same desk which shall be *defined* as *constituting* it the same desk, and there shall be no assumption of a metaphysical substance which is identical throughout. It is a little easier to the untrained mind to conceive of an identity than it is to conceive of a system of correlated particulars, hung one to another by relations of similarity and continuous change and so on. That idea is apparently more complicated, but that is what is empirically given in the real world, and substance, in the sense of something which is continuously identical in the same desk, is not given to you. Therefore in all cases where you seem to have a continuous entity persisting through changes, what you have to do is to ask yourself what makes you consider the successive appearances as belonging to one thing. When you have found out what makes you take the view that they belong to the same thing, you will then see that that which has made you say so, is all that is *certainly* there in the way of unity. Anything that there may be over and above that, I shall recognize as something I cannot know. What I can know is that there are a certain series of appearances linked together, and the series of those appearances I shall define as being a desk. In that way the desk is reduced to being a logical fiction, because a series is a logical fiction. In that way all the ordinary objects of daily life are extruded from the world of what there is, and in their place as what there is you find a number of passing particulars of the kind that one is immediately conscious of in sense. I want to make clear that I am not *denying* the existence of anything; I am

only refusing to affirm it. I refuse to affirm the existence of anything for which there is no evidence, but I equally refuse to deny the existence of anything against which there is no evidence. Therefore I neither affirm nor deny it, but merely say, that is not in the realm of the knowable and is certainly not a part of physics; and physics, if it is to be interpreted, must be interpreted in terms of the sort of thing that can be empirical. If your atom is going to serve purposes in physics, as it undoubtedly does, your atom has got to turn out to be a construction, and your atom will in fact turn out to be a series of classes of particulars. The same process which one applies to physics, one will also apply elsewhere. The application to physics I explained briefly in my book on the *External World*, Chapters III and IV (Open Court Publishing Co., 1914).

I have talked so far about the unreality of the things we think real. I want to speak with equal emphasis about the reality of things we think unreal, such as phantoms and hallucinations. Phantoms and hallucinations, considered in themselves, are, as I explained in the preceding lectures, on exactly the same level as ordinary sense-data. They differ from ordinary sense-data only in the fact that they do not have the usual correlations with other things. In themselves they have the same reality as ordinary sense-data. They have the most complete and absolute and perfect reality that anything can have. They are part of the ultimate constituents of the world, just as the fleeting sense-data are. Speaking of the fleeting sense-data, I think it is very important to remove out of one's instincts any disposition to believe that the real is the permanent. There has been a metaphysical prejudice always that if a thing is really real, it has to last either forever or for a fairly decent length of time. That is to my mind an entire mistake. The things that are really real last a very short time. Again I am not denying that there *may* be things that last forever, or for thousands of years; I only say that those are not within our experience, and that the real things that we know by experience last for a very short time, one tenth or half a second, or whatever it may be. Phantoms and hallucinations are among those, among the ultimate constituents of the world. The things that we call real, like tables and chairs, are systems, series of classes of particulars, and the particulars are the real things, the particulars being sense-data when they happen to be given to you. A table or chair will be a series of classes of particulars, and therefore a logical fiction. Those particulars will be on the same level of reality as a hallucination or a phantom. I ought to explain in what sense a chair is a series of classes. A chair presents at each moment a number of different appearances. All the appearances that it is presenting at a given moment make up a certain class. All those sets of appearances vary from time to time. If I take a chair and smash it, it will present a whole set of different appearances from what it did before, and without going as far

as that, it will always be changing as the light changes, and so on. So you get a series in time of different sets of appearances, and that is what I mean by saying that a chair is a series of classes. That explanation is too crude, but I leave out the niceties, as that is not the actual topic I am dealing with. Now each single particular which is part of this whole system is linked up with the others in the system. Supposing, e.g. I take as my particular the appearance which that chair is presenting to me at this moment. That is linked up first of all with the appearance which the same chair is presenting to any one of you at the same moment, and with the appearance which it is going to present to me at later moments. There you get at once two journeys that you can take away from that particular, and that particular will be correlated in certain definite ways with the other particulars which also belong to that chair. That is what you mean by saying – or what you ought to mean by saying – that what I see before me is a real thing as opposed to a phantom. It means that it has a whole set of correlations of different kinds. It means that that particular, which is the appearance of the chair to me at this moment, is not isolated but is connected in a certain well-known familiar fashion with others, in the sort of way that makes it answer one's expectations. And so, when you go and buy a chair, you buy not only the appearance which it presents to you at that moment, but also those other appearances that it is going to present when it gets home. If it were a phantom chair, it would not present any appearances when it got home, and would not be the sort of thing you would want to buy. The sort one calls real is one of a whole correlated system, whereas the sort you call hallucinations are not. The respectable particulars in the world are all of them linked up with other particulars in respectable, conventional ways. Then sometimes you get a wild particular, like a merely visual chair that you cannot sit on, and say it is a phantom, a hallucination, you exhaust all the vocabulary of abuse upon it. That is what one means by calling it unreal, because "unreal" applied in that way is a term of abuse and never would be applied to a thing that *was* unreal because you would not be so angry with it.

I will pass on to some other illustrations. Take a person. What is it that makes you say, when you meet your friend Jones, "Why, this is Jones"? It is clearly not the persistence of a metaphysical entity inside Jones somewhere, because even if there be such an entity, it certainly is not what you see when you see Jones coming along the street; it certainly is something that you are not acquainted with, not an empirical datum. Therefore plainly there is something in the empirical appearances which he presents to you, something in their relations one to another, which enables you to collect all these together and say, "These are what I call the appearances of one person", and that something that makes you collect them together is not the persistence

of a metaphysical subject, because that, whether there be such a persistent subject or not, is certainly not a datum, and that which makes you say "Why, it is Jones" is a datum. Therefore Jones is not constituted as he is known by a sort of pin-point ego that is underlying his appearances, and you have got to find some correlations among the appearances which are of the sort that make you put all those appearances together and say, they are the appearances of one person. Those are different when it is other people and when it is yourself. When it is yourself, you have more to go by. You have not only what you look like, you have also your thoughts and memories and all your organic sensations, so that you have a much richer material and are therefore much less likely to be mistaken as to your own identity than as to some one else's. It happens, of course, that there are mistakes even as to one's own identity, in cases of multiple personality and so forth, but as a rule you will know that it is you because you have more to go by than other people have, and you would know it is you, not by a consciousness of the ego at all but by all sorts of things, by memory, by the way you feel and the way you look and a host of things. But all those are empirical data, and those enable you to say that the person to whom something happened yesterday was yourself. So you can collect a whole set of experiences into one string as all belonging to you, and similarly other people's experiences can be collected together as all belonging to them by relations that actually are observable and without assuming the existence of the persistent ego. It does not matter in the least to what we are concerned with, what exactly is the given empirical relation between two experiences that makes us say, "These are two experiences of the same person". It does not matter precisely what that relation is, because the logical formula for the construction of the person is the same whatever that relation may be, and because the mere fact that you can know that two experiences belong to the same person proves that there is such an empirical relation to be ascertained by analysis. Let us call the relation R. We shall say that when two experiences have to each other the relation R, then they are said to be experiences of the same person. That is a definition of what I mean by "experiences of the same person". We proceed here just in the same way as when we are defining numbers. We first define what is meant by saying that two classes "have the same number", and then define what a number is. The person who has a given experience x will be the class of all those experiences which are "experiences of the same person" as the one who experiences x. You can say that two events are co-personal when there is between them a certain relation R, namely that relation which makes us say that they are experiences of the same person. You can define the person who has a certain experience as being those experiences that are co-personal with that experience, and it will be better perhaps to take them as a series than as a class, because you

want to know which is the beginning of a man's life and which is the end. Therefore we shall say that a person is a certain series of experiences. We shall not deny that there may be a metaphysical ego. We shall merely say that it is a question that does not concern us in any way, because it is a matter about which we know nothing and can know nothing, and therefore it obviously cannot be a thing that comes into science in any way. What we know is this string of experiences that makes up a person, and that is put together by means of certain empirically given relations, such, e.g. as memory.

I will take another illustration, a kind of problem that our method is useful in helping to deal with. You all know the American theory of neutral monism, which derives really from William James and is also suggested in the work of Mach, but in a rather less developed form. The theory of neutral monism maintains that the distinction between the mental and the physical is entirely an affair of arrangement, that the actual material arranged is exactly the same in the case of the mental as it is in the case of the physical, but they differ merely in the fact that when you take a thing as belonging in the same context with certain other things, it will belong to psychology, while when you take it in a certain other context with other things, it will belong to physics, and the difference is as to what you consider to be its context, just the same sort of difference as there is between arranging the people in London alphabetically or geographically. So, according to William James, the actual material of the world can be arranged in two different ways, one of which gives you physics and other psychology. It is just like rows or columns: in an arrangement of rows and columns, you can take an item as either a member of a certain row or a member of a certain column; the item is the same in the two cases, but its context is different.

If you will allow me a little undue simplicity I can go on to say rather more about neutral monism, but you must understand that I am talking more simply than I ought to do because there is not time to put in all the shadings and qualifications. I was talking a moment ago about the appearances that a chair presents. If we take any one of these chairs, we can all look at it, and it presents a different appearance to each of us. Taken all together, taking all the different appearances that that chair is presenting to all of us at this moment, you get something that belongs to physics. So that, if one takes sense-data and arranges together all those sense-data that appear to different people at a given moment and are such as we should ordinarily say are appearances of the same physical object, then that class of sense-data will give you something that belongs to physics, namely, the chair at this moment. On the other hand, if instead of taking all the appearances that that chair presents to all of us at this moment, I take all the appearances

that the different chairs in this room present to me at this moment, I get quite another group of particulars. All the different appearances that different chairs present to me now will give you something belonging to psychology, because that will give you my experiences at the present moment. Broadly speaking, according to what one may take as an expansion of William James, that should be the definition of the difference between physics and psychology.

We commonly assume that there is a phenomenon which we call seeing the chair, but what I call my seeing the chair according to neutral monism is merely the existence of a certain particular, namely the particular which is the sense-datum of that chair at that moment. And I and the chair are both logical fictions, both being in fact a series of classes of particulars, of which one will be that particular which we call my seeing the chair. That actual appearance that the chair is presenting to me now is a member of me and a member of the chair, I and the chair being logical fictions. That will be at any rate a view that you can consider if you are engaged in vindicating neutral monism. There is no simple entity that you can point to and say: this entity is physical and not mental. According to William James and neutral monists that will not be the case with any simple entity that you may take. Any such entity will be a member of physical series and a member of mental series. Now I want to say that if you wish to test such a theory as that of neutral monism, if you wish to discover whether it is true or false, you cannot hope to get any distance with your problem unless you have at your fingers' end the theory of logic that I have been talking of. You never can tell otherwise what can be done with a given material, whether you can concoct out of a given material the sort of logical fictions that will have the properties you want in psychology and in physics. That sort of thing is by no means easy to decide. You can only decide it if you really have a very considerable technical facility in these matters. Having said that, I ought to proceed to tell you that I have discovered whether neutral monism is true or not, because otherwise you may not believe that logic is any use in the matter. But I do not profess to know whether it is true or not. I feel more and more inclined to think that it may be true. I feel more and more that the difficulties that occur in regard to it are all of the sort that may be solved by ingenuity. But nevertheless there *are* a number of difficulties; there are a number of problems, some of which I have spoken about in the course of these lectures. One is the question of belief and the other sorts of facts involving two verbs. If there are such facts as this, that, I think, may make neutral monism rather difficult, but as I was pointing out, there is the theory that one calls behaviourism, which belongs logically with neutral monism, and that theory would altogether dispense with those facts containing two verbs, and would therefore dispose of that argument against

neutral monism. There is, on the other hand, the argument from emphatic particulars, such as "this" and "now" and "here" and such words as that, which are not very easy to reconcile, to my mind, with the view which does not distinguish between a particular and experiencing that particular. But the argument about emphatic particulars is so delicate and so subtle that I cannot feel quite sure whether it is a valid one or not, and I think the longer one pursues philosophy, the more conscious one becomes how extremely often one has been taken in by fallacies, and the less willing one is to be quite sure that an argument is valid if there is anything about it that is at all subtle or elusive, at all difficult to grasp. That makes me a little cautious and doubtful about all these arguments, and therefore although I am quite sure that the question of the truth or falsehood of neutral monism is not to be solved except by these means, yet I do not profess to know whether neutral monism is true or is not. I am not without hopes of finding out in the course of time, but I do not profess to know yet.

As I said earlier in this lecture, one thing that our technique does, is to give us a means of constructing a given body of symbolic propositions with the minimum of apparatus, and every diminution in apparatus diminishes the risk of error. Suppose, e.g. that you have constructed your physics with a certain number of entities and a certain number of premisses; suppose you discover that by a little ingenuity you can dispense with half of those entities and half of those premisses, you clearly have diminished the risk of error, because if you had before 10 entities and 10 premisses, then the 5 you have now would be all right, but it is not true conversely that if the 5 you have now are all right, the 10 must have been. Therefore you diminish the risk of error with every diminution of entities and premisses. When I spoke about the desk and said I was not going to assume the existence of a persistent substance underlying its appearances, it is an example of the case in point. You have anyhow the successive appearances, and if you can get on without assuming the metaphysical and constant desk, you have a smaller risk of error than you had before. You would not necessarily have a smaller risk of error if you were tied down to *denying* the metaphysical desk. That is the advantage of Occam's Razor, that it diminishes your risk of error. Considered in that way you may say that the whole of our problem belongs rather to science than to philosophy. I think perhaps that is true, but I believe the only difference between science and philosophy is, that science is what you more or less know and philosophy is what you do not know. Philosophy is that part of science which at present people choose to have opinions about, but which they have no knowledge about. Therefore every advance in knowledge robs philosophy of some problems which formerly it had, and if there is any truth, if there is any value in the kind of procedure of mathematical logic, it will follow that a number of problems which had

belonged to philosophy will have ceased to belong to philosophy and will belong to science. And of course the moment they become soluble, they become to a large class of philosophical minds uninteresting, because to many of the people who like philosophy, the charm of it consists in the speculative freedom, in the fact that you can play with hypotheses. You can think out this or that which *may* be true, which is a very valuable exercise until you discover what *is* true; but when you discover what is true the whole fruitful play of fancy in that region is curtailed, and you will abandon that region and pass on. Just as there are families in America who from the time of the Pilgrim Fathers onward had always migrated westward, toward the backwoods, because they did not like civilized life, so the philosopher has an adventurous disposition and likes to dwell in the region where there are still uncertainties. It is true that the transferring of a region from philosophy into science will make it distasteful to a very important and useful type of mind. I think that is true of a good deal of the applications of mathematical logic in the directions that I have been indicating. It makes it dry, precise, methodical, and in that way robs it of a certain quality that it had when you could play with it more freely. I do not feel that it is my place to apologize for that, because if it is true, it is true. If it is not true, of course, I do owe you an apology; but if it is, it is not my fault, and therefore I do not feel I owe any apology for any sort of dryness or dullness in the world. I would say this too, that for those who have any taste for mathematics, for those who like symbolic constructions, that sort of world is a very delightful one, and if you do not find it otherwise attractive, all that is necessary to do is to acquire a taste for mathematics, and then you will have a very agreeable world, and with that conclusion I will bring this course of lectures to an end.

Part III

UNIVERSALS

Few discussions in philosophy have shown such longevity as the argument concerning the status of what are called "universals".

"The Problem of Universals"

12

ON THE RELATIONS OF UNIVERSALS AND PARTICULARS

1911

This paper was written for a specifically philosophical audience at the Aristotelian Society. It was Russell's Presidential Address to the Society, inaugurating the 1911–12 series. It was delivered in October 1911 and published the following year. Russell also wrote on the subject for a more popular audience, at around the same time, in his classic introduction *The Problems of Philosophy* (1912), and it might be useful to begin this Part with a summary of what he said there.

In Chapter 9, "The world of universals", we find Russell giving a Platonic answer to the problem. Particular things are white, just or square if they "participate in a common nature or essence". What it means for particular things to participate in the essence or form has always been obscure and Russell offers no further illumination here. The mark of a universal is that it is "anything which may be shared by many particulars". This is the standard criterion of universals but it was challenged by Ramsey, after he read Russell's 1911 Address (Ramsey 1925). Ramsey says that universals can just as well be regarded as particulars (as when I say "Justice is desirable") and particulars can be regarded as universals or properties of events (there is at least one event which "Socratises").

Russell then presented an argument that proves the existence of universals. It is a difficulty for any nominalism (including trope theory) that denies the existence of universals and is an argument that is cited still (e.g. Daly 1997). The argument runs as follows. Anyone wishing to invoke only particulars, and avoid universals, will at least have to concede that, for instance, white particulars resemble each other. But then the resemblance between this and that white thing must be the same resemblance that is being invoked between this and that red thing. Such pairs of objects resemble each other in different ways but just the single relation of resemblance is involved in each case. It seems, therefore, that at least the relation of resemblance must be a universal. The existence

of relations is thus certain, indeed more certain than the existence of properties.

If universals exist, where and when do they exist? "Nowhere and nowhen", replies Russell. They do not exist in space and time at all, in contrast with particulars. They do not have a mental existence either. They exist apart from our thoughts about them, and Russell again has an argument against the opposite view. Thoughts are particulars that would, therefore, rob the universal of its universality. Rather, universals have a being or subsistence, says Russell, harking back to the Meinongian ontology he had already rejected. In many ways, the world of *being* is superior to the fleeting world of particulars, but the latter has far greater importance to our lives.

Having argued that universals have no existence in space and time, Russell went on, in Chapter 10, "On our knowledge of universals", to make the remarkable claim that nevertheless we are acquainted with them in sense-data. This contradicts Plato's characterisation of universals being supra-sensible, which Russell seemed to cite approvingly in the previous chapter. Rather, it seems, Russell should have said that we can be acquainted only with that which partakes of the forms. However, in Chapter 10, he offers an account of how we attain knowledge of universals despite all our experience being of particular things. We ascend to the Platonic realm from our Earthly one, it seems, via a process of abstraction from many instances.

Exactly how this is done, is left unclear. Any knowledge of general terms abstracted from particular experiences appears highly problematic. How, we might wonder, can a process of abstraction begin without any prior knowledge of the universal in question? How, for instance, might I acquire knowledge of the universal *table*? Russell's answer has similarities with Locke's (1690: Book III, ch. 3). I see one table, then another, then another, and so on, and I take what is common to each one and reject what is particular. The concept I am left with picks out the universal. But why did I not take a table, a chair, a stool and a biscuit tin to arrive at the universal? The answer, presumably, is that I should not include the latter three items because they are not tables. But how am I to know that they are not tables prior to acquiring the concept of a *table*? In short, the problem is that without already having the concept in question, there are no rules for the process of abstraction to follow. The only successful abstraction would be one for which the concept in question was already available. This offers no account, therefore, of how the concept is acquired and, it follows, Russell has not shown how we are acquainted with universals.

Russell's case has an appearance of credibility only because he uses an example where there might be an exact resemblance

between all instances. He gives the example of white things, which could be all of exactly the same shade. Many universals apply to things where all the instances need not exactly resemble, however, and if different degrees of resemblance are possible, within a single universal, then the above argument applies: it is not clear, without already having the universal, which particulars are to be considered, and which not, in the process of abstraction.

The Aristotelian Society paper, presented in this selection, focuses on the basis of the distinction between universals and particulars. What are the criteria for being a particular and for being a universal? Russell concedes that some thinkers, with whom he has a degree of sympathy, allow that there is no fundamental division between universals and particulars. But on this issue, after due consideration, he believes they are wrong. Russell bases the distinction on spatial location. A particular can only be in one place at a time whereas as universal is capable of having instances at different places at the same time (though note Ramsey's challenge).

The distinction becomes immediately problematic, however. Particulars exist in one place. Relations, Russell accepts, exist in no place. But non-relation universals can exist in many places at the same time. Formerly, we have understood relations to be among the universals so why are they treated differently here? That relations are in no place seems to make good sense. John may be taller than Mary. John has a location and Mary has a location. But it makes no sense to say that the relation between them, … *is taller than* …, has a location. A relation being at a place makes such little sense that we might even call it a category mistake. But if we say this of relations, why not say the same of all universals, including those that are not relations? Russell does not say this and his basis for not doing so seems to be that properties have instances which are located. In contrast, even the instances of … *is taller than* … have no location.

There is a complication here of which Russell is aware. The universal, residing in its Platonic realm, as he says in *The Problems of Philosophy*, cannot be said to have a location; rather, only its instances have a location. Each instance has but one location, however, and this suggests that the instances of universals are themselves particulars. Russell's criterion appears inadequate as stated, therefore, even though there is some consideration of this particular challenge. A universal cannot be at more than one place and nor can its instances because they are particulars. Perhaps instead, consistent with his Platonism, Russell should have said that universals, like relations, can be at no place. Indeed, the account given in *The Problems of Philosophy* might have been a better way to make the distinction: particulars are in space and time while universals are not.

However, this view would be unacceptable to those who reject Platonic, "transcendent" realism about universals. Some believe that there are universals but that there is no transcendent realm. Universals are real but here, immanent to the world (Armstrong 1978). Hence, the "immanent realist" may be in a better position to agree with the claim of this paper, that "whiteness is in every place where there is a white thing", than Russell, the transcendent realist, is.

ON THE RELATIONS OF UNIVERSALS AND PARTICULARS

The purpose of the following paper[1] is to consider whether there is a fundamental division of the objects with which metaphysics is concerned into two classes, universals and particulars, or whether there is any method of overcoming this dualism. My own opinion is that the dualism is ultimate; on the other hand, many men with whom, in the main, I am in close agreement, hold that it is not ultimate. I do not feel the grounds in favour of its ultimate nature to be very conclusive, and in what follows I should lay stress rather on the distinctions and considerations introduced during the argument than on the conclusion at which the argument arrives.

It is impossible to begin our discussion with sharp definitions of universals and particulars, though we may hope to reach such definitions in the end. At the beginning, we can only roughly indicate the kind of facts that we wish to analyze and the kind of distinctions that we wish to examine. There are several cognate distinctions which produce confusion by intermingling, and which it is important to disentangle before advancing into the heart of our problem.

The first distinction that concerns us is the distinction between percepts and concepts, i.e. between objects of acts of perception and objects of acts of conception. If there is a distinction between particulars and universals, percepts will be among particulars, while concepts will be among universals. Opponents of universals, such as Berkeley and Hume, will maintain that concepts are derivable from percepts, as faint copies, or in some other way. Opponents of particulars will maintain that the apparent particularity

1 The thesis of the present paper is closely similar to that of Mr. Moore's paper "Identity", read before this Society in 1900–1. My chief reason for thinking that the question demands a fresh discussion is that the statement of the grounds for the thesis appears to require some examination of the nature of sensible space as opposed to physical space.

of percepts is illusory, and that, though the act of perception may differ from the act of conception, yet its objects differ only by their greater complexity, and are really composed of constituents which are, or might be, concepts.

But the distinction of percepts and concepts is too psychological for an ultimate metaphysical distinction. Percepts and concepts are respectively the relata of two different relations, perception and conception, and there is nothing in their definitions to show whether, or how, they differ. Moreover, the distinction of percepts and concepts, in itself, is incapable of being extended to entities which are not objects of cognitive acts. Hence we require some other distinction expressing the intrinsic difference which we seem to feel between percepts and concepts.

A cognate distinction, which effects part at least of what we want, is the distinction between things which exist in time and things which do not. In order to avoid any question as to whether time is relative or absolute, we may say that an entity *x* "exists in time" provided *x* is not itself a moment or part of time, and some such proposition as "*x* is before *y* or simultaneous with *y* or after *y*" is true of *x*. (It is not to be assumed that *before, simultaneous*, and *after* are mutually exclusive: if *x* has duration, they will not be so.) *Prima facie*, a percept exists in time, in the above sense, while a concept does not. The object of perception is simultaneous with the act of perception, while the object of conception seems indifferent to the time of conceiving and to all time. Thus, *prima facie* we have here the non-psychological distinction of which we were in search. But the same controversies will break out as in the case of percepts and concepts. The man who reduces concepts to percepts will say that nothing is really out of time, and that the appearance of this in the case of concepts is illusory. The man who reduces percepts to concepts may either, like most idealists, deny that anything is in time, or, like some realists, maintain that concepts can and do exist in time.

In addition to the above distinction as regards time, there is a distinction as regards space which, as we shall find, is very important in connection with our present question. Put as vaguely as possible, this is a distinction which divides entities into three classes: (*a*) those which are not in any place, (*b*) those which are in one place at one time, but never in more than one, (*c*) those which are in many places at once. To make this threefold division precise, we should have to discuss what we mean by a place, what we mean by "in", and how the different kinds of space – visual, tactile, physical – produce different forms of this threefold division. For the present I will merely illustrate what I mean by examples. Relations, obviously, do not exist anywhere in space. Our bodies, we think, exist in one place at a time, but not in more than one. General qualities, such as whiteness, on

the contrary, may be said to be in many places at once: we may say, in a sense, that whiteness is in every place where there is a white thing. This division of entities will be discussed later; for the present I merely wish to indicate that it requires examination.

In addition to the above psychological and metaphysical distinctions, there are two logical distinctions which are relevant in the present enquiry. In the first place, there is the distinction between relations and entities which are not relations. It has been customary for philosophers to ignore or reject relations, and speak as if all entities were either subjects or predicates. But this custom is on the decline, and I shall assume without further argument that there are such entities as relations. Philosophy has, so far as I know, no common name for all entities which are not relations. Among such entities are included not only all the things that would naturally be called particulars, but also all the universals that philosophers are in the habit of considering when they discuss the relation of particulars to universals, for universals are generally conceived as common properties of particulars, in fact, as predicates. For our purpose it is hardly worth while to invent a technical term *ad hoc*; I shall therefore speak of entities which are not relations simply as *non-relations*.

The second logical distinction which we require is one which may or may not be identical in extension with that between relations and non-relations, but is certainly not identical in intension. It may be expressed as the distinction between verbs and substantives, or, more correctly, between the objects denoted by verbs and the objects denoted by substantives.[2] (Since this more correct expression is long and cumbrous, I shall generally use the shorter phrase to mean the same thing. Thus, when I speak of verbs, I mean the objects denoted by verbs, and similarly for substantives.) The nature of this distinction emerges from the analysis of complexes. In most complexes, if not in all, a certain number of different entities are combined into a single entity by means of a relation, "*A*'s hatred for *B*", for example, is a complex in which *hatred* combines *A* and *B* into one whole; "*C*'s belief that *A* hates *B*" is a complex in which *belief* combines *A* and *B* and *C* and hatred into one whole, and so on. A relation is distinguished as dual, triple, quadruple, etc., or dyadic, triadic, tetradic, etc., according to the number of terms which it unites in the simplest complexes in which it occurs. Thus in the above examples, hatred is a dual relation and belief is a quadruple relation. The capacity for combining terms into a single complex is the

2 This is the distinction which I formerly spoke of as the distinction between *things* and *concepts*, but these terms no longer seem to me appropriate. Cf. *Principles of Mathematics*, §48.

defining characteristic of what I call *verbs*. The question now arises: Are there complexes which consist of a single term and a verb? "*A* exists" might serve as an example of what is possibly such a complex. It is the possibility that there may be complexes of this kind which makes it impossible to decide off-hand that verbs are the same as relations. There may be verbs which are philosophically as well as grammatically intransitive. Such verbs, if they exist, may be called *predicates*, and the propositions in which they are attributed may be called subject-predicate propositions.

If there are no such verbs as those whose possibility we have been considering, i.e. if all verbs are relations, it will follow that subject-predicate propositions, if there are any, will express a *relation* of subject to predicate. Such propositions will then be definable as those that involve a certain relation called *predication*. Even if there are subject-predicate propositions in which the predicate is the verb, there will still be equivalent propositions in which the predicate is related to the subject; thus "*A* exists", for example, will be equivalent to "*A* has existence". Hence the question whether predicates are verbs or not becomes unimportant. The more important question is whether there is a specific relation of predication, or whether what are grammatically subject-predicate propositions are really of many different kinds, no one of which has the characteristics one naturally associates with subject-predicate propositions. This question is one to which we shall return at a later stage.

The above logical distinctions are relevant to our enquiry because it is natural to regard particulars as entities which can only be subjects or terms of relations, and cannot be predicates or relations. A particular is naturally conceived as a *this* or something intrinsically analogous to a *this*; and such an entity seems incapable of being a predicate or a relation. A universal, on this view, will be anything that is a predicate or a relation. But if there is no specific relation of predication, so that there is no class of entities which can properly be called predicates, then the above method of distinguishing particulars and universals fails. The question whether philosophy must recognize two ultimately distinct kinds of entities, particulars and universals, turns, as we shall see more fully later on, on the question whether non-relations are of two kinds, subjects and predicates, or rather terms which can only be subjects and terms which may be either subjects or predicates. And this question turns on whether there is an ultimate simple asymmetrical relation which may be called predication, or whether all apparent subject-predicate propositions are to be analyzed into propositions of other forms, which do not require a radical difference of nature between the apparent subject and the apparent predicate.

The decision of the question whether there is a simple relation of predication ought perhaps to be possible by inspection, but for my part I am

unable to come to any decision in this way. I think, however, that it can be decided in favour of predication by the analysis of *things* and by considerations as to spatio-temporal diversity. This analysis and these considerations will also show the way in which our purely logical question is bound up with the other questions as to particulars and universals which I raised at the beginning of this paper.

The common-sense notion of things and their qualities is, I suppose, the source of the conception of subject and predicate, and the reason why language is so largely based on this conception. But the thing, like other common-sense notions, is a piece of half-hearted metaphysics, which neither gives crude data nor gives a tenable hypothesis as to a reality behind the data. A thing, of the every-day sort, is constituted by a bundle of sensible qualities belonging to various senses, but supposed all to co-exist in one continuous portion of space. But the common space which should contain both visual and tactile qualities is not the space of either visual or tactile perception: it is a constructed "real" space, belief in which has, I suppose, been generated by association. And in crude fact, the visual and tactile qualities of which I am sensible are not in a common space, but each in its own space. Hence if the thing is to be impartial as between sight and touch, it must cease to have the actual qualities of which we are sensible, and become their common cause or origin or whatever vaguer word can be found. Thus the road is opened to the metaphysical theories of science and to the metaphysical theories of philosophy: the thing may be a number of electric charges in rapid motion, or an idea in the mind of God, but it is certainly not what the senses perceive.

The argument against things is trite, and I need not labour it. I introduce it here only in order to illustrate a consequence which is sometimes overlooked. Realists who reject particulars are apt to regard a thing as reducible to a number of qualities co-existing in one place. But, apart from other objections to this view, it is doubtful whether the different qualities in question ever do co-exist in one place. If the qualities are sensible, the place must be in a sensible space; but this makes it necessary that the qualities should belong to only one sense, and it is not clear that genuinely different qualities belonging to one sense ever co-exist in a single place in a perceptual space. If, on the other hand, we consider what may be called "real" space, i.e. the inferred space containing the "real" objects which we suppose to be the causes of our perceptions, then we no longer know what is the nature of the qualities, if any, which exist in this "real" space, and it is natural to replace the bundle of qualities by a collection of pieces of matter having whatever characteristics the science of the moment may prescribe. Thus in any case the bundle of co-existing qualities in the same place is not an admissible substitute for the thing.

For our purposes, the "real" object by which science or philosophy replaces the thing is not important. We have rather to consider the relations of sensible objects in a single sensible space, say that of sight.

The theory of sensible qualities which dispenses with particulars will say, if the same shade of colour is found in two different places, that what exists is the shade of colour itself, and that what exists in the one place is identical with what exists in the other. The theory which admits particulars will say, on the contrary, that two numerically different *instances* of the shade of colour exist in the two places: in this view, the shade of colour itself is a universal and a predicate of both the instances, but the universal does not exist in space and time. Of the above two views, the first, which does not introduce particulars, dispenses altogether with predication as a fundamental relation: according to this view, when we say "this thing is white", the fundamental fact is that whiteness exists here. According to the other view, which admits particulars, what exists here is something of which whiteness is a predicate – not, as for common sense, the thing with many other qualities, but an instance of whiteness, a particular of which whiteness is the only predicate except shape and brightness and whatever else is necessarily connected with whiteness.

Of the above two theories, one admits only what would naturally be called universals, while the other admits both universals and particulars. Before examining them, it may be as well to examine and dismiss the theory which admits only particulars, and dispenses altogether with universals. This is the theory advocated by Berkeley and Hume in their polemic against "abstract ideas". Without tying ourselves down to their statements, let us see what can be made of this theory. The general name "white", in this view, is defined for a given person at a given moment by a particular patch of white which he sees or imagines; another patch is called white if it has exact likeness in colour to the standard patch. In order to avoid making the colour a universal, we have to suppose that "exact likeness" is a simple relation, not analyzable into community of predicates; moreover, it is not the general relation of likeness that we require, but a more special relation, that of colour-likeness, since two patches might be exactly alike in shape or size but different in colour. Thus, in order to make the theory of Berkeley and Hume workable, we must assume an ultimate relation of colour-likeness, which holds between two patches which would commonly be said to have the same colour. Now, *prima facie*, this relation of colour-likeness will itself be a universal or an "abstract idea", and thus we shall still have failed to avoid universals. But we may apply the same analysis to colour-likeness. We may take a standard particular case of colour-likeness, and say that anything else is to be called a colour-likeness if it is exactly like our standard case. It is obvious, however, that such a process leads to an endless

regress: we explain the likeness of two terms as consisting in the likeness which their likeness bears to the likeness of two other terms, and such a regress is plainly vicious. Likeness at least, therefore, must be admitted as a universal, and, having admitted one universal, we have no longer any reason to reject others. Thus the whole complicated theory, which had no motive except to avoid universals, falls to the ground. Whether or not there are particulars, there must be relations which are universals in the sense that (*a*) they are concepts, not percepts; (*b*) they do not exist in time; (*c*) they are verbs, not substantives.

It is true that the above argument does not prove that there are universal qualities as opposed to universal relations. On the contrary, it shows that universal qualities *can*, so far as logic can show, be replaced by exact likenesses of various kinds between particulars. This view has, so far as I know, nothing to recommend it beyond its logical possibility. But from the point of view of the problem whether there are particulars, it has no bearing on the argument. It is a view which is only possible if there are particulars, and it demands only an easy re-statement of subject-predicate propositions: instead of saying that an entity has such and such a predicate, we shall have to say that there are entities to which it has such and such a specific likeness. I shall therefore in future ignore this view, which in any case assumes our main thesis, namely the existence of particulars. To the grounds in favour of this thesis we must now return.

When we endeavoured to state the two theories as to sensible qualities, we had occasion to consider two white patches. On the view which denies particulars, whiteness itself exists in both patches: a numerically single entity, whiteness, exists in all places that are white. Nevertheless, we speak of *two* white patches, and it is obvious that, in some sense, the patches are two, not one. It is this spatial plurality which makes the difficulty of the theory that denies particulars.

Without attempting, as yet, to introduce all the necessary explanations and distinctions, we may state the argument for particulars roughly as follows. It is logically possible for two exactly similar patches of white, of the same size and shape, to exist simultaneously in different places. Now, whatever may be the exact meaning of "existing in different places", it is self-evident that, in such a case, there are two different patches of white. Their diversity might, if we adopted the theory of absolute position, be regarded as belonging, not to the white itself which exists in the two places, but to the complexes "whiteness in this place" and "whiteness in that place". This would derive their diversity from the diversity of this place and that place; and since places cannot be supposed to differ as to qualities, this would require that the places should be particulars. But if we reject absolute position, it will become impossible to distinguish the two patches as two,

unless each, instead of being the universal whiteness, is an *instance* of whiteness. It might be thought that the two might be distinguished by means of other qualities in the same place as the one but not in the same place as the other. This, however, presupposes that the two patches are already distinguished as numerically diverse, since otherwise what is in the same place as the one must be in the same place as the other. Thus the fact that it is logically possible for precisely similar things to co-exist in two different places, but that things in different places at the same time cannot be numerically identical, forces us to admit that it is particulars, i.e. *instances* of universals, that exist in places, and not universals themselves.

The above is the outline of our argument. But various points in it have to be examined and expanded before it can be considered conclusive. In the first place, it is not necessary to assert that there ever are two exactly similar existents. It is only necessary to perceive that our judgment that this and that are two different existents is not necessarily based on any difference of qualities, but may be based on difference of spatial position alone; and that difference of qualities, whether or not it always in fact accompanies numerical difference, is not logically necessary in order to insure numerical difference where there is difference of spatial position.

Again, it is not easy to state exactly what sort of spatial distribution in perceived space warrants us in asserting plurality. Before we can use space as an argument for particulars, we must be clear on this point. We are accustomed to concede that a thing cannot be in two places at once, but this common-sense maxim, unless very carefully stated, will lead us into inextricable difficulties. Our first business, therefore, is to find out how to state this maxim in an unobjectionable form.

In rational dynamics, where we are concerned with matter and "real" space, the maxim that nothing can be in two places at once is taken rigidly, and any matter occupying more than a point of space is regarded as at least theoretically divisible. Only what occupies a bare point is simple and single. This view is straightforward, and raises no difficulties as applied to "real" space.

But as applied to perceived space, such a view is quite inadmissible. The immediate object of (say) visual perception is always of finite extent. If we suppose it to be, like the matter corresponding to it in "real" space, composed of a collection of entities, one for each point which is not empty, we shall have to suppose two things, both of which seem incredible, namely: (1) that every immediate object of visual (or tactile) perception is infinitely complex; (2) that every such object is always composed of parts which are by their very nature imperceptible. It seems quite impossible that the immediate object of perception should have these properties. Hence we must suppose that an indivisible object of visual perception may occupy a finite

extent of visual space. In short, we must, in dividing any complex object of visual perception, reach, after a finite number of steps, a *minimum sensibile*, which contains no plurality although it is of finite extent. Visual space may, in a sense, be infinitely *divisible*, for, by attention alone, or by the microscope, the immediate object of perception can be changed in a way which introduces complexity where formerly there was simplicity; and to this process no clear limit can be set. But this is a process which substitutes a new immediate object in place of the old one, and the new object, though more subdivided than the old one, will still consist of only a finite number of parts. We must therefore admit that the space of perception is not infinitely divided, and does not consist of points, but is composed of a finite though constantly varying number of surfaces or volumes, continually breaking up or joining together according to the fluctuations of attention. If there is a "real" geometrical space corresponding to the space of perception, an infinite number of points in the geometrical space will have to correspond to a single simple entity in the perceived space.

It follows from this that, if we are to apply to the immediate objects of perception the maxim that a thing cannot be in two places at once, a "place" must not be taken to be a point, but must be taken to be the extent occupied by a single object of perception. A white sheet of paper, for example, may be seen as a single undivided object, or as an object consisting of two parts, an upper and a lower or a right hand and a left hand part, or again as an object consisting of four parts, and so on. If we on this account consider that, even when the sheet appeared as an undivided object, its upper and lower halves were in different places, then we shall have to say that the undivided object was in both these places at once. But it is better to say that, when the sheet appeared as an undivided object, this object occupied only one "place", though the place corresponded to what were afterwards two places. Thus a "place" may be defined as the space occupied by one undivided object of perception.

With this definition, the maxim that a thing cannot be in two places at once might seem to reduce to a tautology. But this maxim, though it may need rewording, will still have a substantial significance, to be derived from the consideration of spatial relations. It is obvious that perceived spatial relations cannot hold between points, but must hold between the parts of a single complex object of perception. When the sheet of paper is perceived as consisting of two halves, an upper and a lower, these two halves are combined into a complex whole by means of a spatial relation which holds directly between the two halves, not between supposed smaller subdivisions which in fact do not exist in the immediate object of perception. Perceived spatial relations, therefore, must have a certain roughness, not the neat smooth properties of geometrical relations between points. What, for example,

shall we say of distance? The distance between two simultaneously perceived objects will have to be defined by the perceived objects between them; in the case of two objects which touch, like the two halves of the sheet of paper, there is no distance between them. What remains definite is a certain order; by means of right and left, up and down, and so on, the parts of a complex object of perception acquire a spatial order, which is definite, though not subject to quite the same laws as geometrical order. The maxim that a thing cannot be in two places at once will then become the maxim that every spatial relation implies diversity of its terms, i.e. that nothing is to the right of itself, or above itself, and so on. In that case, given two white patches, one of which is to the right of the other, it will follow that there is not a single thing, whiteness, which is to the right of itself, but that there are two different things, instances of whiteness, of which one is to the right of the other. In this way our maxim will support the conclusion that there must be particulars as well as universals. But the above outline of an argument needs some amplification before it can be considered conclusive. Let us therefore examine, one by one, the steps of the argument.

Let us suppose, for the sake of definiteness, that within one field of vision we perceive two separated patches of white on a ground of black. It may then be taken as quite certain that the two patches are two and not one. The question is: Can we maintain that there are two if what exists in each is the universal whiteness?

If absolute space is admitted, we can of course say that it is the difference of place that makes the patches two; there is whiteness in this place, and whiteness in that place. From the point of view of our main problem, which is as to the existence of particulars, such a view would prove our thesis, since this place and that place would be or imply particulars constituting absolute space. But from the point of view of our immediate problem, which is concerned with plurality in perceived space, we may reject the above view on the ground that, whatever may be the case with "real" space, perceived space is certainly not absolute, i.e. absolute positions are not among objects of perception. Thus the whiteness here and the whiteness there cannot be distinguished as complexes of which this place and that place are respectively constituents.

Of course the whitenesses may be of different shapes, say one round and one square, and then they could be distinguished by their shapes. It will be observed that, with the view adopted above as to the nature of perceived space, it is perfectly possible for a simple object of perception to have a shape: the shape will be a quality like another. Since a simple object of perception may be of finite extent, there is no reason to suppose that a shape must imply spatial divisibility in the object of perception. Hence our two patches may be respectively round and square, and yet not be spatially

divisible. It is obvious, however, that this method of distinguishing the two patches is altogether inadequate. The two patches are just as easily distinguished if both are square or both are round. So long as we can see both at once, no degree of likeness between them causes the slightest difficulty in perceiving that there are two of them. Thus difference of shape, whether it exists or not, is not what makes the patches two entities instead of one.

It may be said that the two patches are distinguished by the difference in their relations to other things. For example, it may happen that a patch of red is to the right of one and to the left of the other. But this does not imply that the patches are two unless we know that one thing cannot be both to the right and to the left of another. This, it might be said, is obviously false. Suppose a surface of black with a small white space in the middle. Then the whole of the black may form only one simple object of perception, and would seem to be both to the right and to the left of the white space which it entirely surrounds. I think it would be more true to say, in this case, that the black is neither to the right nor to the left of the white. But right and left are complicated relations involving the body of the percipient. Let us take some other simpler relation, say that of surrounding, which the black surface has to the white patch in our example. Suppose we have another white patch, of exactly the same size and shape, entirely surrounded by red. Then, it may be said, the two patches of white are distinguished by difference of relation, since one is surrounded by black and the other by red. But if this ground of distinction is to be valid, we must know that it is impossible for one entity to be both wholly and immediately surrounded by black and wholly and immediately surrounded by red. I do not mean to deny that we do know this. But two things deserve notice – first, that it is not an analytic proposition; second, that it presupposes the numerical diversity of our two patches of white.

We are so accustomed to regarding such relations as "inside" and "outside" as incompatible that it is easy to suppose a *logical* incompatibility, although in fact the incompatibility is a characteristic of space, not a result of logic. I do not know what are the unanalyzable spatial relations of objects of perception, whether visual or tactile, but whatever they are they must have the kind of characteristics which are required in order to generate an order. They, or some of them, must be asymmetrical, i.e. such that they are incompatible with their converses: for example, supposing "inside" to be one of them, a thing which is inside another must not also be outside it. They, or some of them, must also be transitive, i.e. such that, for example, if x is inside y and y is inside z, then x is inside z – supposing, for the sake of illustration, "inside" to be among fundamental spatial relations. Probably some further properties will be required, but these at least are essential, in view of the fact that there is such a thing as spatial order. It follows that some at

least of the fundamental spatial relations must be such as no entity can have to itself. It is indeed self-evident that spatial relations fulfil these conditions. But these conditions are not demonstrable by purely logical considerations: they are synthetic properties of perceived spatial relations.

It is in virtue of these self-evident properties that the numerical diversity of the two patches of white is self-evident. They have the relation of being outside each other, and this requires that they should be two, not one. They may or may not have intrinsic differences – of shape, or size, or brightness, or any other quality – but whether they have or not they are two, and it is obviously logically possible that they should have no intrinsic differences whatever. It follows from this that the terms of spatial relations cannot be universals or collections of universals, but must be particulars capable of being exactly alike and yet numerically diverse.

It is very desirable, in such discussions as that on which we are at present engaged, to be able to talk of "places" and of things or qualities "occupying" places, without implying absolute position. It must be understood that, on the view which adopts relative position, a "place" is not a precise notion. But its usefulness arises as follows: – Suppose a set of objects, such as the walls and furniture of a room, to retain their spatial relations unchanged for a certain length of time, while a succession of other objects, say people who successively sit in a certain chair, have successively a given set of spatial relations to the relatively fixed objects. Then the people have, one after the other, a given set of properties, consisting in spatial relations to the walls and furniture. Whatever has this given set of properties at a given moment is said to "occupy" a certain place, the "place" itself being merely a fixed set of spatial relations to certain objects whose spatial relations to each other do not change appreciably during the time considered. Thus when we say that one thing can only be in one place at one time, we mean that it can only have one set of spatial relations to a given set of objects at one time.

It might be argued that, since we have admitted that a simple object of perception may be of finite extent, we have admitted that it may be in many places at once, and therefore may be outside itself. This, however, would be a misunderstanding. In perceived space, the finite extent occupied by a simple object of perception is not divided into many places. It is a single place occupied by a single thing. There are two different ways in which this place may "correspond" to many places. First, if there is such a thing as "real" space with geometrical properties, the one place in perceived space will correspond to an infinite number of points in "real" space, and the single entity which is the object of perception will correspond to many physical entities in "real" space. Secondly, there is a more or less partial correspondence between perceived space at one time and perceived space at another.

Suppose that we attend closely to our white patch, and meanwhile no other noticeable changes occur in the field of vision. Our white patch may, and often does, change as the result of attention – we may perceive differences of shade or other differentiations, or, without differences of quality, we may merely observe parts in it which make it complex and introduce diversity and spatial relations within it. We consider, naturally, that we are still looking at the same thing as before, and that what we see now was there all along. Thus we conclude that our apparently simple white patch was not really simple. But, in fact, the object of perception is not the same as it was before; what may be the same is the physical object supposed to correspond to the object of perception. This physical object is, of course, complex. And the perception which results from attention will be in one sense more correct than that which perceived a simple object, because, if attention reveals previously unnoticed differences, it may be assumed that there are corresponding differences in the "real" object which corresponds to the object of perception. Hence the perception resulting from attention gives more information about the "real" object than the other perception did: but the object of perception itself is no more and no less real in the one case than in the other – that is to say, in both cases it is an object which exists when perceived, but which there is no reason to believe existent except when it is perceived.

In perceived space, the spatial unit is not a point, but a simple object of perception or an ultimate constituent in a complex object of perception. This as the reason why, although two patches of white which are visibly separated from each other must be two, a continuous area of white may not be two. A continuous area, if not too large, may be a single object of perception not consisting of parts, which is impossible for two visibly separated areas. The spatial unit is variable, constantly changing its size, and subject to every fluctuation of attention, but it must occupy a continuous portion of perceived space, since otherwise it would be perceived as plural.

The argument as to numerical diversity which we have derived from perceived space may be reinforced by a similar argument as regards the contents of different minds. If two people are both believing that two and two are four, it is at least theoretically possible that the meanings they attach to the words *two* and *and* and *are* and *four* are the same, and that therefore, so far as the objects of their beliefs are concerned, there is nothing to distinguish the one from the other. Nevertheless, it seems plain that there are two entities, one the belief of the one man and the other the belief of the other. A particular belief is a complex of which something which we may call a subject is a constituent; in our case, it is the diversity of the subjects that produces the diversity of the beliefs. But these subjects cannot be mere

bundles of general qualities. Suppose one of our men is characterized by benevolence, stupidity, and love of puns. It would not be correct to say: "Benevolence, stupidity, and love of puns believe that two and two are four". Nor would this become correct by the addition of a larger number of general qualities. Moreover, however many qualities we add, it remains possible that the other subject may also have them; hence qualities cannot be what constitutes the diversity of the subjects. The only respect in which two different subjects *must* differ is in their relations to particulars: for example, each will have to the other relations which he does not have to himself. But it is not logically impossible that everything concerning one of the subjects and otherwise only concerning universals might be true of the other subject. Hence, even when differences in regard to such propositions occur, it is not these differences that constitute the diversity of the two subjects. The subjects, therefore, must be regarded as particulars, and as radically different from any collection of those general qualities which may be predicated of them.

It will be observed that, according to the general principles which must govern any correspondence of real things with objects of perception, any principle which introduces diversity among objects of perception must introduce a corresponding diversity among real things. I am not now concerned to argue as to what grounds exist for assuming a correspondence, but, if there is such a correspondence, it must be supposed that diversity in the effects – i.e. the perceived objects – implies diversity in the causes – i.e. the real objects. Hence if I perceive two objects in the field of vision, we must suppose that at least two real objects are concerned in causing my perception.

The essential characteristic of particulars, as they appear in perceived space, is that they cannot be in two places at once. But this is an unsatisfactory way of stating the matter, owing to the doubt as to what a "place" is. The more correct statement is that certain perceptible spatial relations imply diversity of their terms; for example, if x is above y, x and y must be different entities. So long, however, as it is understood that this is what is meant, no harm is done by the statement that a thing cannot be in two places at once.

We may now return to the question of particulars and universals with a better hope of being able to state precisely the nature of the opposition between them. It will be remembered that we began with three different oppositions: (1) that of percept and concept, (2) that of entities existing in time and entities not existing in time, (3) that of substantives and verbs. But in the course of our discussion a different opposition developed itself, namely (4) that between entities which can be in one place, but not in more than one, at a given time, and entities which either cannot be anywhere or

can be in several places at one time. What makes a particular patch of white particular, whereas whiteness is universal, is the fact that the particular patch cannot be in two places simultaneously, whereas the whiteness, if it exists at all, exists wherever there are white things. This opposition, as stated, might be held not to apply to thoughts. We might reply that a man's thoughts are in his head; but without going into this question, we may observe that there certainly is some relation between a man's thoughts and his head (or some part of it) which there is not between his thoughts and other things in space. We may extend our definition of particulars so as to cover this relation. We may say that a man's thought "belongs to" the place where his head is. We may then define a particular in our fourth sense as an entity which cannot be in or belong to more than one place at one time, and a universal as an entity which either cannot be in or belong to any place, or can be in or belong to many places at once. This opposition has certain affinities with the three earlier oppositions, which must be examined.

(1) Owing to the admission of particulars in our fourth sense, we can make an absolute division between percepts and concepts. The universal whiteness is a concept, whereas a particular white patch is a percept. If we had not admitted particulars in our fourth sense, percepts would have been identical with certain concepts.

(2) For the same reason, we are able to say that such general qualities as whiteness never exist in time, whereas the things that do exist in time are all particulars in our fourth sense. The converse, that all particulars in our fourth sense exist in time, holds in virtue of their definition. Hence the second and fourth senses of the opposition of particulars and universals are co-extensive.

(3) The third opposition, that of substantives and verbs, presents more difficulties, owing to the doubt whether predicates are verbs or not. In order to evade this doubt, we may substitute another opposition, which will be co-extensive with substantives and verbs if predicates are verbs, but not otherwise. This other opposition puts predicates and relations on one side, and everything else on the other. What is not a predicate or relation is, according to one traditional definition, a substance. It is true that, when substance was in vogue, it was supposed that a substance must be indestructible, and this quality will not belong to our substances. For example, what a man sees when he sees a flash of lightning is a substance in our sense. But the importance of indestructibility was metaphysical, not logical. As far as logical properties are concerned, our substances will be fairly analogous to traditional substances. Thus we have the opposition of substances on the one hand and predicates and relations on the other hand. The theory which rejects particulars allows entities commonly classed as predicates – e.g. white – to

exist; thus the distinction between substances and predicates is obliterated by this theory. Our theory, on the contrary, preserves the distinction. In the world we know, substances are identical with particulars in our fourth sense, and predicates and relations with universals.

It will be seen that, according to the theory which assumes particulars, there is a specific relation of subject to predicate, unless we adopt the view – considered above in connection with Berkeley and Hume – that common sensible qualities are really derivative from specific kinds of likeness. Assuming this view to be false, ordinary sensible qualities will be predicates of the particulars which are instances of them. The sensible qualities themselves do not exist in time in the same sense in which the instances do. Predication is a relation involving a fundamental logical difference between its two terms. Predicates may themselves have predicates, but the predicates of predicates will be radically different from the predicates of substances. The predicate, on this view, is never part of the subject, and thus no true subject-predicate proposition is analytic. Propositions of the form "All *A* is *B*" are not really subject-predicate propositions, but express relations of predicates; such propositions may be analytic, but the traditional confusion of them with true subject-predicate propositions has been a disgrace to formal logic.

The theory which rejects particulars, and assumes that, e.g. whiteness itself exists wherever (as common-sense would say) there are white things, dispenses altogether with predication as a fundamental relation. "This is white", which, on the other view, expresses a relation between a particular and whiteness, will, when particulars are rejected, really state that whiteness is one of the qualities in this place, or has certain spatial relations to certain other qualities. Thus the question whether predication is an ultimate simple relation may be taken as distinguishing the two theories; it is ultimate if there are particulars, but not otherwise. And if predication is an ultimate relation, the best definition of particulars is that they are entities which can only be subjects of predicates or terms of relations, i.e. that they are (in the logical sense) substances. This definition is preferable to one introducing space or time, because space and time are accidental characteristics of the world with which we happen to be acquainted, and therefore are destitute of the necessary universality belonging to purely logical categories.

We have thus a division of all entities into two classes: (1) particulars, which enter into complexes only as the subjects of predicates or the terms of relations, and, if they belong to the world of which we have experience, exist in time, and cannot occupy more than one place at one time in the space to which they belong; (2) universals, which can occur as predicates or relations in complexes, do not exist in time, and have no relation to one

place which they may not simultaneously have to another. The ground for regarding such a division as unavoidable is the self-evident fact that certain spatial relations imply diversity of their terms, together with the self-evident fact that it is logically possible for entities having such spatial relations to be wholly indistinguishable as to predicates.

13

THE PROBLEM OF UNIVERSALS

1946

Russell returned to philosophy later in life after a long period when political campaigning, education and popular writing were his main concerns. He had a spell of serious philosophy in the immediate post-war years that culminated in his last major philosophical work, *Human Knowledge: Its Scope and Limits* (1948). This piece on the problem of universals was prompted by an article by, and discussions with, Rupert Crawshay-Williams, who had read Russell on universals in *An Inquiry into Meaning and Truth* (1940).

Upon returning to the problem of universals, Russell saw it as "almost ripe for a definitive solution". In this paper, Russell does not quite offer that solution but merely expresses confidence that the techniques of modern logic can uncover the solution. The paper has a polemical style, which was fitting as it was published in a journal called *Polemic*. The ensuing discussion departs from the topic of universals in many places. The principle of individuation is discussed (see also paper 21) and Russell ventures the view that a "thing" is nothing more than a bundle of qualities. This is a familiar idea in metaphysics, which has received much discussion and criticism. It has the implication that two things cannot be exactly alike, in all their qualities (including relational qualities), as being so would entail that they were in reality one thing. To allow there to be two things exactly alike in all their qualities, would be to reject this as a principle of individuation. The possibility of a world containing just two objects, alike in all their non-relational properties and symmetrically arranged, hence alike in all their relational qualities as well, seems to be a counter-example but Russell does not quite consider it. He says instead that two things that appear alike can nevertheless be distinguished by their different positions in perceptual space. The possible world in the counter-example contains no perceivers, however. The example is controversial, but it seems that anyone who wants to pursue Russell's line has to provide some response to it.

On universals specifically, Aristotle is dismissed briefly. This is unfortunate in so far as immanent realism, which many take to be following Aristotle, is currently seen as a highly credible solution to the problem of universals. The Platonic view of transcendent universals was regarded as more credible in Russell's day but more of current metaphysics is sympathetic to Aristotle, seeing far greater subtlety and conceptual resources in his work than in Plato's.

Russell goes on to highlight the vagueness of universal words. A wide variety of colours may nevertheless all count as red. To define red, therefore, we have to start with a shade that is certainly red, and all other red things are red by reference to it. "Colour", in contrast, has no vagueness, in that nothing can be more or less a colour. However, things can have colour in virtue of having different specific colours. Why are these all collected together under the quality of colour? Russell proposes that they are bundled on the basis of similarity – a "similarity-chain" – where one can pass from any one specific colour to any other via a chain of closely similar colours.

Few discussions in philosophy have shown such longevity as the argument concerning the status of what are called "universals", which began with Plato, and is still in full swing at the present day. There has been, from time to time, a change in the way of formulating the problem and in the kind of analysis used in discussing it, but the essential core of the problem has changed little since Aristotle. For my part – although I may be accused of foolish optimism – I think it is almost ripe for definitive solution. Some things are already clear. The solution must be technical, and dependent upon considerations derived from modern logic; and it will have no bearing whatever upon the large problems of religious philosophy with which, ever since Plato, the discussion has been connected. Plato thought that his "ideas" afforded a proof of immortality; Leibniz deduced the existence of God from the "eternal truths"; Hume thought it essential to his empiricism to deny "abstract ideas". In these respects these men seem to me to have been mistaken. The acceptance or rejection of universals has, in my opinion, no bearing on religious belief or on the truth or falsehood of the empiricist philosophy. If the subject is to be discussed scientifically, these wider issues must be ignored during the discussion.

Plato's doctrine of ideas was open to objections of which he himself, in time, became aware, as appears from the argument about them in the *Parmenides*. The ideal Man was not a universal, but another particular, ethically and aesthetically superior to ordinary men, but logically and syntactically on a level with them. What needed to be considered was the adjective "human", which clearly does not denote any one man, whether earthly or heavenly. Socrates is human, and Socrates has two legs, but it is

nonsense to say "*human* has two legs". The universal, if there is such an entity, is what is signified by the word "human", and this, if it is anything, must be something quite different from any particular man.

It was Aristotle who first invented appropriate nomenclature for our problem, and who first stated it in a form approaching logical correctness. On the linguistic level, his definition of a "universal" is quite precise: "By the term 'universal' I mean that which is of such a nature as to be predicated of many subjects, by 'individual' that which is not thus predicated." Confining ourselves, for the moment, to the linguistic level, we make use of the familiar distinction of subject and predicate. In the sentence "Socrates is human", the word "Socrates" is the subject, and the word "human" is the predicate. The predicate "human" attaches also to many other subjects besides "Socrates"; and even if, in fact, the world contained only one man, the statement "this is human" would still be significant, though false, if "this" were a cat or a dog. Thus it is not necessary that a universal should be *truly* predicable of many things, but only that it should be *significantly* predicable. In any language having a subject-predicate syntax this distinction is clear.

But there remains a metaphysical question, irresistibly suggested by the linguistic distinction of subject and predicate. The word "Socrates" is a name, and most people think there is no difficulty as to what the word means: it means a certain person who lived a long time ago, and annoyed the Athenians by being too argumentative. If we had been Athenians of the fifth century B.C., his friends could have pointed him out to us and said "Look, that's Socrates". Thus it is clear that the word "Socrates" means something which is not a word, but a man. But how about the word "human"? This is a word, not an empty noise; it certainly has meaning. But is there *something* that it means? Clearly, if so, it is not something having a definite position in space and time, for it is present in every man, not only in Socrates. You cannot point to it as you can to Socrates. It is not born and it does not die. It has a name, but not a local habitation. Obviously it is something rather queer. The metaphysical problem of universals, as opposed to the linguistic problem, is to find something which can be meant by words that are predicates – or at least this is the form in which the problem first presented itself.

Aristotle proceeds to tell us that a universal is not a substance; it is a "such", not a "this", a kind, not an individual thing. "It seems impossible", he says, "that any universal term should be the name of a substance. For … the substance of each thing is that which is peculiar to it, which does not belong to anything else; but the universal is common, since that is called universal which is such as to belong to more than one thing". We may agree to this, but it does not take us very far; it is, in fact, purely negative.

The question of the metaphysical status of universals having been left open by Aristotle, the scholastics, while following him, were free to suggest various theories. Realists held that there are universals, and that predicates mean universals more or less as proper names means persons or things; nominalists held that universals are created by language, and that there is nothing outside language that is the meaning of a predicate. Between these two extremes various compromises were suggested. Both Aquinas and Occam, who in the main are nominalists, admit that God must have had universals in mind before creating things of different kinds. When, for example, He decided to create two human beings, He must have had the concept "human being", although there were as yet no instances from which to abstract it. But as far as *human* knowledge is concerned, they held that universals have no being except in the things that are instances of them.

In modern philosophy the old controversies, though with a somewhat changed vocabulary, continued unabated. Broadly speaking, realism prevailed on the Continent, nominalism in the British Isles. Berkeley and Hume denied not only universals in the world, but even abstract ideas in the mind. They held that I can have no idea "man", but only ideas of particular men, of a particular stature and complexion, and, in fact, with all the individuality belonging to the sensible appearance of an actual man. The word "man", they held, becomes general by being used as a sign, i.e. by our ignoring whatever features in our momentary idea of a man we do not regard as common to all men. This is the extreme antithesis to realism, and on the whole it represents the prevailing view among empiricist philosophers of the eighteenth and nineteenth centuries. They instance a geometrical demonstration, in which we draw (say) a particular triangle, but refrain from making use of any property which it does not share with all other triangles. In this way, they maintain, it is possible to suppose that the world contains nothing but impressions and particular ideas.

The controversy as to universals was connected with another medieval controversy, namely that as to the "principle of individuation". This question has, to my mind, been too little considered in modern times. I will try to state it in modern language, and to give my own theory as to the correct answer.

The question is this: Given two different "things", A and B, what is it that makes them two? Is it a difference of quality, or is it a mere difference of position in space and time? Some maintained that there could be, in different places, two things exactly alike in all their qualities, and yet certainly two; others maintained that if they were *exactly* alike they would be one, not two. The latter party held that an individual person or thing is completely defined when all his or its qualities are given. This was, in particular, the view of Leibniz. But their opponents maintained that individuality is

conferred by spatio-temporal position, not by qualities; it would be logically possible, they held, for identical twins to be *exactly* alike, except for occupying different parts of space. Broadly speaking, those who inclined towards empiricism made spatio-temporal position the "principle of individuation", while the *a priori* philosophers, who were also the champions of universals, considered that a thing could be completely defined by its qualities.

For my part, I hold that a "thing" is nothing but a bundle of qualities, and that, therefore, two different things cannot be exactly alike. But I hold this only because I regard position in space as defined by means of certain qualities not usually recognized as such. If I see two things at once, they cannot both be in the centre of my field of vision; if one is so, the other is to the right or left of it, and above or below it. In an immediate object of my visual perception, there is one of a series of qualities of more-or-less-right-or-leftness and of more-or-less-up-or-downness. It is in virtue of these qualities that my visual field has spatial characteristics. The space of physics is partly constructed, partly inferred, from the space of visual and other immediate objects of perception. I shall give my reasons for this view shortly.

The qualities of which, in my view, "things" are composed would count traditionally as universals. But in my view the names of qualities do not have the *syntactical* characteristics usually assigned to universal-words, though there are other words which do have these properties. We must, therefore, after considering qualities, resume the discussion of universals on the basis of the new syntactical analysis that I advocate. But throughout the earlier part of the following discussion I shall follow usage in regarding qualities as universals.

All the various disputants in this long history have conceived the problem much too simply, and all their various solutions can be seen to be inadequate if viewed without party spirit. To introduce the discussion, let us begin with a parable.

There was once a company of philosophers of various schools travelling in an out-of-the-way region on the Continent. They found an unpretentious inn and ordered dinner; the innkeeper promised them a joint of beef. But the joint, when it came, was unappetizing. One of the philosophers, a disciple of Hume and an experienced traveller, summoned mine host and said: "This is not beef, it is horse." He did not know that the innkeeper had seen better days, but had neglected his affairs and come down in the world through devotion to philosophy; he was therefore amazed when the innkeeper replied: "Sir, I am surprised to hear you saying something which you believe to be devoid of meaning. 'Beef' and 'horse', according to you, are only words, and do not denote anything in the non-linguistic world.

The dispute is therefore only about words. If you prefer the word 'horse', well and good; but I find the word 'beef' more profitable."

This reply set all the philosophers talking at once. "The innkeeper is right", said a disciple of Roscelin, "beef and 'horse' are only sounds uttered by human breath, and neither can denote this abominable piece of very tough meat." "Nonsense", retorted a Platonist, "this joint comes from an animal which, when alive, was a copy of the eternal horse in heaven, and not of the eternal ox." An Augustinian remarked: "'Beef' and 'horse' are ideas in the mind of God, and I am sure the divine idea of beef is something very different from this." There was only one point on which they were all agreed, and that was, that any person who sold such nasty stuff under the name of "beef" deserved to be prosecuted for fraud. At this the innkeeper, who knew the local magistrate to be no philosopher, became frightened and produced another joint, which gave universal satisfaction.

The sole point of this parable is that the question of universals is not merely one of words, but one which arises through the attempt to state facts.

The broad distinction between proper names and predicates is not quite what it appears to be in the traditional statements. It is this: that a proper name denotes something occupying a roughly continuous portion of space and time, while a predicate denotes something occupying discontinuous portions. Such words as "Socrates", "France", or "the sun", which we commonly think of as names of single entities, are really names of series of occurrences, each series having spatio-temporal continuity. Obviously France is made up of parts, each of which can have a proper name, as in fact many of them have. With human beings this is less common, but there are instances: "Abram" and "Abraham" are names for different parts of one man's life, and so are "Saul" and "Paul" or "Octavius" and "Augustus". It used to be thought that a mind or an atom was a single persistent entity, having different properties at different times, but psychology and physics have made this view untenable. We must now regard a mind as a series of thoughts, and an atom as a series of physical occurrences; in fact, persistence in time, just as much as extension in space, must, where things or persons are concerned, be taken as a sign of divisibility into parts.

The ultimate indivisible parts of which divisible things are composed are unknown to us, and have no names in any language. The question arises: what leads us to collect certain events into a group to which we give a proper name? Suppose I have a brother called "Tom"; how do I know, when I see him, that it is Tom? The spatio-temporal continuity which I believe him to possess is inferential; he might cease to exist when asleep alone. It is not by continuity, as a rule, that I recognize him, but by his qualities; his eyes and hair, the shape of his nose, etc., are familiar. But these

qualities by which I recognize him are (in traditional language) universals: red hair, blue eyes, big nose, and so on. The name "Tom" applies primarily to whatever has all these qualities. If I discover that Tom has a double, recognition becomes doubtful, and a Comedy of Errors becomes possible. Such errors can only be avoided by discovering *some* qualitative difference between Tom and his double; if they were *exactly* alike, we could never know which we were seeing. The word "Tom" thus denotes primarily a bundle of universals. Whether this analysis is ultimate, I leave for the present an open question.

When we consider the way in which children learn the meaning of words, it becomes obvious that our primary stock of words must all denote universals. An infant, as yet ignorant of language, cannot learn the meaning of words by means of verbal definitions. It is no use to tell him: "A dodecagon is a plane figure bounded by twelve straight lines" or "a cat is a domestic carnivorous quadruped of the genus *Felis*". The child learns the meaning of the word "cat" by frequently hearing the word at a time when a cat is occupying his attention. It is in the same way that he learns the meaning of the words "mother" and "father", which, though for him they are proper names of his own parents, are yet originally universals naming a series of similar appearances, just as much as "cat" is, as we saw in the case of "Tom". Certain similar occurrences become associated with the word "cat", certain others with the word "mother"; there is no important distinction between the two. To say that there is only one mother but there are many cats is only to say that mother occupies a continuous portion of space-time whereas the regions rightly designated "cat" do not. But from the point of view of logic and syntax this difference is not of fundamental importance.

The process by which we learn the meaning of a word without the help of a verbal definition is called "ostensive definition". It is obvious that it depends upon repetition and recognition. The child sees cats repeatedly, and at the same time hears the word "cat". At last, seeing another cat, the child recognizes it and says "cat". Now obviously this process of repetition and recognition is only possible in the case of qualities or bundles of qualities. If there be in a person a material identity of substance, this is certainly not obvious to the senses, and is not what enables us to recognize the person; we recognize the person by the similarity of his personal appearance to previous personal appearances, that is to say, by similarity of *qualities*. Whether there is or is not an identity of substance is a question for the metaphysician, and an otiose question, since no evidence either way is imaginable. What is given empirically, and what is involved in learning the use of words, is qualitative similarity between appearances on different occasions.

It follows that our primary vocabulary consists entirely of words denoting universals; words or phrases denoting particulars, if they are possible at all, must be learned words or phrases, involving in their use much analysis and scientific inference. "We cannot step twice into the same rivers", said Heraclitus; he was right if he was thinking of substantial identity. But if a river is defined qualitatively, which is the only way that makes recognition possible, then it can be the same river on two occasions. Substantial identity, and the concept of "substance" generally, are merely useless metaphysical lumber. The known world is composed of qualities with spatio-temporal and other relations. The qualities are what would formerly have been called "universals", and so far from our knowledge of them being in doubt, it is the easiest and most primitive knowledge that we possess, since it is the knowledge involved in the beginnings of our acquisition of language.

"But", you may say, "this theory will not do. There might be a *this* and a *that* which were exactly similar, and yet two. That is why we need the concept of 'substance' in addition to qualities." For my part, I entirely deny that there could be a *this* and a *that* that were exactly similar and yet two. Those who believe this possible have omitted to take account of the qualities by which space is defined. Consider first the space of your momentary visual field. If you keep your eyes fixed upon a stationary scene, some of the things you see are in the centre of your field of vision, some are above or below the centre, and some to right or left of it. The visual objects that compose your momentary field have different qualities according to their position relatively to the centre of the field; they are more or less up or more or less down, and more or less to the right or more or less to the left. No two parts of your visual field are identical as regards these qualities of up-or-down, right-or-left; only one part is a given amount up and a given amount to the right. When we take account of these qualities, it is impossible that two things seen simultaneously should be exactly alike.

Very similar considerations apply to the sense of touch. If a spider is crawling on you in the dark, you know what part of your body he is crawling on by the quality of the sensation of touch; no one would scratch his toe when the spider was on his hand. No two simultaneous sensations of touch have the same local quality, and therefore no two things that touch my body at the same time can be precisely similar in quality.

The conception of space is built up, by an elaborate and largely unconscious process, from such qualities as those that we have been considering in the case of sight and touch together with other factors. We thus arrive at the belief that a "thing" cannot be in two places at once – or rather, since every "thing" that we know of occupies a finite volume, that a "thing" must occupy a continuous portion of space and time, and not be dotted about as

redness is, for example. "Particulars", therefore, will be those bundles of qualities which include enough spatio-temporal qualities to insure that they occupy one continuous region. But in logical and syntactical status they will not differ importantly from such qualities as redness, which have no spatio-temporal continuity.

If the above theory is correct, many words that are commonly regarded as predicates cease to be such, and become substantives. Take such a statement as "this rose is red". Here the word "this" denotes certain spatial qualities, such as may be indicated by pointing, existing at the time when the sentence is pronounced. The word "rose" indicates a bundle of qualities, such as are given in the dictionary under the appropriate heading. The word "red" indicates a quality which is present in some roses and not in others. Thus what we are really saying is something like this: "At the present time, certain spatial qualities, and the qualities by which the word 'rose' is defined, and redness, all coexist, or are *compresent*." This is not a subject-predicate proposition. There are, however, subject-predicate propositions; they are such as "red is a colour", "C-sharp is a note". Thus, from the point of view of syntax, what were formerly universal, such as *red*, take the place of particulars as being denoted by substantives, while the place of universal is taken by such words as "colour" or "note", which are properties belonging to many qualities. The problem of universals is thus transferred to these properties of qualities.

Before considering this new problem, something must be said about redness, and its various shades, for it may be said that "redness", like "colour", is a class-name, denoting a number of different qualities. I am suggesting that the world is a pattern of qualities, but if this view is to be tenable the qualities must not have the vague generality indicated by such a word as "red", but the exactness of one particular shade. This, however, lands us in a fresh set of difficulties, since we know that we cannot distinguish one shade of colour from another that is only slightly different from it. This knowledge results from the fact that, if we see simultaneously three coloured patches, *A, B, C*, the colour of *A* may be indistinguishable from that of *B*, and that of *B* from that of *C*, and yet the colour of *A* may be distinguishable from that of *C*. It follows that there are differences of colour too minute to be perceptible. The same sort of thing happens with all perceptible qualities. Suppose *A* is a giant and *Z* is a dwarf: one could find a series of people *B, C, D*, … each of whom would not be noticeably taller or shorter than his neighbour on either side, but who, between them, would extend, in height, all the way from *A* to *Z*. None of our senses have the exactness that the mathematical physicist supposes to exist in nature.

Suppose, now, you have a certain length *A*, which is approximately a metre, and you wish to find out if there is reason to think it less than

a metre, or to think it more than a metre. No direct measurement shows it to be either more or less than a metre. But now you take another length *B*, which, so far as you can discover, is the same length as *A*, but you find that *B* is noticeably less than a metre. You then decide that *A* also must be less than a metre, though nearer to a metre than *B* is. If, on the other hand, you find that all the lengths indistinguishable from *A* are indistinguishable from a metre, you will say that, in the present state of the technique of measurement, *A* is to count as one metre long.

Exactly the same applies to the definition of a given shade of colour. Given two patches of colour *A* and *B*, they are to count as being of the same shade if every colour indistinguishable from *A* is indistinguishable from *B*, and vice versa.

By such devices we give greater exactness to our concepts than is to be derived from one or two unaided perceptions. The process is not quite conclusive in practice, because we cannot compare a given length or colour with *all* approximately similar lengths or colours. But it gives a method of approximation to exactness, and shows that there is no more difficulty about a given shade of colour than about a given length, such as a metre.

The words for qualities that we actually use have not even the degree of precision that is possible. We use words such as "red", which cover a wide range of different shades of colour. The definition of such a word as "red" has an essential vagueness, similar to that of such words as "bald" or "tall". Some shades of colour are certainly red, others are certainly not red, but at the boundary there are shades concerning which it is doubtful whether they are to be called "red" or not. The definition of "red" will have to start from a shade of colour that is certainly red, and say: "'Red' means a shade more or less like this." All the ordinary names of qualities, such as "hard" and "soft", "sweet" and "sour", have similar definitions, with the same essential vagueness.

It is otherwise with such a word as "colour", which seems to have no vagueness at all. We cannot imagine anything that is more or less a colour. Everything in the visual field has a colour, and nothing else has. The word "colour" is a true universal, and its instances are the various shades. The problem of universals is better dealt with in connection with such a word as "colour" than in connection with such a word as "red".

At first sight, it seems easier to take up a nominalist position in regard to such a word as "colour" than in regard to such a word as "red". It might seem that we could describe the world completely without using the word "colour". If we could state exactly what shade of colour exists at what point in everybody's visual field, we should have stated all the primary facts about colours, i.e. all the facts that have to do with particular regions of space-time; and in stating these facts we should not need the word "colour", but only words for the various particular shades of colour.

There are, however, other facts about colours besides those that have to do with particular places. Suppose *A, B, C* are three shades of colour which I see at once; suppose *A* is a greenish-yellow, *B* is green, and *C* is a greenish-blue. Then I can see, if I choose, that *A* is more like *B* than like *C*. This is not a dated fact, since it is true of the shades *A, B, C* whenever and wherever they occur; but although not dated, it is yet, in an important sense, a fact of perception. In Leibniz's sense it is an "eternal truth", and yet it is in some sense given empirically.

Now there seems to be something similar to say about "*A* is a colour", where *A* is some particular shade. The word "colour" is not defined by enumeration of the various colours – black and white, grey and brown, and the colours of the rainbow – as appears from the fact that if I see a shade of colour I never saw before, I shall have no doubt that it is a shade of colour. This shows that there must be some property common and peculiar to colours, by means of which the word "colour" should be defined. This property need not be a simple quality; it may be a rather complicated relational property. For example, the following would be a possible definition (not the *only* possible definition): starting from some datum *A*, which is to be a colour by definition, we shall say that if *B* closely resembles *A, B* is to be called a colour; and generally, if *M* is to be called a colour, and *N* closely resembles *M*, then *N* is to be called a colour. But this definition has a defect: it only gives the right extension if the *A* from which I started was in fact the sort of thing that we call a colour. This defect, however, is perhaps not fatal. Whether it is so, we must now consider.

The various qualities that are given in perception can be arranged in bundles, each bundle being what we may call a similarity-chain. A similarity-chain is defined as follows: starting from some datum *A*, it includes whatever is closely similar to *A*, or closely similar to something closely similar to *A*, or etc. That is to say, *Z* belongs to the similarity-chain of *A* if there is a finite number of terms *B, C, D*, ... *Y*, such that *A* is closely similar to *B, B* to *C*, ... and *Y* to *Z*. For "closely similar" we may, if we choose, substitute "indistinguishable". We can pass by imperceptible gradations from any one shade of colour to any other, from any sound to any other, from any sensation of touch to any other, from the taste of good roast beef to the taste of rotten egg – though in this last case there would be need of more culinary art than the world at present possesses. There is no such gradual transition from the data of one sense to those of another; between colours and sounds, for instance, there is an impassable gulf.

Having recognized the existence of such bundles, we can give them names. One bundle we call "colours", one "sounds", one "tastes", and so on. In this way we can avoid admitting such universal as "colour", and substitute the relation of close similarity or indistinguishability.

The introduction of "similarity" brings us to the consideration of relation-words. Hitherto we have been concerned with the meaning of what would commonly be considered adjectives; we have been led to the conclusion that some of them ought to be regarded as substantives, while others can be defined by means of the relation of similarity. This relation has a special and peculiar importance, but before concentrating on it something must be said about relations in general.

Let us begin with purely linguistic matters. There are certain words which are called "logical words"; such are "not", "or", "and", "if", "all", "some". These words are characterized by the fact that sentences in which they occur all presuppose the existence of simpler sentences in which they do not occur. A sentence containing no logical words I call an "atomic sentence". Thus "if it rains I shall get wet" is not atomic, but contains two atomic parts "it rains" and "I shall get wet". An atomic sentence contains no part which is a sentence, and also does not contain the word "all" or the word "some" or any equivalent.

Every atomic sentence, if fully expressed, contains two kinds of words, which are roughly nouns and verbs. An atomic sentence cannot contain more than one verb, but may, in theory, contain any number of nouns. "Caesar died" contains one noun and one verb; "*A* loves *B*" contains two nouns and one verb; "*A* is jealous of *B* on account of *C*" contains three nouns and one verb; "*A* exchanged *B* for *C* with *D*" contains four nouns and one verb. When an atomic sentence contains more than one noun, the verb is called a "relation-word", "dyadic" if there are two nouns, "triadic" if three, "tetradic" if four, and so on. I cannot think of any actual relation-word of higher order than the tetradic, but there is no reason except our intellectual limitations for the absence of such words.

A relation-word, if it occurs in an atomic sentence, must have the appropriate number of nouns. The sentence "*A* loves" may be interpreted so as to be significant, but is then not atomic; it may mean "*A* loves some one", i.e. "For some *B*, *A* loves *B*". This contains the word "some", and is therefore not atomic. Any truly atomic sentence containing the verb "loves" must contain two nouns, one denoting the lover, the other the beloved. In this respect relation-words differ from nouns, for every noun can occur in atomic sentences of all possible forms.

Prima facie, one would be inclined to say that nouns, when they are proper names, mean particulars, while adjectives, general nouns, and relation-words mean universals. It is impossible to construct a language that will say the things we want to say if our language consists of proper names alone. The nominalist is therefore confronted with the problem of showing how universal-words contribute to the significance of the sentences in which they occur. I do not wish to prejudge the question whether this problem is

soluble; I say only that the nominalist's position is not secure unless he can solve it.

But the problem is somewhat altered if we adopt the theory, advocated above, that what have usually been regarded as particulars are really bundles of qualities. In that case, the nouns with the simplest meaning are not such words as "Socrates", but words denoting qualities, such as "redness" or "sweetness". Two qualities may have a relation called "compresence"; this happens when they partially or completely overlap in space-time. "Place" is defined in terms of "compresence"; a "place" is a bundle of qualities all compresent with each other, and not all compresent with anything outside the bundle.

The question of universals, in its non-linguistic form now becomes: What must there be in the world besides qualities, in order that qualities may form the pattern in which we find them? Clearly we cannot describe our experience by merely mentioning qualities; we must also take account of relations, for it is in virtue of relations that qualities are arranged in patterns. Take the most obvious example: shape. The colours in a photograph are different from those in the original, but form the same pattern because relations are unchanged. Dark colours become black, light colours become white; if in the original a dark colour is above or to the right of a light colour, then in the photograph black is above or to the right of white. The relations up-and-down, right-and-left, are just as much part of what is perceived as are the actual colours; if this were not so, photographs would not seem as like their originals as they do. I do not see how we can escape the conclusion that "above" or "to the right of", in a suitable context, have as much factual meaning as any words can possibly have.

But this alone by no means decides the status of universals. It is clear that we need the word "above", but it is by no means clear that it means anything in isolation. In a sentence "*A* is above *B*", the word "above" contributes to the total significance, but it does not follow that the word by itself denotes an actual ingredient of the fact asserted. If I say "*A* is greater than *B*", the word "than" serves a purpose, but no one supposes that there is something in the world denoted by the word "than". Perhaps "above" is in the same case. Let us see whether we can make this hypothesis tenable.

Let us simplify as far as possible. I shall suppose that you are in the dark, but that there is an apparatus by means of which two bright dots, one directly above the other, can be made to appear straight in front of you; and I shall suppose that one of these dots is green, one red. You will then be able to say "green above red" or "red above green" as the case may be, and in so saying you are describing a fact. There are also the two simpler facts, "red" and "green"; but it is the complex relational fact with which we are concerned. Our purpose is, if possible, to deprive "above" of any significance except in a context.

There is an obvious difference of status between "above" and the two dots. Either dot could continue to exist without the other, indeed it would be logically possible for either to constitute the whole universe, but unless at least two dots exist there cannot be a fact requiring the word "above" for its statement. Thus "above" is dependent upon its terms in a way in which they are not dependent upon it. It would be natural to suggest that "above" is really a quality of the whole composed of the two dots. But this suggestion brings us up against the difficulty that there are *two* possible wholes, red-above-green and green-above-red. There must therefore be two different qualities that the whole may have, one when red is above, and the other when green is above. Is this theory defensible?

If we adopt it, we shall have to say something of the following sort: A complex whole is an entity whose existence is logically (not merely causally) dependent upon the existence of certain other entities, called its "parts". In our case, the two dots, which we will call *A* and *B*, constitute the whole visual field, and the whole which they form has various characteristics, expressed by "*A*-above-*B*", "*A*-brighter-than-*B*", perhaps "*A*-just-before-*B*". All these, in spite of the linguistic form, are to be regarded, not as facts about *A* and *B*, but about the wholes which they compose. If *A* is above *B*, then the whole composed of *A* and *B* is a whole of vertical order; so it is if *B* is above *A*. If *A* is brighter than *B*, the whole composed of *A* and *B* is a whole of relative luminosity; so it is if *B* is brighter than *A*. And so on. But in what way does the whole differ according as *A* or *B* is the higher? It seems impossible to state the difference without mentioning *A* or *B*, but if we have to mention them, as opposed to the whole, our suggested way of avoiding relations apparently fails. Whether it really fails I shall consider again presently.

There is yet another difficulty. Our theory requires that a whole should be of many different kinds – of as many kinds, in fact, as there are relations between its parts. The whole composed of *A* and *B* is a whole of temporal order if *A* is before or after *B*, a whole of vertical order if *A* is above or below *B*, a whole of horizontal order if *A* is to the right or left of *B*. Every whole will belong to many different kinds, corresponding to what would commonly be called the different relations between its parts. Each kind can be defined either by a common property, or by similarity; for instance, all the wholes consisting of one thing above another may be held to have a common property called "vertical order", or may be held to be a group of things that resemble each other in a certain respect. Neither suggestion enables us to dispense with universals, in the one case the property called "vertical order", in the other case "similarity-in-a-certain-respect".

On the whole, it seems so far (though our conclusion is not yet definitive) that we cannot escape from relations by substituting properties of the

wholes composed of their terms. That the Norman Conquest was earlier than the discovery of America is, it would seem, a fact about those two events, and not merely about the whole which they constitute. Temporal order is something which exists independently of language, and equally whether we notice it or not. We cannot describe the world or our own experience without relation-words, and clearly these words are not destitute of meaning. So far, I think we are on firm ground.

Although relation-words are necessary, it may be maintained that they are only necessary in their verbal or adjectival form, not as substantives. We must be able to say "*A* resembles *B*" or "*A* is similar to *B*", but perhaps we do not need the word "similarity". When I say that we do not need it, I mean that everything we know of can be described without using this word. Very often, it is used in such a way that the word "similar" can replace it. For instance, one might say "similarity is the rule among identical twins", and this can be replaced by "in most cases, a pair of identical twins are similar to each other". But when we ask metaphysical questions, such as "is there, apart from language, a relation of similarity?" we cannot translate our question into a form using "similar" instead of "similarity". It is suggested that such untranslatable uses of the word "similarity" are meaningless.

This characteristic, which is assigned to relation-words, is called that of being "syncategorematic". This means that a relation-word can never be used significantly except when terms are supplied to the relation. I can say "*A* is like *B*" or "some men are very like some other men", but if I say anything containing the word "likeness", my sentence will be meaningless unless there is a way of restating it so as to contain instead the word "like". If this is correct, such a question as "is likeness a universal?" is meaningless. I incline to hope that this is true, but it does not help us to answer the question: "What feature of the world is it that makes it necessary to use relation-words in describing it?" Following our rule of language, we must not answer by saying "there are relations", but we may answer by saying "things are related". Even this, however, will need careful interpretation; it cannot mean "between the two things *A* and *B* some relation holds", for the phrase "some relation", on the face of it, requires us to treat a relation as a substantive.

Let us revert to our visual field consisting of a red light *A* vertically above a green light *B*. This visual field is a complex whole, which we will call *W*. With the same ingredients *A* and *B* we could form another whole *W′* in which B would be above *A*. The use of the word "above" can be learnt by the process of ostensive definition, just as the use of the word "cat" is learnt: a parent can point first to *A* and then to *B*, and say: "This is *above* that." In time the child learns the correct use of the word by means of many instances. The difference between "*A* is above *B*" and "*B* is below *A*" is

subjective, in the order of attention: in the former case we attend first to *A*, then let our eyes sink to *B*, while in the latter case we attend first to *B* and then let our eyes rise to *A*. The fact asserted is independent of this order of attention, and is the same in either case. The use of the two words "above" and "below" is therefore unnecessary; either alone suffices to describe the facts.

We shall have to say that a relational fact has a *sense* or *direction*. In "*A* is above *B*" the sense is from *A* to *B*; in "*B* is above *A*" the sense is from *B* to *A*. This will enable us to deal with asymmetry.

If we wish to adhere to the view that "above" is "syncategorematic", we shall have to say that, given two visual dots *A* and *B* which form part (or the whole) of one visual field, two complex wholes are logically possible, one of which we describe as "*A*-above-*B*" and the other as "*B*-above-*A*". We shall have to say that all complex wholes correctly described by the use of the word "above" have either a common quality or a specific kind of resemblance. If we take the latter view, we have the word "similar" to deal with. We shall have to say that the complex *A*-above-*B* is similar in a certain respect to *C*-above-*D*; this is again a relation, namely "similar-in-a-certain-respect". This seems to lead to an endless regress. The suggestion of a common quality is perhaps slightly less objectionable, but, as we saw, it also has its difficulties.

I confess that I do not know the answer to these problems, though I am convinced that only clear thinking is needed in order to find it. Until it is found, some parts of the problem of universals must remain unsolved.

It remains to add a few general remarks. In the first place, the problem is a technical one as to the interpretation of language. It is almost wholly independent of our views as to the general nature of the universe; in the above discussion, there has been not a word which either assumed or denied, directly or by implication, solipsism, idealism, materialism, or any other such doctrine. The describing of our own experiences raises the whole of the problem, quite regardless of the question as to what else there may be in the world.

In the second place, unadulterated nominalism suffers from having overlooked the fact that a word is not a single entity, but a class. There are many cats, but there are also many instances of the word "cat". If there is no such universal as felinity, it will follow, by exactly the same arguments, that there is no such thing as the word "cat", but only instances of the word, spoken or heard, written or read. The relation of a word to its meaning is that of one set of particulars to another set of particulars: seeing before you an instance of the species cat, you give utterance to an instance of the species "cat" (the word). Thus there is exactly the same generality about the word as about what it means. This makes it impossible to suppose that there

are universals in language, but not in things. If the species *cat* is a logical figment, then so is the word "cat".

In the third place, if the problem is to be solved, it must not be approached with bias or party spirit. We must not let ourselves think, while we are investigating it, such thoughts as: "If I admit universals I shall be helping superstition and obscurantism", or: "If I deny universals I shall be doing my bit towards depriving people of a source of dignified comfort, and perhaps paving the way to social anarchy." Such considerations lead to bad philosophy, just as hard cases lead to bad law. I think myself that when once the problem has been brought, as it should be, into the region of detailed technicality, it becomes evident that no tenable solution has any social consequences whatever. It is true that there have been doctrines – of which Plato's is the most notable – which appeared to have important social consequences. But all such doctrines are easily shown to be false, and of the remainder, which cannot at present be proved or disproved, it makes no political or theological difference which we adopt.

What does make a difference, however, and a very profound difference, is the acquisition of a habit of truthfulness and of a sense of relevance, which can be acquired by a resolute refusal to confuse one kind of question with another, and by the use of what might be called a mental microscope. Physicists are allowed to investigate minute phenomena, which to the layman have no obvious importance; they have acquired this license because it has been found that minute phenomena often yield the key to the general nature of many physical processes. But philosophers, having fewer results to show, are still thought ill of if they concern themselves with points that to the outsider look trivial, and if they are not at every moment telling us how we should live or whether the soul is immortal or something of the same obvious emotion interest. It is only by temporarily forgetting such preoccupations, and acquiring the patience and devotion to detail to which experimental science has accustomed us, that philosophy can hope to achieve solid results which shall be genuine contributions to the sum of human knowledge.

Part IV

CAUSALITY AND LAWS

The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm.

"On The Notion of Cause"

14

ON THE NOTION OF CAUSE

1913

This paper was written for the Aristotelian Society as Russell's second Presidential Address. A more popular version, that covers much of the same ground, was published as the final chapter of Russell's book *Our Knowledge of the External World* (1914). The version here contains the detail and precision required for a purely philosophical audience.

Russell came in an empiricist tradition that started with Locke, Berkeley and Hume. Russell places himself in this tradition by beginning with a rejection of modal truths; that is truths of necessity and possibility. He allows that propositions can have a truth value only, they cannot have a modal value as well. Realists about causation allow causal truths to be modal truths. That *B* follows *A* might be not just true, but also necessarily so.

Hume is famously associated with a deflationist view of causation in which, he claims, we have no good reason to believe that there is such a thing. Russell does not rely explicitly on this argument. His starting point, rather, is that science itself manages to say all it needs to say without any notion of cause. Nevertheless, the 'law of causality' survives as a relic of a bygone age. The relic survives because we mistakenly believe it to do no harm. In fact, Russell contends, a notion of causality does do harm and the latter sections of the paper show some of the confusions connected with the belief in this erroneous notion.

Russell also produces some relatively novel criticisms of the standard Hume-derived notion of a cause and its corollaries. Russell denies, for instance, the asymmetry between cause and effect and this makes it hard to sustain the thesis that the cause determines the effect and not vice versa. Indeed, Russell says that it is a mistake to think of a cause as "operating" on an effect. In place of these dismissed notions, Russell notes that science utilises, instead, formulae. These are equations in which nothing can be identified as cause and nothing as effect; for example, F = ma.

Such equations have little in common with the law of causality so they cannot be what the law of causality is really about. In more modern terms, causality is not "reducible" to such formulae – by being a more accurate way of talking about it. Rather, the formulae "eliminate" causality altogether.

Two later discussions are worthy of note. The first concerns the alleged uniformity of nature principle. Russell takes it that this has only an inductive justification whereas in the contemporary Dretske–Tooley–Armstrong theory of laws (Dretske 1977; Tooley 1977; Armstrong 1983) it is implied that the uniformity of nature principle is in some sense necessary. Second, in a passage that resembles recent claims made by Nancy Cartwright, Russell says that where one event *a* is said to cause another event *b*, we are treating these as if in a practically isolated system (Cartwright 1999: Part I). We are abstracting them away from all the many possible interfering factors that almost always influence the process in reality.

There is a pay-off for the issue of free will in Russell's substitution of symmetrical differential equations for the traditional asymmetrical concept of causality. There are at least two new points to be made. First, where there have previously been discussions of whether matter causes mind (carrying a threat to free will) or mind causes matter (carrying a threat to the universality of physical laws), there is no such issue of psycho-physical determinism under Russell's notion. The two will co-vary – where there is a change in mind there is simultaneously a change in matter – without justifying the claim that one determines the other. Second, the future is fixed but so is the past (Russell rejects objective chance: the only indeterminacy is ignorance). The future's fixedness does not mean that we will not act on our desires in the future, however, just as the fixedness of the past does not mean that we were not acting on our desires in the past. Hence we have a development in Russell's solution to the free will problem presented in papers 1 and 3.

ON THE NOTION OF CAUSE

In the following paper, I wish, first, to maintain that the word "cause" is so inextricably bound up with misleading associations as to make its complete extrusion from the philosophical vocabulary desirable; secondly, to inquire what principle, if any, is employed in science in place of the supposed "law of causality" which philosophers imagine to be employed; thirdly, to exhibit

certain confusions, especially in regard to teleology and determinism, which appear to me to be connected with erroneous notions as to causality.

All philosophers, of every school, imagine that causation is one of the fundamental axioms or postulates of science, yet, oddly enough, in advanced sciences such as gravitational astronomy, the word "cause" never occurs. Dr. James Ward, in his *Naturalism and Agnosticism*, makes this a ground of complaint against physics: the business of those who wish to ascertain the ultimate truth about the world, he apparently thinks, should be the discovery of causes, yet physics never even seeks them. To me it seems that philosophy ought not to assume such legislative functions, and that the reason why physics has ceased to look for causes is that in fact there are no such things. The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm.

In order to find out what philosophers commonly understand by "cause", I consulted Baldwin's *Dictionary*, and was rewarded beyond my expectations, for I found the following three mutually incompatible definitions:

CAUSALITY. (1) The necessary connection of events in the time-series. ...

CAUSE (notion of). Whatever may be included in the thought or perception of a process as taking place in consequence of another process. ...

CAUSE AND EFFECT. (1) Cause and effect ... are correlative terms denoting any two distinguishable things, phases, or aspects of reality, which are so related to each other, that whenever the first ceases to exist, the second comes into existence immediately after, and whenever the second comes into existence, the first has ceased to exist immediately before.

Let us consider these three definitions in turn. The first, obviously, is unintelligible without a definition of "necessary". Under this head, Baldwin's *Dictionary* gives the following:

NECESSARY. That is necessary which not only is true, but would be true under all circumstances. Something more than brute compulsion is, therefore, involved in the conception; there is a general law under which the thing takes place.

The notion of cause is so intimately connected with that of necessity that it will be no digression to linger over the above definition, with a view to discovering, if possible, *some* meaning of which it is capable; for, as it stands, it is very far from having any definite signification.

The first point to notice is that, if any meaning is to be given to the phrase "would be true under all circumstances", the subject of it must be a propositional function, not a proposition.[1] A proposition is simply true or false, and that ends the matter: there can be no question of "circumstances". "Charles I's head was cut off" is just as true in summer as in winter, on Sundays as on Mondays. Thus when it is worth saying that something "would be true under all circumstances", the something in question must be a propositional function, i.e. an expression containing a variable, and becoming a proposition when a value is assigned to the variable; the varying "circumstances" alluded to are then the different values of which the variable is capable. Thus if "necessary" means "what is true under all circumstances", then "if x is a man, x is mortal" is necessary, because it is true for any possible value of x. Thus we should be led to the following definition:

NECESSARY is a predicate of a propositional function, meaning that it is true for all possible values of its argument or arguments.

Unfortunately, however, the definition in Baldwin's *Dictionary* says that what is necessary is not only "true under all circumstances" but is also "true". Now these two are incompatible. Only propositions can be "true", and only propositional functions can be "true under all circumstances". Hence the definition as it stands is nonsense. What is meant seems to be this: "A proposition is necessary when it is a value of a propositional function which is true under all circumstances, i.e. for all values of its argument or arguments." But if we adopt this definition, the same proposition will be necessary or contingent according as we choose one or other of its terms as the argument to our propositional function. For example, "if Socrates is a man, Socrates is mortal", is necessary if Socrates is chosen as argument, but not if *man* or *mortal* is chosen. Again, "if Socrates is a man, Plato is mortal", will be necessary if either Socrates or *man* is chosen as argument, but not if Plato or *mortal* is chosen. However, this difficulty can be overcome by specifying the constituent which is to be regarded as argument, and we thus arrive at the following definition:

> A proposition is *necessary* with respect to a given constituent if it remains true when that constituent is altered in any way compatible with the proposition remaining significant.

1 A propositional function is an expression containing a variable, or undetermined constituent, and becoming a proposition as soon as a definite value is assigned to the variable. Examples are: "A is A", "x is a number". The variable is called the *argument* of the function.

We may now apply this definition to the definition of causality quoted above. It is obvious that the argument must be the time at which the earlier event occurs. Thus an instance of causality will be such as: "If the event e_1, occurs at the time t_1, it will be followed by the event e_2". This proposition is intended to be necessary with respect to t_1, i.e. to remain true however t_1 may be varied. Causality, as a universal law, will then be the following: "Given any event e_1, there is an event e_2 such that, whenever e_1 occurs, e_2 occurs later". But before this can be considered precise, we must specify how much later e_2 is to occur. Thus the principle becomes:

> Given any event e_1, there is an event e_2 and a time-interval τ such that, whenever e_1 occurs, e_2 follows after an interval τ.

I am not concerned as yet to consider whether this law is true or false. For the present, I am merely concerned to discover what the law of causality is supposed to be. I pass, therefore, to the other definitions quoted above.

The second definition need not detain us long, for two reasons. First, because it is psychological: not the "thought or perception" of a process, but the process itself, must be what concerns us in considering causality. Secondly, because it is circular: in speaking of a process as "taking place in consequence of" another process, it introduces the very notion of cause which was to be defined.

The third definition is by far the most precise; indeed as regards clearness it leaves nothing to be desired. But a great difficulty is caused by the temporal contiguity of cause and effect which the definition asserts. No two instants are contiguous, since the time-series is compact; hence either the cause or the effect or both must, if the definition is correct, endure for a finite time; indeed, by the wording of the definition it is plain that both are assumed to endure for a finite time. But then we are faced with a dilemma: if the cause is a process involving change within itself, we shall require (if causality is universal) causal relations between its earlier and later parts; moreover it would seem that only the later parts can be relevant to the effect, since the earlier parts are not contiguous to the effect, and therefore (by the definition) cannot influence the effect. Thus we shall be led to diminish the duration of the cause without limit, and however much we may diminish it, there will still remain an earlier part which might be altered without altering the effect, so that the true cause, as defined, will not have been reached, for it will be observed that the definition excludes plurality of causes. If, on the other hand, the cause is purely static, involving no change within itself, then, in the first place, no such cause is to be found in nature, and in the second place, it seems strange – too strange to be accepted, in spite of bare logical possibility – that the cause, after

existing placidly for some time, should suddenly explode into the effect, when it might just as well have done so at any earlier time, or have gone on unchanged without producing its effect. This dilemma, therefore, is fatal to the view that cause and effect can be contiguous in time: if there are causes and effects, they must be separated by a finite time-interval τ, as was assumed in the above interpretation of the first definition.

What is essentially the same statement of the law of causality as the one elicited above from the first of Baldwin's definitions is given by other philosophers. Thus John Stuart Mill says:

> The Law of Causation, the recognition of which is the main pillar of inductive science, is but the familiar truth, that invariability of succession is found by observation to obtain between every fact in nature and some other fact which has preceded it.[2]

And Bergson, who has rightly perceived that the law as stated by philosophers is worthless, nevertheless continues to suppose that it is used in science. Thus he says:

> Now, it is argued, this law [the law of causality] means that every phenomenon is determined by its conditions, or, in other words, that the same causes produce the same effects.[3]

And again:

> We perceive physical phenomena, and these phenomena obey laws. This means: (1) that phenomena *a, b, c, d*, previously perceived, can occur again in the same shape; (2) that a certain phenomenon *P*, which appeared after the conditions *a, b, c, d*, and after these conditions only, will not fail to recur as soon as the same conditions are again present.[4]

A great part of Bergson's attack on science rests on the assumption that it employs this principle. In fact, it employs no such principle, but philosophers – even Bergson – are too apt to take their views on science from each other, not from science. As to what the principle is, there is a fair consensus among philosophers of different schools. There are, however, a number of difficulties which at once arise. I omit the question of plurality of causes for the present, since other graver questions have to be considered. Two of

2 *Logic*, Book. III, ch. V, §2.
3 *Time and Free Will*, p. 199.
4 Ibid., p. 202.

these, which are forced on our attention by the above statement of the law, are the following:

(1) What is meant by an "event"?
(2) How long may the time-interval be between cause and effect?

(1) An "event", in the statement of the law, is obviously intended to be something that is likely to recur, since otherwise the law becomes trivial. It follows that an "event" is not a particular, but some universal of which there may be many instances. It follows also that an "event" must be something short of the whole state of the universe, since it is highly improbable that this will recur. What is meant by an "event" is something like striking a match, or dropping a penny into the slot of an automatic machine. If such an event is to recur, it must not be defined too narrowly: we must not state with what degree of force the match is to be struck, nor what is to be the temperature of the penny. For if such considerations were relevant, our "event" would occur at most once, and the law would cease to give information. An "event", then, is a universal defined sufficiently widely to admit of many particular occurrences in time being instances of it.

(2) The next question concerns the time-interval. Philosophers, no doubt, think of cause and effect as contiguous in time, but this, for reasons already given, is impossible. Hence, since there are no infinitesimal time-intervals, there must be some finite lapse of time τ between cause and effect. This, however, at once raises insuperable difficulties. However short we make the interval τ, something may happen during this interval which prevents the expected result. I put my penny in the slot, but before I can draw out my ticket there is an earthquake which upsets the machine and my calculations. In order to be sure of the expected effect, we must know that there is nothing in the environment to interfere with it. But this means that the supposed cause is not, by itself, adequate to insure the effect. And as soon as we include the environment, the probability of repetition is diminished, until at last, when the whole environment is included, the probability of repetition becomes almost *nil*.

In spite of these difficulties, it must, of course, be admitted that many, fairly dependable regularities of sequence occur in daily life. It is these regularities that have suggested the supposed law of causality; where they are found to fail, it is thought that a better formulation could have been found which would have never failed. I am far from denying that there may be such sequences which in fact never do fail. It may be that there will never be an exception to the rule that when a stone of more than a certain mass, moving with more than a certain velocity, comes in contact with a pane of

glass of less than a certain thickness, the glass breaks. I also do not deny that the observation of such regularities, even when they are not without exceptions, is useful in the infancy of a science: the observation that unsupported bodies in air usually fall was a stage on the way to the law of gravitation. What I deny is that science assumes the existence of invariable uniformities of sequence of this kind, or that it aims at discovering them. All such uniformities, as we saw, depend upon a certain vagueness in the definition of the "events". That bodies fall is a vague qualitative statement; science wishes to know how fast they fall. This depends upon the shape of the bodies and the density of the air. It is true that there is more nearly uniformity when they fall in a vacuum; so far as Galileo could observe, the uniformity is then complete. But later it appeared that even there the latitude made a difference, and the altitude. Theoretically, the position of the sun and moon must make a difference. In short, every advance in a science takes us farther away from the crude uniformities which are first observed, into greater differentiation of antecedent and consequent, and into a continually wider circle of antecedents recognized as relevant.

The principle "same cause, same effect", which philosophers imagine to be vital to science, is therefore utterly otiose. As soon as the antecedents have been given sufficiently fully to enable the consequent to be calculated with some exactitude, the antecedents have become so complicated that it is very unlikely they will ever recur. Hence, if this were the principle involved, science would remain utterly sterile.

The importance of these considerations lies partly in the fact that they lead to a more correct account of scientific procedure, partly in the fact that they remove the analogy with human volition which makes the conception of cause such a fruitful source of fallacies. The latter point will become clearer by the help of some illustrations. For this purpose I shall consider a few maxims which have played a great part in the history of philosophy.

(1) "Cause and effect must more or less resemble each other". This principle was prominent in the philosophy of occasionalism, and is still by no means extinct. It is still often thought, for example, that mind could not have grown up in a universe which previously contained nothing mental; and one ground for this belief is that matter is too dissimilar from mind to have been able to cause it. Or, more particularly, what are termed the nobler parts of our nature are supposed to be inexplicable unless the universe always contained something at least equally noble which could cause them. All such views seem to depend upon assuming some unduly simplified law of causality; for, in any legitimate sense of "cause" and "effect", science seems to show that they are usually very widely dissimilar, the "cause" being, in fact, two states of the whole universe, and the "effect" some particular event.

(2) "Cause is analogous to volition, since there must be an intelligible *nexus* between cause and effect". This maxim is, I think, often unconsciously in the imaginations of philosophers who would reject it when explicitly stated. It is probably operative in the view we have just been considering, that mind could not have resulted from a purely material world. I do not profess to know what is meant by "intelligible"; it seems to mean "familiar to imagination". Nothing is less "intelligible", in any other sense, than the connection between an act of will and its fulfilment. But obviously the sort of nexus desired between cause and effect is such as could only hold between the "events" which the supposed law of causality contemplates; the laws which replace causality in such a science as physics leave no room for any two events between which a nexus could be sought.

(3) "The cause *compels* the effect in some sense in which the effect does not compel the cause". This belief seems largely operative in the dislike of determinism; but as a matter of fact, it is connected with our second maxim, and falls as soon as that is abandoned. We may define "compulsion" as follows: "Any set of circumstances is said to compel *A* when *A* desires to do something which the circumstances prevent, or to abstain from something which the circumstances cause". This presupposes that some meaning has been found for the word "cause" – a point to which I shall return later. What I want to make clear at present is that compulsion is a very complex notion, involving thwarted desire. So long as a person does what he wishes to do, there is no compulsion, however much his wishes may be calculable by the help of earlier events. And where desire does not come in, there can be no question of compulsion. Hence it is in general misleading to regard the cause as compelling the effect.

A vaguer form of the same maxim substitutes the word "determine" for the word "compel": we are told that the cause *determines* the effect in a sense in which the effect does not *determine* the cause. It is not quite clear what is meant by "determining"; the only precise sense, so far as I know, is that of a function or one-many relation. If we admit plurality of causes, but not of effects, that is, if we suppose that, given the cause, the effect must be such and such, but given the effect, the cause may have been one of many alternatives, then we may say that the cause determines the effect, but not the effect the cause. Plurality of causes, however, results only from conceiving the effect vaguely and narrowly and the cause precisely and widely. Many antecedents may "cause" a man's death, because his death is vague and narrow. But if we adopt the opposite course, taking as the "cause" the drinking of a dose of arsenic, and as the "effect" the whole state of the world five minutes later, we shall have plurality of effects instead of plurality of causes. Thus the supposed lack of symmetry between "cause" and "effect" is illusory.

(4) "A cause cannot operate when it has ceased to exist, because what has ceased to exist is nothing". This is a common maxim, and a still more common unexpressed prejudice. It has, I fancy, a good deal to do with the attractiveness of Bergson's "*durée*": since the past has effects now, it must still exist in some sense. The mistake in this maxim consists in the supposition that causes "operate" at all. A volition "operates" when what it wills takes place; but nothing can operate except a volition. The belief that causes "operate" results from assimilating them, consciously or unconsciously, to volitions. We have already seen that, if there are causes at all, they must be separated by a finite interval of time from their effects, and thus cause their effects after they have ceased to exist.

It may be objected to the above definition of a volition "operating" that it only operates when it "causes" what it wills, not when it merely happens to be followed by what it wills. This certainly represents the usual view of what is meant by a volition "operating", but as it involves the very view of causation which we are engaged in combating, it is not open to us as a definition. We may say that a volition "operates" when there is some law in virtue of which a similar volition in rather similar circumstances will usually be followed by what it wills. But this is a vague conception, and introduces ideas which we have not yet considered. What is chiefly important to notice is that the usual notion of "operating" is not open to us if we reject, as I contend that we should, the usual notion of causation.

(5) "A cause cannot operate except where it is". This maxim is very widespread: it was urged against Newton, and has remained a source of prejudice against "action at a distance". In philosophy it has led to a denial of transient action, and thence to monism or Leibnizian monadism. Like the analogous maxim concerning temporal contiguity, it rests upon the assumption that causes "operate", i.e. that they are in some obscure way analogous to volitions. And as in the case of temporal contiguity, the inferences drawn from this maxim are wholly groundless.

I return now to the question what law or laws can be found to take the place of the supposed law of causality?

First, without passing beyond such uniformities of sequence as are contemplated by the traditional law, we may admit that, if any such sequence has been observed in a great many cases, and has never been found to fail, there is an inductive probability that it will be found to hold in future cases. If stones have hitherto been found to break windows, it is probable that they will continue to do so. This of course assumes the inductive principle, of which the truth may reasonably be questioned; but as this principle is not our present concern, I shall in this discussion treat it as

indubitable. We may then say, in the case of any such frequently observed sequence, that the earlier event is the *cause* and the later event the *effect*.

Several considerations, however, make such special sequences very different from the traditional relation of cause and effect. In the first place, the sequence, in any hitherto unobserved instance, is no more than probable, whereas the relation of cause and effect was supposed to be necessary. I do not mean by this merely that we are not sure of having discovered a true case of cause and effect; I mean that, even when we have a case of cause and effect in our present sense, all that is meant is that, on grounds of observation, it is probable that when one occurs the other will also occur. Thus in our present sense, *A* may be the cause of *B* even if there actually are cases where *B* does not follow *A*. Striking a match will be the cause of its igniting, in spite of the fact that some matches are damp and fail to ignite.

In the second place, it will not be assumed that *every* event has some antecedent which is its cause in this sense: we shall only believe in causal sequences where we find them, without any presumption that they always are to be found.

In the third place, *any* case of sufficiently frequent sequence will be causal in our present sense; for example, we shall not refuse to say that night is the cause of day. Our repugnance to saying this arises from the ease with which we can imagine the sequence to fail, but owing to the fact that cause and effect must be separated by a finite interval of time, *any* such sequence *might* fail through the interposition of other circumstances in the interval. Mill, discussing this instance of night and day, says:

> It is necessary to our using the word cause, that we should believe not only that the antecedent always *has* been followed by the consequent, but that as long as the present constitution of things endures, it always *will* be so.[5]

In this sense, we shall have to give up the hope of finding causal laws such as Mill contemplated; any causal sequence which we have observed may at any moment be falsified without a falsification of any laws of the kind that the more advanced sciences aim at establishing.

In the fourth place, such laws of probable sequence, though useful in daily life and in the infancy of a science, tend to be displaced by quite different laws as soon as a science is successful. The law of gravitation will illustrate what occurs in any advanced science. In the motions of mutually gravitating bodies, there is nothing that can be called a cause, and nothing

5 Loc. cit., §6.

that can be called an effect; there is merely a formula. Certain differential equations can be found, which hold at every instant for every particle of the system, and which, given the configuration and velocities at one instant, or the configurations at two instants, render the configuration at any other earlier or later instant theoretically calculable. That is to say, the configuration at any instant is a function of that instant and the configurations at two given instants. This statement holds throughout physics, and not only in the special case of gravitation. But there is nothing that could be properly called "cause" and nothing that could be properly called "effect" in such a system.

No doubt the reason why the old "law of causality" has so long continued to pervade the books of philosophers is simply that the idea of a function is unfamiliar to most of them, and therefore they seek an unduly simplified statement. There is no question of repetitions, of the "same" cause producing the "same" effect; it is not in any sameness of causes and effects that the constancy of scientific laws consists, but in sameness of relations. And even "sameness of relations" is too simple a phrase; "sameness of differential equations" is the only correct phrase. It is impossible to state this accurately in non-mathematical language; the nearest approach would be as follows: "There is a constant relation between the state of the universe at any instant and the rate of change in the rate at which any part of the universe is changing at that instant; and this relation is many-one, i.e. such that the rate of change in the rate of change is determinate when the state of the universe is given". If the "law of causality" is to be something actually discoverable in the practice of science, the above proposition has a better right to the name than any "law of causality" to be found in the books of philosophers.

In regard to the above principle, several observations must be made.

(1) No one can pretend that the above principle is à priori or self-evident or a "necessity of thought". Nor is it, in any sense, a premiss of science: it is an empirical generalization from a number of laws which are themselves empirical generalizations.

(2) The law makes no difference between past and future: the future "determines" the past in exactly the same sense in which the past "determines" the future. The word "determine", here, has a purely logical significance: a certain number of variables "determine" another variable if that other variable is a function of them.

(3) The law will not be empirically verifiable unless the course of events within some sufficiently small volume will be approximately the same in any two states of the universe which only differ in regard to what is at a considerable distance from the small volume in question. For example, motions of planets in the solar system must be approximately the same however the

fixed stars may be distributed, provided that all the fixed stars are very much farther from the sun than the planets are. If gravitation varied directly as the distance, so that the most remote stars made the most difference to the motions of the planets, the world might be just as regular and just as much subject to mathematical laws as it is at present, but we could never discover the fact.

(4) Although the old "law of causality" is not assumed by science something which we may call the "uniformity of nature" is assumed, or rather is accepted on inductive grounds. The uniformity of nature does not assert the trivial principle "same cause, same effect", but the principle of the permanence of laws. That is to say, when a law exhibiting, e.g. an acceleration as a function of the configuration has been found to hold throughout the observable past, it is expected that it will continue to hold in the future, or that, if it does not itself hold, there is some other law, agreeing with the supposed law as regards the past, which will hold for the future. The ground of this principle is simply the inductive ground that it has been found to be true in very many instances; hence the principle cannot be considered certain, but only probable to a degree which cannot be accurately estimated.

The uniformity of nature, in the above sense, although it is assumed in the practice of science, must not, in its generality, be regarded as a kind of major premiss, without which all scientific reasoning would be in error. The assumption that *all* laws of nature are permanent has, of course, less probability than the assumption that this or that particular law is permanent; and the assumption that a particular law is permanent for all time has less probability than the assumption that it will be valid up to such and such a date. Science, in any given case, will assume what the case requires, but no more. In constructing the *Nautical Almanac* for 1915 it will assume that the law of gravitation will remain true up to the end of that year; but it will make no assumption as to 1916 until it comes to the next volume of the almanac. This procedure is, of course, dictated by the fact that the uniformity of nature is not known a priori, but is an empirical generalization, like "all men are mortal". In all such cases, it is better to argue immediately from the given particular instances to the new instance, than to argue by way of a major premiss: the conclusion is only probable in either case, but acquires a higher probability by the former method than by the latter.

In all science, we have to distinguish two sorts of laws: first, those that are empirically verifiable but probably only approximate, secondly, those for that are not verifiable, but may be exact. The law of gravitation, for example, in its applications to the solar system, is only empirically verifiable when it is assumed that matter outside the solar system may be

ignored for such purposes; we believe this to be only approximately true, but we cannot empirically verify the law of universal gravitation which we believe to be exact. This point is very important in connection with what we may call "relatively isolated systems". These may be defined as follows:

A system relatively isolated during a given period is one which, within some assignable margin of error, will behave in the same way throughout that period however the rest of the universe may be constituted.

A system may be called "practically isolated" during a given period if, although there *might* be states of the rest of the universe which would produce more than the assigned margin of error, there is reason to believe that such states do not in fact occur.

Strictly speaking, we ought to specify the respect in which the system is relatively isolated. For example, the earth is relatively isolated as regards falling bodies, but not as regards tides; it is *practically* isolated as regards economic phenomena, although, if Jevons's sun-spot theory of commercial crises had been true, it would not have been even practically isolated in this respect.

It will be observed that we cannot prove in advance that a system is isolated. This will be inferred from the observed fact that approximate uniformities can be stated for this system alone. If the complete laws for the whole universe were known, the isolation of a system could be deduced from them; assuming, for example, the law of universal gravitation, the practical isolation of the solar system in this respect can be deduced by the help of the fact that there is very little matter in its neighbourhood. But it should be observed that isolated systems are only important as providing a possibility of *discovering* scientific laws; they have no theoretical importance in the finished structure of a science.

The case where one event A is said to "cause" another event B, which philosophers take as fundamental, is really only the most simplified instance of a practically isolated system. It may happen that, as a result of general scientific laws, whenever A occurs throughout a certain period, it is followed by B; in that case, A and B form a system which is practically isolated throughout that period. It is, however, to be regarded as a piece of good fortune if this occurs; it will always be due to special circumstances, and would not have been true if the rest of the universe had been different though subject to the same laws.

The essential function which causality has been supposed to perform is the possibility of inferring the future from the past, or, more generally, events at any time from events at certain assigned times. Any system in which such inference is possible may be called a "deterministic" system. We may define a deterministic system as follows:

A system is said to be "deterministic" when, given certain data e_1, e_2, … , e_n, at times t_1, t_2, …, t_n respectively, concerning this system, if E_t is

the state of the system at any time t, there is a functional relation of the form

$$E_t = f(e_1, t_1, e_2, t_2, \ldots, e_n, t_n, t). \qquad \text{(A)}$$

The system will be "deterministic throughout a given period" if t, in the above formula, may be any time within that period, though outside that period the formula may be no longer true. If the universe as a whole is such a system, determinism is true of the universe; if not, not. A system which is part of a deterministic system I shall call "determined"; one which is not part of any such system I shall call "capricious".

The events $e_1, e_2, \ldots, e_n$ I shall call "determinants" of the system. It is to be observed that a system which has one set of determinants will in general have many. In the case of the motions of the planets, for example, the configurations of the solar system at any two given times will be determinants.

We may take another illustration from the hypothesis of psycho-physical parallelism. Let us assume, for the purposes of this illustration, that to a given state of brain a given state of mind always corresponds, and vice versa, i.e. that there is a one–one relation between them, so that each is a function of the other. We may also assume, what is practically certain, that to a given state of a certain brain a given state of the whole material universe corresponds, since it is highly improbable that a given brain is ever twice in exactly the same state. Hence there will be a one–one relation between the state of a given person's mind and the state of the whole material universe. It follows that, if n states of the material universe are determinants of the material universe, then n states of a given man's mind are determinants of the whole material and mental universe – assuming, that is to say, that psycho-physical parallelism is true.

The above illustration is important in connection with a certain confusion which seems to have beset those who have philosophized on the relation of mind and matter. It is often thought that, if the state of the mind is determinate when the state of the brain is given, and if the material world forms a deterministic system, then mind is "subject" to matter in some sense in which matter is not "subject" to mind. But if the state of the brain is also determinate when the state of the mind is given, it must be exactly as true to regard matter as subject to mind as it would be to regard mind as subject to matter. We could, theoretically, work out the history of mind without ever mentioning matter, and then, at the end, deduce that matter must meanwhile have gone through the corresponding history. It is true that if the relation of brain to mind were many–one, not one–one, there would be a one-sided dependence of mind on brain, while conversely, if the relation were one–many, as Bergson supposes, there would be a one-sided dependence of brain on mind. But the dependence involved is, in any case,

only logical; it does not mean that we shall be compelled to do things we desire not to do, which is what people instinctively imagine it to mean.

As another illustration we may take the case of mechanism and teleology. A system may be defined as "mechanical" when it has a set of determinants that are purely material, such as the positions of certain pieces of matter at certain times. It is an open question whether the world of mind and matter, as we know it, is a mechanical system or not; let us suppose, for the sake of argument, that it is a mechanical system. This supposition – so I contend – throws no light whatever on the question whether the universe is or is not a "teleological" system. It is difficult to define accurately what is meant by a "teleological" system, but the argument is not much affected by the particular definition we adopt. Broadly, a teleological system is one in which purposes are realized, i.e. in which certain desires – those that are deeper or nobler or more fundamental or more universal or what not – are followed by their realization. Now the fact – if it be a fact – that the universe is mechanical has no bearing whatever on the question whether it is teleological in the above sense. There might be a mechanical system in which all wishes were realized, and there might be one in which all wishes were thwarted. The question whether, or how far, our actual world is teleological, cannot, therefore, be settled by proving that it is mechanical; and the desire that it should be teleological is no ground for wishing it to be not mechanical.

There is, in all these questions, a very great difficulty in avoiding confusion between what we can infer and what is in fact determined. Let us consider, for a moment, the various senses in which the future may be "determined". There is one sense – and a very important one – in which it is determined quite independently of scientific laws, namely the sense that it will be what it will be. We all regard the past as determined simply by the fact that it has happened; but for the accident that memory works backward and not forward, we should regard the future as equally determined by the fact that it will happen. "But", we are told, "you cannot alter the past, while you can to some extent alter the future". This view seems to me to rest upon just those errors in regard to causation which it has been my object to remove. You cannot make the past other than it was – true, but this is a mere application of the law of contradiction. If you already know what the past was, obviously it is useless to wish it different. But also you cannot make the future other than it will be; this again is an application of the law of contradiction. And if you happen to know the future – e.g. in the case of a forthcoming eclipse – it is just as useless to wish it different as to wish the past different. "But", it will be rejoined, "our wishes can *cause* the future, sometimes, to be different from what it would be if they did not exist, and they can have no such effect upon the past". This again is a mere

tautology. An effect being *defined* as something subsequent to its cause, obviously we can have no *effect* upon the past. But that does not mean that the past would not have been different if our present wishes had been different. Obviously, our present wishes are conditioned by the past, and therefore could not have been different unless the past had been different; therefore, if our present wishes were different, the past would be different. Of course, the past cannot be different from what it was, but no more can our present wishes be different from what they are; this again is merely the law of contradiction. The facts seem to be merely (1) that wishing generally depends upon ignorance, and is therefore commoner in regard to the future than in regard to the past, (2) that where a wish concerns the future, it and its realization very often form a "practically independent system", i.e. many wishes regarding the future are realized. But there seems no doubt that the main difference in our feelings arises from the accidental fact that the past but not the future can be known by memory.

Although the sense of "determined" in which the future is determined by the mere fact that it will be what it will be is sufficient (at least so it seems to me) to refute some opponents of determinism, notably M. Bergson and the pragmatists, yet it is not what most people have in mind when they speak of the future as determined. What they have in mind is a formula by means of which the future can be exhibited, and at least theoretically calculated, as a function of the past. But at this point we meet with a great difficulty, which besets what has been said above about deterministic systems, as well as what is said by others.

If formulae of *any* degree of complexity, however great, are admitted, it would seem that any system, whose state at a given moment is a function of certain measurable quantities, *must* be a deterministic system. Let us consider, in illustration, a single material particle, whose coordinates at time t are x_t, y_t, z_t. Then however the particle moves, there must be, theoretically, functions f_1, f_2, f_3, such that

$$x_t = f_1(t), \quad y_t = f_2(t), \quad z_t = f_3(t).$$

It follows that, theoretically, the whole state of the material universe at time t must be capable of being exhibited as a function of t. Hence our universe will be deterministic in the sense defined above. But if this be true, no information is conveyed about the universe in stating that it is deterministic. It is true that the formulae involved may be of strictly infinite complexity, and therefore not practically capable of being written down or apprehended. But except from the point of view of our knowledge, this might seem to be a detail; in itself, if the above considerations are sound, the material universe *must* be deterministic, *must* be subject to laws.

This, however, is plainly not what was intended. The difference between this view and the view intended may be seen as follows. Given some formula which fits the facts hitherto – say the law of gravitation – there will be an infinite number of other formulae, not empirically distinguishable from it in the past, but diverging from it more and more in the future. Hence, even assuming that there are persistent laws, we shall have no reason for assuming that the law of the inverse square will hold in future: it may be some other hitherto indistinguishable law that will hold. We cannot say that *every* law which has held hitherto must hold in the future, because past facts which obey one law will also obey others, hitherto indistinguishable but diverging in future. Hence there must, at every moment, be laws hitherto unbroken which are now broken for the first time. What science does, in fact, is to select the *simplest* formula that will fit the facts. But this, quite obviously, is merely a methodological precept, not a law of Nature. If the simplest formula ceases, after a time, to be applicable, the simplest formula that remains applicable is selected, and science has no sense that an axiom has been falsified. We are thus left with the brute fact that, in many departments of science, quite simple laws have hitherto been found to hold. This fact cannot be regarded as having any a priori ground, nor can it be used to support inductively the opinion that the same laws will continue; for at every moment laws hitherto true are being falsified, though in the advanced sciences these laws are less simple than those that have remained true. Moreover it would be fallacious to argue inductively from the state of the advanced sciences to the future state of the others, for it may well be that the advanced sciences are advanced simply because, hitherto, their subject-matter has obeyed simple and easily ascertainable laws, while the subject-matter of other sciences has not done so.

The difficulty we have been considering seems to be met partly, if not wholly, by the principle that the *time* must not enter explicitly into our formulae. All mechanical laws exhibit acceleration as a function of configuration, not of configuration and time jointly; and this principle of the irrelevance of the time may be extended to all scientific laws. In fact we might interpret the "uniformity of nature" as meaning just this, that no scientific law involves the time as an argument, unless, of course, it is given in an integrated form, in which case *lapse* of time, though not absolute time, may appear in our formulae. Whether this consideration suffices to overcome our difficulty completely, I do not know; but in any case it does much to diminish it.

It will serve to illustrate what has been said if we apply it to the question of free will.

(1) Determinism in regard to the will is the doctrine that our volitions belong to some deterministic system, i.e. are "determined" in the sense

defined above. Whether this doctrine is true or false, is a mere question of fact: no a priori considerations (if our previous discussions have been correct) can exist on either side. On the one hand, there is no a priori category of causality, but merely certain observed uniformities. As a matter of fact, there are observed uniformities in regard to volitions; thus there is some empirical evidence that volitions are determined. But it would be very rash to maintain that the evidence is overwhelming, and it is quite possible that some volitions, as well as some other things, are not determined, except in the sense in which we found that everything must be determined.

(2) But, on the other hand, the subjective sense of freedom, sometimes alleged against determinism, has no bearing on the question whatever. The view that it has a bearing rests upon the belief that causes compel their effects, or that nature enforces obedience to its laws as governments do. These are mere anthropomorphic superstitions, due to assimilation of causes with volitions and of natural laws with human edicts. We feel that our will is not compelled, but that only means that it is not other than we choose it to be. It is one of the demerits of the traditional theory of causality that it has created an artificial opposition between determinism and the freedom of which we are introspectively conscious.

(3) Besides the general question whether volitions are determined, there is the further question whether they are *mechanically* determined, i.e. whether they are part of what was above defined as a mechanical system. This is the question whether they form part of a system with purely material determinants, i.e. whether there are laws which, given certain material data, make all volitions functions of those data. Here again, there is empirical evidence up to a point, but it is not conclusive in regard to all volitions. It is important to observe, however, that even if volitions are part of a mechanical system, this by no means implies any supremacy of matter over mind. It may well be that the same system which is susceptible of material determinants is also susceptible of mental determinants; thus a mechanical system may be determined by sets of volitions, as well as by sets of material facts. It would seem, therefore, that the reasons which make people dislike the view that volitions are mechanically determined are fallacious.

(4) The notion of *necessity*, which is often associated with determinism, is a confused notion not legitimately deducible from determinism. Three meanings are commonly confounded when necessity is spoken of:

(α) An *action* is necessary when it will be performed however much the agent may wish to do otherwise. Determinism does not imply that actions are necessary in this sense.

(β) A *propositional function* is necessary when all its values are true. This sense is not relevant to our present discussion.

(γ) A *proposition* is necessary with respect to a given constituent when it is the value, with that constituent as argument, of a necessary propositional function, in other words, when it remains true however that constituent may be varied. In this sense, in a deterministic system, the connection of a volition with its determinants is necessary, if the time at which the determinants occur be taken as the constituent to be varied, the time-interval between the determinants and the volition being kept constant. But this sense of necessity is purely logical, and has no emotional importance.

We may now sum up our discussion of causality. We found first that the law of causality, as usually stated by philosophers, is false, and is not employed in science. We then considered the nature of scientific laws, and found that, instead of stating that one event A is always followed by another event B, they stated functional relations between certain events at certain times, which we called determinants, and other events at earlier or later times or at the same time. We were unable to find any a priori category involved: the existence of scientific laws appeared as a purely empirical fact, not necessarily universal, except in a trivial and scientifically useless form. We found that a system with one set of determinants may very likely have other sets of a quite different kind, that, for example, a mechanically determined system may also be teleologically or volitionally determined. Finally we considered the problem of free will: here we found that the reasons for supposing volitions to be determined are strong but not conclusive, and we decided that even if volitions are mechanically determined, that is no reason for denying freedom in the sense revealed by introspection, or for supposing that mechanical events are not determined by volitions. The problem of free will versus determinism is therefore, if we were right, mainly illusory, but in part not yet capable of being decisively solved.

15

CAUSAL LAWS IN PHYSICS

1927

There are a number of interesting features in this selection taken from Russell's book *Outline of Philosophy* (1927a). The basic position is that we do not experience what we think we do in respect of causal laws and of physical objects.

Russell proposes a reduction of objects to events. We believe there are physical objects persisting in time, whereas Russell points out that physics has no such conception. In place of persisting objects, there is merely a series of events. The illusion of permanence is illustrated by analogy with the "objects" represented on a cinema screen. They are created by a series of events: a succession of fleeting images. When viewed in sequence, they appear to us as a single object, continuing through time and undergoing changes. Why, then, do we believe in the permanence of physical objects and must we do so?

The belief is metaphysical, in so far as the empirical evidence alone is insufficient to dictate either a metaphysic of permanent objects or of successions of events ("there is nothing in the actual experience to show whether there is a persistent entity or not"). Might it be that we have an in-built predisposition to believe in permanent physical objects? Psychologists refer to a belief in "object permanence" that emerges in children at about nine months (Gopnik, Meltzoff and Kuhl 1999: ch. 3). This is supposed to be a belief that objects continue to exist unperceived. Perhaps the belief bestows an evolutionary advantage. If we do have an in-built metaphysic of permanent objects, it might be seen as a weighty reply to Locke's famous denial of innate ideas in his *Essay* (1690: Book I).

But why would we have an in-built disposition towards a metaphysic of substance, a metaphysic that Russell believes to be in scientific error? Given no empirical difference between a metaphysics of substance and of fleeting events, it would seem hard to justify, on empirical evidence, any claim that humans favour one metaphysic over another. How do we know that infants

are not Berkeleyans and believe only in a coherent and connected set of appearances? Why should a convinced Berkeleyan have less evolutionary advantage than a substance theorist? A belief in the permanent possibility of sensation seems just as advantageous as a belief in permanent objects that exist whether perceived or not.

On causation, Russell's emphasis in this selection is on the denial of forces acting between bodies. This, he explains, is no part of modern physics. It is, however, an idea that has proved hard to eliminate. We have employed it as "a rationalising of natural processes", but it is a "fruitless one" and it is no part of our experience. Russell appears to believe that, for all scientific purposes, a reduction of "force-talk" to non- "force-talk" is possible. Again, the empiricist-leaning Russell has rejected a class of modal truth and rid the world of more causal action.

In the last chapter [not included] we spoke about the substitution of space-time for space and time, and the effect which this has had in substituting strings of events for "things" conceived as substances. In this chapter we will deal with cause and effect as they appear in the light of modern science. It is at least as difficult to purge our imagination of irrelevances in this matter as in regard to substance. The old-fashioned notion of cause appeared in dynamics as "force". We still speak of forces just as we still speak of the sunrise, but we recognise that this is nothing but a convenient way of speaking, in the one case as in the other.

Causation is deeply embedded in language and common sense. We say that people build houses or make roads: to "build" and to "make" are both notions involving causality. We say that a man is "powerful", meaning that his volitions are causes over a wide range. Some examples of causation seem to us quite natural, others less so. It seems natural that our muscles should obey our will, and only reflection makes us perceive the necessity of finding an explanation of this phenomenon. It seems natural that when you hit a billiard-ball with a cue it moves. When we see a horse pulling a cart, or a heavy object being dragged by a rope, we *feel* as if we understood all about it. It is events of this sort that have given rise to the common-sense belief in causes and forces.

But as a matter of fact the world is incredibly more complicated than it seems to common sense. When we think we understand a process – I mean by "we" the non-reflective part in each of us – what really happens is that there is some sequence of events so familiar through past experience that at each stage we expect the next stage. The whole process seems to us peculiarly intelligible when human desires enter in, for example, in watching a game: what the ball does and what the players do seems "natural", and we feel as if we quite understood how the stages succeed each other. We thus

arrive at the notion of what is called "necessary" sequence. The text-books say that *A* is the cause of *B* if *A* is "necessarily" followed by *B*. This notion of "necessity" seems to be purely anthropomorphic, and not based upon anything that is a discoverable feature of the world. Things happen according to certain rules; the rules can be generalised, but in the end remain brute facts. Unless the rules are concealed conventions or definitions, no reason can be given why they should not be completely different.

To say that *A* is "necessarily" followed by *B* is thus to say no more than that there is some general rule, exemplified in a very large number of observed instances, and falsified in none, according to which events such as *A* are followed by events such as *B*. We must not have any notion of "compulsion", as if the cause *forced* the effect to happen. A good test for the imagination in this respect is the reversibility of causal laws. We can just as often infer backwards as forwards. When you get a letter, you are justified in inferring that somebody wrote it, but you do not feel that your receiving it *compelled* the sender to write it. The notion of compulsion is just as little applicable to effects as to causes. To say that causes compel effects is as misleading as to say that effects compel causes. Compulsion is anthropomorphic: a man is compelled to do something when he wishes to do the opposite, but except where human or animal wishes come in the notion of compulsion is inapplicable. Science is concerned merely with what happens, not with what *must* happen.

When we look for invariable rules of sequence in nature, we find that they are not such as common sense sets up. Common sense says: thunder follows lightning, waves at sea follow wind, and so on. Rules of this sort are indispensable in practical life, but in science they are all only approximate. If there is any finite interval of time, however short, between the cause and the effect, something may happen to prevent the effect from occurring. Scientific laws can only be expressed in differential equations. This means that, although you cannot tell what may happen after a finite time, you can say that, if you make the time shorter and shorter, what will happen will be more and more nearly according to such-and-such a rule. To take a very simple case: I am now in this room; you cannot tell where I shall be in another second, because a bomb may explode and blow me sky-high, but if you take any two small fragments of my body which are now very close together, you can be sure that, after *some* very short finite time, they will still be very close together. If a second is not short enough, you must take a shorter time; you cannot tell in advance how short a time you may have to take, but you may feel fairly certain that there is a short enough time.

The laws of sequence in physics, apart from quantum phenomena, are of two sorts, which appeared in traditional dynamics as laws of velocity and laws of acceleration. In a very short time, the velocity of a body alters very

little, and if the time is taken short enough, the change of velocity diminishes without limit. This is what, in the last chapter, we called an "intrinsic" causal law. Then there is the effect of the outer world, as it appeared in traditional dynamics, which is shown in acceleration. The small change which does occur in the velocity in a short time is attributed to surrounding bodies, because it is found to vary as they vary, and to vary according to ascertained laws. Thus we think of surrounding bodies as exerting an influence, which we call "force", though this remains as mysterious as the influence of the stars in astrology.

Einstein's theory of gravitation has done away with this conception in so far as gravitational forces are concerned. In this theory, a planet moving round the sun is moving in the nearest approach to a straight line that the neighbourhood permits. The neighbourhood is supposed to be non-Euclidean, that is to say, to contain no straight lines such as Euclid imagined. If a body is moving freely, as the planets do, it observes a certain rule. Perhaps the simplest way to state this rule is as follows: Suppose you take any two events which happen on the earth, and you measure the time between them by ideally accurate clocks which move with the earth. Suppose some traveller on a magic carpet had meanwhile cruised about the universe, leaving the earth at the time of the first event and returning at the time of the second. By his clocks the period elapsed will be less than by the terrestial clocks. This is what is meant by saying that the earth moves in a "geodesic", which is the nearest approach to a straight line to be found in the region in which we live. All this is, so to speak, geometrical, and involves no "forces". It is not the sun that makes the earth go round, but the nature of space-time where the earth is.

Even this is not quite correct. Space-time does not make the earth go round the sun; it makes us *say* the earth goes round the sun. That is to say, it makes this the shortest way of describing what occurs. We could describe it in other language, which would be equally correct, but less convenient.

The abolition of "force" in astronomy is perhaps connected with the fact that astronomy depends only upon the sense of sight. On the earth, we push and pull, we touch things, and we experience muscular strains. This all gives us a notion of "force", but this notion is anthropomorphic. To imagine the laws of motion of heavenly bodies, think of the motions of objects in a mirror; they may move very fast, although in the mirror world there are no forces.

What we really have to substitute for force is laws of correlation. Events can be collected in groups by their correlations. This is all that is true in the old notion of causality. And this is not a "postulate" or "category", but an observed fact – lucky, not necessary.

As we suggested before, it is these correlations of events that lead to the definition of permanent "things". There is no essential difference, as regards substantiality, between an electron and a light-ray. Each is really a string of events or of sets of events. In the case of the light-ray, we have no temptation to think otherwise. But in the case of the electron, we think of it as a single persistent entity. There *may* be such an entity, but we can have no evidence that there is. What we can discover is (*a*) a group of events spreading outwards from a centre – say, for definiteness, the events constituting a wave of light – and attributed, hypothetically, to a "cause" in that centre; (*b*) more or less similar groups of events at other times, connected with the first group according to the laws of physics, and therefore attributed to the same hypothetical cause at other times. But all that we ought to assume is series of groups of events, connected by discoverable laws. These series we may *define* as "matter". Whether there is matter in any other sense, no one can tell.

What is true in the old notion of causality is the fact that events at different times are connected by laws (differential equations). When there is a law connecting an event *A* with an event *B*, the two have a definite unambiguous time-order. But if the events are such that a ray of light starting from *A* would arrive at any body which was present at *B* after *B* had occurred, and *vice versa*, then there is no definite time order, and no possible causal law connecting *A* and *B*. *A* and *B* must then be regarded as separate facts of geography.

Perhaps the scope and purpose of this and the foregoing chapters may be made clearer by showing their bearing upon certain popular beliefs which may seem self-evident but are really, in my opinion, either false or likely to lead to falsehood. I shall confine myself to objections which have actually been made to me when trying to explain the philosophical outcome of modern physics.[1]

"We cannot conceive of movement apart from some *thing* as moving". This is, in a sense, a truism; but in the sense in which it is usually meant, it is a falsehood. We speak of the "movement" of a drama or piece of music, although we do not conceive either as a "thing" which exists complete at every moment of the performance. This is the sort of picture we must have in our minds when we try to conceive the physical world. We must think of a string of events connected together by certain causal connections, and having enough unity to deserve a single name. We then begin to imagine

1 These objections are quoted (with kind permission) from a letter written to me by a well-known engineer, Mr. Percy Griffith, who is also a writer on philosophical subjects.

that the single name denotes a single "thing", and if the events concerned are not all in the same place, we say the "thing" has "moved". But this is only a convenient shorthand. In the cinema, we seem to see a man falling off a skyscraper, catching hold of the telegraph wires, and reaching the ground none the worse. We know that in fact, there are a number of different photographs, and the appearance of a single "thing" moving is deceptive. In this respect, the real world resembles the cinema.

In connection with motion one needs to emphasise the very difficult distinction between *experience* and *prejudice*. Experience, roughly, is what you see, and prejudice is what you only think you see. Prejudice tells you that you see the *same* table on two different occasions; you *think* that experience tells you this. If it really were experience, you could not be mistaken; yet a similar table may be substituted without altering the experience. If you look at a table on two different occasions, you have very similar sensations, and memory tells you that they are similar; but there is nothing to show that one identical entity causes the two sensations. If the table is in a cinema, you know that there is not such an entity, even though you can watch it changing with apparent continuity. The experience is just like that with a "real" table; so in the case of a "real" table also, there is nothing in the actual experience to show whether there is a persistent entity or not. I say, therefore: I do not know whether there is a persistent entity, but I do know that my experiences can be explained without assuming that there is. Therefore it can be no part of legitimate science to assert or deny the persistent entity; if it does either, it goes beyond the warrant of experience.

The following is a verbally cited passage in the letter referred to objecting to what was said above about "force":

"The concept of Force is not of physical but of psychological origin. Rightly or wrongly it arises in the most impersonal contemplation of the Stellar Universe, where we observe an infinite number of spherical bodies revolving on their own axes and gyrating in orbits round each other. Rightly or wrongly, we naturally conceive of these as having been so constituted and so maintained by some Force or Forces."

We do not, in fact, "observe" what it is here said that we observe; all this is *inferred*. What we observe, in astronomy, is a two-dimensional pattern of points of light, with a few bright surfaces of measurable size when seen through the telescope (the planets), and of course the larger bright surfaces that we call the sun and moon. Most of this pattern (the fixed stars) rotates round the earth once in every twenty-three hours and fifty-six minutes. The sun rotates in varying periods, which average twenty-four hours and never depart very far from the average. The moon and planets have apparent motions which are more irregular. These are the *observed* facts. There is no logical impossibility about the mediæval doctrine of spheres rotating

round the earth, one for each planet and one for the stars. The modern doctrines are simpler, but not one whit more in accordance with observed facts; it is our passion for *simple* laws that has made us adopt them.

The last sentence of the above quotation raises some further points of interest. "Rightly or wrongly", the writer says, "we naturally conceive of these as having been so constituted and so maintained by some Force or Forces". I do not deny this. It is "natural", and it is "right or wrong" – more specifically, it is wrong. "Force" is part of our love of explanations. Everyone knows about the Hindu who thought that the world does not fall because it is supported by an elephant, and the elephant does not fall because it is supported by a tortoise. When his European interlocutor said "But how about the tortoise?" he replied that he was tired of metaphysics and wanted to change the subject. "Force", as an explanation, is no better than the elephant and the tortoise. It is an attempt to get at the "why" of natural processes, not only at the "how". What we observe, to a limited extent, is what happens, and we can arrive at laws according to which observable things happen, but we cannot arrive at a reason for the laws. If we invent a reason, it needs a reason in its turn, and so on. "Force" is a rationalising of natural processes, but a fruitless one since "force" would have to be rationalised also.

When it is said, as it often is, that "force" belongs to the world of experience, we must be careful to understand what can be meant. In the first place, it may be meant that calculations which employ the notion of force work out right in practice. This, broadly speaking, is admitted: no one would suggest that the engineer should alter his methods, or should give up working out stresses and strains. But that does not prove that there are stresses and strains. A medical man works his accounts in guineas, although there are none; he obtains a real payment, though he employs a fictitious coin. Similarly, the engineer is concerned with the question whether his bridge will stand: the fact of experience is that it stands (or does not stand), and the stresses and strains are only a way of explaining what sort of bridge will stand. They are as useful as guineas, but equally imaginary.

But when it is said that force is a fact of experience, there is something quite different that may be meant. It may be meant that we experience force when we experience such things as pressure or muscular exertion. We cannot discuss this contention adequately without going into the relation of physics to psychology, which is a topic we shall consider at length at a later stage. But we may say this much: if you press your finger-tip upon a hard object, you have an experience which you attribute to your finger-tip, but there is a long chain of intermediate causes in nerves and brain. If your finger were amputated you could still have the same experience by a suitable operation on the nerves that formerly connected the finger with the brain,

so that the force between the finger-tip and the hard object, as a fact of experience, may exist when there is no fingertip. This shows that force, in this sense, cannot be what concerns physics.

As the above example illustrates, we do not, in fact, experience many things that we think we experience. This makes it necessary to ask, without too much assurance, in what sense physics can be based upon experience, and what must be the nature of its entities and its inferences if it is to make good its claim to be empirically grounded.

16

PHYSICS AND METAPHYSICS

1928

In this essay, Russell develops his contention that physics has no idea of permanent substance. An elimination of permanent objects in favour of events is, says Russell, the metaphysics that inspired Einstein and modern physics.

Russell makes a number of points that will be familiar to readers of twentieth-century philosophy of science. First, Russell makes the point that a precise statement is probably false. Imprecision is required for truth. But the following response might come from Karl Popper: a true but vague statement is useless. We should aim to eliminate such imprecision. Best science aims at precise, and therefore highly falsifiable, theories rather then true but vague, and therefore less falsifiable, statements (Popper 1934). An interesting twist on this has been supplied by Nancy Cartwright in her influential book *How The Laws of Physics Lie* (1983). She claims that the most useful statements used in science are indeed false (they are lies, so the laws of physics really are lies). Explanatory power does not imply truth. Indeed, she claims, the laws we use in science are literally false in that they are approximations and idealisations. But only approximations will have use. This would suggest replacing Russell's slogan that imprecision is required for truth with the slogan that imprecision is required for use. It may sound instrumentalistic and, although Russell finds that view repugnant, he accepts a kind of instrumentalism in the last paragraph: "knowing about anything is being aware of the behaviour suitable in its presence".

There is in our time a widening gulf between the scientific specialist and the ordinary intelligent man. This was very much less the case in former times. In the eighteenth century everybody of intellectual pretensions, at least in France, was more or less acquainted with Newton, who had the same kind of vogue that Freud has in our day. In the nineteenth century the most sensational piece of science was Darwinian evolution, which was

easily intelligible to any educated man. But during the present century the things of most importance in science have occurred in physics, and have involved for their understanding exceedingly difficult and abstruse mathematical reasoning. So abstruse and difficult has this reasoning become that even distinguished physicists have had to stand aside. The general reader knows that odd things have been discovered about relativity and about atoms, but he is aware that whatever attempts he may have made to find out what it is all about have not met with as much success as he might have hoped. Now all this would not have mattered if the new physics had had for its content merely specialized advances in the later parts of a very well developed science. If that had been the case, the matter could have been left to the expert, and the general public would have confined itself to benevolent neutrality.

The new theories in physics, however, are not only new in mathematical technique or in the results at which they arrive; they are new also in their general point of view, in the character of their reasoning, and in the metaphysic which inspires their hypotheses. This aspect of their work is only accidentally involved in mathematical technicalities, and can with a little trouble be so presented that intelligent laymen, at any rate the younger of them, can understand it and appreciate its importance. There is of course, as in any rapidly growing subject, still controversy and uncertainty, but both sides in a controversy of physicists are advocating theories alien to educated common sense, and the theories of both sides have something in common which distinguishes them from the theories of former times.

Applied physics, that is to say, everything that is used as yet in industry, and even in such modern developments as radio, still uses older conceptions which are on the whole consonant with educated common sense. The newer physics as yet has not become embodied in anything commercially useful. There is no reason to suppose that this state of affairs will continue. The theory upon which radio depends existed for many years before it received practical application. I am not, however, concerned in this article with the impossible attempt to forecast the inventions which may hereafter utilize the modern developments of physics. What I want to do is to suggest the changes in popular metaphysics which must come about, if the new physics is to produce the same kind of influence upon our outlook that was produced in the eighteenth century by Newton, and in the nineteenth by Darwin.

I think that if we were to search for one short phrase to characterize the difference between the newer physics and that of past times, I should choose the following: *The world is not composed of "things"*. To the metaphysician this is no new idea, but in the past the metaphysician could not point to the technique of science as being on his side, and he was therefore unable to

combat the popular metaphysics which survived contentedly alongside of his speculations. Nowadays, physicists, the most hard-headed of mankind, the people associated more than any others with the intellectual and mechanical triumphs that distinguish our epoch, have embodied in their technique this insubstantiality which some of the metaphysicians have so long urged in vain. "We are such stuff as dreams are made on" was once a piece of poetic imagination; now it is among the presuppositions of physics.

It seems probable that if the ideas underlying modern physics ever become apprehended by the ordinary educated and half educated person, they will greatly modify his outlook upon life and politics, not to mention religion and criminal law. I shudder when I think of the revolutions required to adapt ourselves to the ideas of Heisenberg and Schrödinger, which are in many ways more strange than those of Einstein. We should have to begin by altering grammar completely; our grammar is based upon the belief in permanent things. A series of different apparitions, changing gradually as time goes on, are linked together under one name, say John Jones, and are said to constitute one person. If one of these apparitions runs away with a leg of mutton, it is thought right and just that one of the others should be shut within the four walls of a prison cell. If we did not imagine that the person we imprison is the same as the person who stole the leg of mutton, we should be less convinced of our right to shut him up, and if we went on to realize that there are no legs of mutton and no prison walls, we should feel still more reconstruction of our traditional notions to be called for. I do not mean to say that there are no prisons, but as has been said before: "Stone walls do not a prison make, nor iron bars a cage". I think, however, that the poet did not quite know what does make a prison, and if he did, he might have been puzzled to express it is verse. A prison consists of a very large number of matrices, and a matrix is an infinite rectangle of integers. To define a cage is a most complicated problem in a very modern branch of mathematics called topology, which is only properly understood in two universities, one that of Princeton, the other that of Moscow. If any of my readers wishes to know what a cage is, I advise him to write to Professor Veblen of the former university, but I cannot guarantee that the reply will be intelligible.

When I say that there are no "things", I shall perhaps at first convey no definite idea. I will try to explain what I mean. Suppose that on a dark night you see the beam from a searchlight, or a lighthouse, moving about the sky, or sweeping over the sea, the beam in some sense preserves its identity, and yet you do not think of it as a "thing". Or again, suppose you hear "The Star-Spangled Banner" sung; it is one tune, but you would not think of it as a "thing": it is a series of notes, and the notes themselves are essentially brief.

When I say that there are no "things" I mean that tables and chairs and loaves of bread and so on are really just like the beam of light and the song. They are a series of more or less similar phenomena, connected, not by substantial identity, but by certain causal connections.

This, however, is no new suggestion. It was made by the early Buddhists in the time of King Asoka. At this point, however, we arrive at another innovation, an innovation concerned with the idea of cause. The early Buddhists had a firm belief in causation, and upon this belief they rested their metaphysics and their ethics. Newtonian physicists had a firm belief in causation, which was an essential element in eighteenth-century materialism. The man of science, and still more the philosopher interpreting science, would have told you, until lately, that a belief in rigid causation is an indispensable postulate of science. All this seems to have grown doubtful. Causation, like every other traditional notion, appears to be concerned with what happens to things in the mass, not with what happens to them individually. Whether causation is, or is not, rigidly applicable to the most minute phenomena of which we are at present aware is a doubtful question, which is being discussed by German physicists with a detachment as astonishing as it is praiseworthy. Perhaps the reign of law will get itself re-established, perhaps not. Apparently science can adapt itself equally well to either alternative. The laws of science, as they have been known, are concerned with what usually happens approximately, not with what always happens exactly. Men have been misled by the precision of the mathematical instrument into the notion that the mathematical laws of physics were not only precise but exact. Perhaps the difference is not at once clear, but a simple illustration will make it so. If I say that a man is six feet high, my statement is precise, but is practically certain not to be exact, that is to say, I am making an assertion which is not exactly true. If I want to make an assertion which is exactly true, I must make one which is less precise, such as that he is within a quarter of an inch of six feet one way or another. The traditional laws of physics are like the statement that the man is just six feet high: they are precise, but probably not quite true. The newer laws may perhaps be quite true, but they have lost something of the old precision. The older physics was based upon somewhat gross observations of large objects. (I mean by a large object anything bigger than an atom.) It was found that certain precise mathematical laws fitted the behaviour of these large objects within the limits of observation as they then were, and it was assumed that these precise laws were not only precise, but exact. This latter assumption is being dropped, and the older laws are being regarded in the light of statistical probabilities analogous to the statement that if you toss a coin often it will come equally often heads and tails. In fact it seems that everything we see is a statistical probability. A coloured surface, for

example, represents the statistical probability of quantum changes in a certain region. Continuity, which used to be thought to be of the essence of nature, is now thought by some to be only a continuity of probability. The individual phenomena according to these men are discontinuous jumps within atoms, but the probability of a jump occurring in any particular place varies continuously with the place, and this probability is really what we see when we think we see a table or a chair. When Dr. Johnson kicked a stone in order to disprove Berkeley, he was, if we are to believe these physicists, kicking a statistical probability, and the consequent pain in his toe represented the statistical probability of an upset to the atoms in that part of his foot.

Let us not, however, suppose that we are still to be allowed to believe in atoms and electrons, except as convenient fictions like John Jones. An electron consists of a series of sets of phenomena in places where it isn't. What are these phenomena? The only ones of which we have any direct knowledge are our own perceptions. If there are others we know little about them beyond the mathematical laws which they approximately obey.

The normal beliefs of the normal man, as well as all his standards of value, depend upon a radically different outlook. A man wants, we will say, to be a "dynamic personality", but would perhaps feel this desire less strongly if he realized that nothing is "dynamic" and there are no "personalities". When I say this, I mean it of course in a Pickwickian sense. For certain rough practical purposes something not grossly untrue may be conveyed by calling a man a "dynamic personality", but if we were to try to state at all accurately what it really amounts to to be a "dynamic personality", I think the charm exercised by the phrase would largely disappear. I will omit the word "personality" for the moment, since I do not wish to touch upon psychology. The word "dynamic" will afford quite enough food for thought. When we call a person "dynamic", what we mean, as nearly as I can gather, is that he causes a great deal of motion of matter. Seeing that there is no such thing as matter, and that motion has become a completely vague idea, it is evident that the word "dynamic" cannot retain the force that it used to have. Perhaps, however, we might give the word a different meaning. We might say that a "dynamic" person is one who desires some state of affairs considerably different from that which he finds existing, and whose desire is subsequently realized in a manner which common sense regards as partly caused by him. This definition, however, is not quite sufficient. It would, for example, apply to a man who in a drunken fit knocked over a policeman. We do not regard him as a "dynamic personality", because we realize that his desire to knock down a policeman is connected chemically with the properties of alcohol. I think it is essential to our conception of a man as a "dynamic personality" that we should think that we

know the effects of his actions, but that we should not know their causes. There are therefore two different salient criticisms. There is in the first place an older type of criticism illustrated by what we said just now about the drunken man, according to which a man's desires and actions are caused by things outside himself, so that he is not more dynamic than a water pipe; and then there is the newer type of criticism, according to which the old idea of cause is vague and popular. I should not go so far as to say that a man actually is what he seems to be to other people, because it would lead to the questions "What are the other people?" Nevertheless there would be some truth in such a definition. We can define a piece of matter as a series of effects in places where it isn't, and we may say the same of a man in so far as he is regarded as forceful, or dynamic, or influential. The real man is the man regarded psychologically, not the man regarded as having an effect upon his surroundings.

We all naturally believe that matter is characterized by a certain property which Dr. Whitehead calls "pushiness". When you watch a game of billiards, you think that one ball hits another and pushes the other away. This, however, is an illusion. The balls never touch; the outer electrons in the one repel the outer electrons in the other; these outer electrons try to move inward, but are in turn repelled by those that they are thus compelled to approach. There gets to be an unpleasant sense of overcrowding, and there is nothing for the whole lot to do except to move away and look for a place where they do not have such unpleasant neighbours. The whole process is analogous to what happens to a man who likes solitude, and who is perpetually moving away because someone else has built a house within two hundred yards of him. It is not that the electrons are actually in contact during the jostle; it is merely that they dislike the too close proximity of other electrons. All this of course is only roughly speaking. If you try to speak more accurately, you will have to express the whole occurrence in language which replaces electrons by various forms of radiant energy, or by Schrödinger's waves, which are not *in* the ocean but actually constitute the ocean. In any case, the game of billiards cannot be taken as giving the clue to the universe. This is of course the death blow to materialism, for the essence of materialism has always been the belief that the world consisted of billiard balls. Or, to take another illustration, we spoke earlier of the beam of light sent out by a lighthouse on a dark night. To the materialist it is evident that the lighthouse is real, and solid, and tangible, whereas the beam of light in the night sky is merely an evanescent reflection caused by the state of the matter in the lighthouse. The modern view would be more nearly represented by saying that, while the beam of light is indubitably real so long as it lasts, the lighthouse is only an inference from the beam, and a precarious inference at that. If you say that you can go and touch the

walls of the lighthouse, that is merely to say that there can be produced in you a sensation of hardness, a sensation which clearly is in you, not in the lighthouse, so that what you learn of the lighthouse from the sense of touch is really as indirect as what you learn of it by watching the beam of light in the sky. All this as metaphysics is ancient history, but as physics it is modern. For ages there has been a gulf between physics and metaphysics, but the gulf is closing up. Metaphysicians have grown perhaps a little less arrogant, and the physicists have been forced by purely experimental necessities to become increasingly metaphysical. The consequence is that there is a certain solidity in the philosophy to be distilled out of modern physics, which there never was in the philosophy of former times.

It must be confessed, however, that there is far less pomp and circumstance about modern mathematical reasoning than about that of earlier times. The model was set by Euclid, who professed to prove all his assertions with the exception of certain self-evident axioms. We now know that there is no reason to suppose his axioms true, and the consequences which he deduced from them can be known, if at all, only on a basis of experiment. The general view nowadays is that the propositions of Euclid are approximately true of bodies of ordinary size, but not of very small bodies, nor yet of regions comparable in size with the universe, which is supposed to be finite. In Newton's days the prestige of Euclid was undiminished, and Newton's mathematical style is based upon the Greek tradition. Nowadays a mathematical physicist is less impressed than he used to be with the demonstrative powers of his mathematical instrument. His mathematical reasoning is intended as a rule to lead up to some crucial experiment. We might say that mathematical style has altered in a way analogous to that in which prose style has altered. Compare the style of Sir Thomas Browne and that of Mr. H. G. Wells, and you will get some idea of the difference between the mathematical style of Newton and that of, say, Niels Bohr, who first applied the quantum theory to the atom.

It is a curious fact that as the practical applications of science have become more remarkable, its intellectual claims have diminished. I believe that if Empedocles could have a conversation with Einstein, he would be astonished at Einstein's ignorance in regard to many things which he himself had always supposed that he knew. Nor is it necessary to go so far back as Empedocles. Lord Kelvin, I fancy, would suffer the same astonishment. I wish I could think, however, that he would not regard Einstein as an ignoramus. Perhaps the belief that we know has always been the phantasy designed to compensate us for our feelings of helplessness. Knowledge is power, and therefore the phantasy of knowledge is the phantasy of power. Modern science, in proportion as it has given us real power, has diminished the need for the compensatory phantasy of supposed knowledge. We have

become willing to admit our ignorance of the real nature of things, because we feel that we know how to handle them.

The tendency of this attitude is of course to encourage what is called the instrumental theory of knowledge, which is roughly speaking the theory that knowing about anything is being aware of the behaviour suitable in its presence. If you know when to bow to a king, when to call him "Your Majesty", and how to flatter him, you know all that the instrumental theory thinks there is to know; and similarly, if you are able to use a telephone, to make it and repair it, you know all there is to be known about telephones, according to this theory. I confess that the theory is deeply repugnant to me, although I am inclined to think that it may be true. Perhaps the psychological source of this repugnance is the association of knowledge with love. When a man is in love, he is not content to study the outward behaviour of the lady, but wishes to know what, at such a moment, he probably calls her soul, and her inmost thoughts become profoundly interesting to him. A man of science has been in the past a man in love with Nature in this sense, and something of the satisfaction which has been found in the pursuit of knowledge will be lost if knowledge is found to consist merely of tricks for leading Nature to grant her favours unintentionally. Perhaps the more profound intellectual triumphs, upon which in the long run the practical successes of applied science depend, will become psychologically very difficult, if the mystical conception of knowledge as a kind of union of knower and known completely dies out. However, I hold no brief for mysticism, and am merely throwing out a suggestion for what it may be worth.

17

CAUSAL LAWS

1948

In a rather marked contrast to Russell's earlier view, of the preceding papers, in his last seriously philosophical book, *Human Knowledge: Its Scope and Limits* (1948), Russell tells us that the power of science rests in the discovery of causal laws. They appear to have undergone a rehabilitation, in Russell's view – at the very least a rehabilitation in emphasis if not detail. The concept no longer requires elimination. Rather, something like the common metaphysical notion is what legitimises science as causal laws are what allow us to infer one thing about a region of space-time from another. Thus it appears that scientific knowledge assumes causal laws rather than rejects them ("Science assumes causality in some sense").

Such laws, Russell now concedes, may state only what is probable. But probabilistic laws allow the "crude generalisations from which common sense starts", such as "bread nourishes" and "dogs bark". These statements permit exceptions without being thereby falsified: poisoned bread will not nourish, a lazy dog will not bark. Nevertheless, they are true because they are not intended to be universal in scope (see Lowe 1980; 1982 for more on this).

The view of causation Russell offers in this selection was part of a bigger project developing the view that science in general rested upon a number of basic assumptions. He called these assumptions the postulates of science (see Russell 1948*a*). What, asks Russell in the present paper, are the postulates upon which causality in science rests? A postulate of "causal lines" is proposed as the safest such assumption. This is explained in terms of permitted inference, however, which many will see as a purely epistemological concern. The postulate, in turn, must depend upon others, as Russell shows, which concern the uniformity of nature and the existence of natural kinds.

The practical utility of science depends upon its ability to foretell the future. When the atomic bombs were dropped, it was expected that large

numbers of Japanese would die, and they did. Such highly satisfactory results have led, in our day, to an admiration of science, which is due to the pleasure we derive from the satisfaction of our lust for power. The most powerful communities are the most scientific, though it is not the men of science who wield the power conferred by their knowledge. On the contrary, the actual men of science are rapidly sinking into the position of state prisoners, condemned to slave labour by brutal masters, like subject djinns in the Arabian Nights. But we must not waste any more time upon such pleasant topics. The power of science is due to its discovery of causal laws, and it is causal laws that are to occupy us in this chapter.

A "causal law", as I shall use the term, may be defined as a general principle in virtue of which, given sufficient data about certain regions of space-time, it is possible to infer something about certain other regions of space-time. The inference may be only probable, but the probability must be considerably more than a half if the principle in question is to be considered worthy to be called a "causal law".

I have purposely made the above definition very wide. In the first place, the region to which we infer need not be later than those from which we infer. There are, it is true, some laws – notably the second law of thermodynamics – which allow inferences forwards more readily than backwards, but this is not a general characteristic of causal laws. In geology, for example, the inferences are almost all backwards. In the second place, we cannot lay down rules as to the number of data that may be involved in stating a law. If it should ever become possible to state the laws of embryology in terms of physics, enormously complex data would be required. In the third place, the inference may be only to some more or less general characteristic of the inferred event or events. In the days before Galileo it was known that unsupported heavy bodies fall, which was a causal law; but it was not known how fast they fall, so that when a weight was dropped it was impossible to say accurately where it would be after a given lapse of time. In the fourth place, if the law states a high degree of probability it may be almost as satisfactory as if it stated a certainty. I am not thinking of the probability of the law being true; causal laws, like the rest of our knowledge, may be mistaken. What I am thinking of is that some laws *state* probabilities, for example the statistical laws of quantum theory. Such laws supposing them completely true, make inferred events only probable, but this does not prevent them from counting as causal laws according to the above definition.

One advantage of admitting laws which only confer probability is that it enables us to incorporate in science the crude generalizations from which common sense starts, such as "fire burns" "bread nourishes", "dogs bark", or "lions are fierce". All these are causal laws, and all are liable to exceptions,

so that in a given case they confer only probability. The fire on a plum pudding does not burn you, poisoned bread does not nourish, some dogs are too lazy to bark, and some lions grow so fond of their keepers that they cease to be fierce. But in the great majority of cases the above generalizations will be a sound guide in action. There are a large number of such approximate regularities which are assumed in our every-day behaviour, and it is from them that the conception of causal laws arose. Scientific laws, it is true, are no longer so simple: they have become complicated in the endeavour to give them a form in which they are not liable to exceptions. But the old simpler laws remain valid so long as they are only regarded as asserting probabilities.

Causal laws are of two sorts, those concerned with persistence and those concerned with change. The former kind are often not regarded as causal, but this is a mistake. A good example of a law of persistence is the first law of motion. Another example is the persistence of matter. After the discovery of oxygen, when the process of combustion came to be understood, it was possible to regard all matter as indestructible. It has now become doubtful whether this is quite true, but it remains true for most practical purposes. What appears to be more exactly true is the persistence of energy. The gradual development of laws stating persistence started from the common-sense belief, based on pre-scientific experience, that most solid objects continue to exist until they crumble from old age or are destroyed by fire, and that, when this happens, it is possible to suppose that their small parts survive in a new arrangement. It was this pre-scientific point of view that gave rise to the belief in material substance.

Causal laws concerned with change were found by Galileo and Newton to demand statement in terms of acceleration, i.e. change of velocity in magnitude or direction or both. The greatest triumph of this point of view was the law of gravitation, according to which every particle of matter causes in every other an acceleration directly proportional to the mass of the attracting particle and inversely proportional to the square of the distance between them. But Einstein's form of the law of gravitation made it more analogous to the law of inertia, and, in a sense, a law of persistence rather than a law of change. According to Einstein, space-time is full of what we may call hills; each hill grows steeper as you go up, and has a piece of matter at the top. The result is that the easiest route from place to place is one which winds round the hills. The law of gravitation consists in the fact that bodies always take the easiest route, which is what is called a "geodesic". There is a law of cosmic laziness called the "principle of least action", which states that when a body moves from one place to another it will choose the route involving least work. By means of this principle gravitation is absorbed into the geometry of space-time.

The essential laws of change in modern physics are those of quantum theory, which govern transitions from one form of energy to another. An atom can emit energy in the form of light, which then travels on unchanged until it meets another atom, which may absorb the energy of the light. Such interchanges are governed by certain rules, which do not suffice to say what will happen on a given occasion, but can predict, with a very high degree of probability, the statistical distribution of possible happenings among a very large number of interchanges. This is as near as physics can get at present to the ultimate character of causal laws.

Everything that we believe ourselves to know about the physical world depends entirely upon the assumption that there are causal laws. Sensations, and what we optimistically call "perceptions", are events in us. We do not actually see physical objects, any more than we hear electromagnetic waves when we listen to the wireless. What we directly experience might be all that exists, if we did not have reason to believe that our sensations have external causes. It is important, therefore, to inquire into our belief in causation. Is it mere superstitition, or has it a solid foundation?

The question of the justification of our belief in causality belongs to theory of knowledge, and I shall therefore postpone it for the present. My purpose in this Part is the interpretation of science, not an inquiry into the grounds for supposing science valid. Science assumes causality in some sense, and our present question is: in what sense is causality involved in scientific method?[1]

Broadly speaking, scientific method consists in inventing hypotheses which fit the data, which are as simple as is compatible with this requirement, and which make it possible to draw inferences subsequently confirmed by observation. The theory of probability shows that the validity of this process depends upon an assumption which may be roughly stated as the postulate that there are general laws of certain kinds. This postulate, in a suitable form, can make scientific laws probable, but without it they do not even achieve probability. We have therefore to examine this assumption, to find out the most possible form in which it is both effective and possibly true.

If there is no limit to the complexity of possible laws, every imaginable course of events will be subject to laws, and therefore the assumption that there are laws will become a tautology. Take, for example, the numbers of all the taxis that I have hired in the course of my life, and the times when I have hired them. We have here a finite set of integers and a finite number

1 The following pages anticipate, in an abbreviated form, the fuller discussions of Parts V and VI [not included here].

of corresponding times. If n is the number of the taxi that I hired at the time t, it is certainly possible, in an infinite number of ways, to find a function f such that the formula

$$n = f(t)$$

is true for all the values of n and t that have hitherto occurred. An infinite number of these formulae will fail for the next taxi that I hire, but there will still be an infinite number that remain true. By the time I die, it will be possible to close the account, and there will still remain an infinite number of possible formulae, each of which might claim to be a law connecting the number of a taxi with the time when I hire it.

The merit of this example, for my present purpose, is its obvious absurdity. In the sense in which we believe in natural laws we should say that there is no law connecting the n and t of the above formula, and that, if any suggested formula happens to work, that is a mere chance. If we had found a formula that worked in all cases up to the present, we should not expect it to work in the next case. Only a superstitious person whose emotions are involved will believe an induction of this sort; gamblers at Monte Carlo practise inductions which no man of science would sanction. But it is not altogether easy to state the difference between the inductions of the superstitious gambler and the inductions of the prudent man of science. Obviously there is a difference, but in what does it consist? And is the difference such as to affect logical validity, or does it consist merely in a difference as to the obviousness of the appeal to the emotions? Is the faith in scientific method merely the scientist's superstition appropriate to his kind of gambling? These questions, however, belong to the theory of knowledge. For the present I want to discover not why we believe, but what we believe, when we believe in natural laws.

It is customary to speak of *induction* as what is needed to make the truth of scientific laws probable. I do not think that induction, pure and simple, is fundamental. The above example of the numbers of taxis illustrates this. All past observations as to these numbers are compatible with a number of laws of the form $n = f(t)$, and these will, as a rule, give different values for the next n. We cannot therefore use them all for prediction, and in fact we have no inclination to believe in any of them. Generalizing, we may say: Every finite set of observations is compatible with a number of mutually inconsistent laws, all of which have exactly the same inductive evidence in their favour. Therefore pure induction is invalid, and is, moreover, not what we in fact believe.

Whenever inductive evidence seems to us to make a suggested law very probable, the law is one which had suggested itself more or less independently of the evidence, and had seemed to us in some way likely to be true. When

this is the case, subsequent confirmatory evidence is found astonishingly convincing.

This, however, is only partially true. If a law is suggested of which the consequences are very different from what we should expect, and it then is confirmed by observation, we are more prone to believe in it than if its results were commonplace. But in such a case the law itself may seem plausible, although its consequences, when mentioned, are found surprising. Perhaps one of the most important effects of scientific education is to modify the hypotheses that appear *prima facie* probable. It was this cause, not direct negative evidence, that led the belief in witchcraft to decay. If you had a number of outwardly similar boxes, of which some contained gyrostats, and you showed them to a savage, saying that by uttering a magic formula you could make any one of them impossible to turn round, the inductive evidence would soon persuade him that you were right, but a man of educated scientific outlook would search for some other explanation in spite of repeated apparent verifications of your "law".

Induction, moreover, does not validate many of the inferences in which science feels most confidence. We are all convinced that, when a number of people hear a sound simultaneously, their common experience has an external source, which is propagated through the intervening medium by sound waves. There cannot be inductive evidence (unless in some extended sense) for something outside human experience, such as a sound wave. Our experience will be the same whether there really are sound-waves, or, though there are none, auditory sensations occur as they would if there were sound-waves; no inductive evidence can ever favour one of these hypotheses rather than the other. Nevertheless every one in fact accepts the realist alternative – even the idealist philosopher except in his professional moments. We do this on grounds that have nothing to do with induction – partly because we like laws to be as simple as possible, partly because we believe that causal laws must have spatio-temporal continuity, i.e. must not involve action at a distance.

In the establishment of scientific laws experience plays a twofold part. There is the obvious confirming or confuting of a hypothesis by observing whether its calculated consequences take place, and there is the previous experience which determines what hypotheses we shall think antecedently probable. But behind these influences of experience there are certain vague general expectations, and unless these confer a finite *a priori* probability on certain kinds of hypotheses, scientific inferences are not valid. In clarifying scientific method it is essential to give as much precision as possible to these expectations, and to examine whether the success of science in any degree confirms their validity. After being made precise the expectations are, of course, no longer quite what they were while they remained vague, but so long as they remain vague the question whether they are true or false is also vague.

It seems to me that what may be called the "faith" of science is more or less of the following sort: there are formulae (causal laws) connecting events, both perceived and unperceived; these formulae exhibit spatio-temporal continuity, i.e. involve no direct unmediated relation between events at a finite distance from each other; a suggested formula having the above characteristics becomes highly probable if, in addition to fitting in with all past observations, it enables us to predict others which are subsequently confirmed and which would be very improbable if the formula were false.

The justification of this "faith", if any, belongs to theory of knowledge. Our present task is completed in having stated it. But there is still need of some discussion as to the origin and growth of this "faith".

There are various possible postulates which can be taken as the basis of scientific method, but it is difficult to state them with the necessary precision. There is the law of causality; there is the uniformity of nature; there is the reign of law: there is the belief in natural kinds, and Keynes's principle of limited variety; and there is structural constancy with spatio-temporal continuity. It ought to be possible, out of all these somewhat vague assumptions, to distil some definite axiom or axioms which, if true, will confer the desired degree of probability on scientific inferences.

The principle of causality appears in the works of almost all philosophers in an elementary form which it never takes in any advanced science. They suppose science to assume that, given any suitable class of events A, there is always some other class of events B such that every A is "caused" by a B; moreover every event belongs to some such class.

Most philosophers have held that "cause" means something different from "invariable antecedent". The difference may be illustrated by Geulincx's two clocks, which both keep perfect time; when one points to the hour, the other strikes, but we do not think that the one has "caused" the other to strike. A non-scientific Fellow of my College lately remarked in despair: "The barometer has ceased to have any effect on the weather." This was felt to be a joke, but if "cause" meant "invariable antecedent" it would not be. It is supposed that when A is *caused* by B the sequence is not merely a fact, but is in some sense *necessary*. This conception is bound up with the controversy about free will and determinism, summed up by the poet in the following lines:

> There was a young man who said: Damn!
> I learn with regret that I am
> A creature that moves
> In predestinate grooves,
> In short, not a bus, but a tram.

As against this view most empiricists have held that "cause" means nothing but "invariable antecedent". The difficulty of this view, and indeed of any suggestion that scientific laws are of the form "*A* causes *B*", is that such sequences are seldom invariable, and, even if they are invariable in fact, circumstances can easily be imagined which would prevent them from being so. As a rule, if you tell a man he is a silly fool he will be angry, but he may be a saint, or may happen to die of apoplexy before he has time to lose his temper. If you strike a match on a box it usually lights but sometimes it breaks or is damp. If you throw a stone in the air it usually falls down again, but it may be swallowed by an eagle under the impression that it is a bird. If you will to move your arm it usually moves, but not if you are paralysed. In such ways all laws of the form "*A* causes *B*" are liable to exceptions, since something may intervene to prevent the expected result.

Nevertheless, there are reasons, of which the strength will appear in Part VI [not included], for admitting laws of the form "A causes B", provided that we do so with suitable safeguards and limitations. The concept of more or less permanent physical objects, in its common-sense form, involves "substance", and when "substance" is rejected we have to find some other way of defining the identity of a physical object at different times. I think this must be done by means of the concept "causal line". I call a series of events a "causal line", if given some of them, we can infer something about the others without having to know anything about the environment. For example, if my doors and windows are shut, and at intervals I notice my dog asleep on the hearthrug, I infer that he was there, or at least somewhere in the room, at the times when I was not noticing him. A photon which travels from a star to my eye is a series of events obeying an intrinsic law, but ceasing to obey this law when it reaches my eye. When two events belong to one causal line, the earlier may be said to "cause" the later. In this way laws of the form "*A* causes *B*" may preserve a certain validity. They are important in connection both with perception and with persistence of material objects.

It is the possibility of something intervening that has led physics to state its laws in the form of differential equations, which may be regarded as stating what is tending to happen. And as already explained, classical physics, when presented with several causes acting simultaneously, represents the resultant as a vector sum, so that, in a sense, each cause produces its effect as if no other cause were acting. But in fact the whole conception of "cause" is resolved into that of "law". And laws, as they occur in classical physics, are concerned with tendencies at an instant. What actually happens is to be inferred by taking the vector sum of all the tendencies at an instant, and then integrating to find out the result after a finite time.

All empirical laws are inferred from a finite number of observations, eked out by interpolation and extrapolation. The part played by interpolation is not always adequately realized. Take for example the apparent motions of the planets. We assume that, during the day-time, they pursue a smooth course which fits in easily with their observed courses during the preceding and succeeding nights. It would be a possible hypothesis that planets only exist when they are observed, but this would make the laws of astronomy very complicated. If it is objected that planets can be photographed fairly continuously, the same problem arises as regards the photographs: do they exist when no one is looking at them? This again is a question of interpolation, and the interpolation is justified by the fact that it gives the simplest laws compatible with what has been observed.

Exactly the same principle applies to extrapolation. Astronomy makes assertions, not only about what planets have done at all times since there were astronomers, but about what they will do and what they did before there was any one to notice them. This extrapolation is often spoken of as if it involved some principle other than that involved in interpolation, but in fact the principle is one and the same: to choose the simplest law that fits the known facts.

As a postulate, however, this is open to grave objections. "Simple" is a vague conception. Moreover it often happens that a simple law turns out, after a time, to be *too* simple, and that the correct law is more complicated. But in such cases the simple law is usually *approximately* right. If, therefore, we only assert that a law is approximately right, we cannot be convicted of error when some other law is found to be a still better approximation.

The uniformity of nature, which is a principle sometimes invoked, has no definite meaning except in connection with natural laws. If it is already granted that there are natural laws, the principle of the uniformity of nature states that time and place must not appear explicitly in the formulation of laws: the laws must be the same in one part of space-time as in another. This principle may or may not be true, but in any case it is insufficient as a postulate, since it presupposes the existence of laws.

The existence of natural kinds underlies most pre-scientific generalizations, such as "dogs bark" or "wood floats". The essence of a "natural kind" is that it is a class of objects all of which possess a number of properties that are not known to be logically interconnected. Dogs bark and growl and wag their tails, while cats mew and purr and lick themselves. We do not know why all the members of an animal species should share so many common qualities, but we observe that they do, and base our expectations on what we observe. We should be amazed if a cat began to bark.

Natural kinds are not only of biological importance. Atoms and molecules are natural kinds; so are electrons, positrons, and neutrons. Quantum theory has introduced a new form of natural kinds in its discrete series of energy levels. It is now possible to conceive the ultimate structure of the physical world not as a continuous flux, in the manner of conventional hydrodynamics, but in a more Pythagorean fashion, in which models are derived from analogy with a heap of shot. Evolution, which in Darwin's time "broadened slowly down from precedent to precedent", now takes revolutionary leaps by means of mutants or freaks. Perhaps wars and revolutions have made us impatient of gradualness; however that may be, modern scientific theories are much more jolty and jagged than the smooth cosmic stream of ordered progress imagined by the Victorians.

The bearing of all this on induction is of considerable importance. If you are dealing with a property which is likely to be characteristic of a natural kind, you can generalize fairly safely after very few instances. Do seals bark? After hearing half a dozen do so, you confidently answer "yes", because you are persuaded in advance that either all seals bark or no seals bark. When you have found that a few pieces of copper are good conductors of electricity, you unhesitatingly assume that this is true of all copper. In such cases a generalization has a finite a priori probability, and induction is less precarious than in other problems.

Keynes has a postulate by which, in his opinion, inductive arguments might be justified; he calls it the principle of limited variety. It is a form of the assumption of natural kinds. This is one of the expedients in the way of a general assumption which, if true, validates scientific method. I shall have more to say about it at a later stage. What has been said in this chapter is only by way of anticipation.

Part V

ONTOLOGY (II)

Complete metaphysical agnosticism is not compatible with the maintenance of linguistic propositions.

"Language and Metaphysics"

18

VAGUENESS

1923

"Language has many properties which are not shared by things in general", Russell says near the start of his essay on vagueness. His particular concern in this essay is to point out the fallacy of inferring vague things from vague language. Vagueness is a property of words only.

Suppose I say of a colleague that he is bald. It is possible that the truth of my claim is uncertain. Russell attempted to clarify why such things might be uncertain. Being bald is vague but what exactly is this vagueness? Is it something in the world that is vague? Are there vague objects or vague properties of objects? Is it, rather, merely that the *concept* of baldness is vague? Russell supports this latter option. His imagined opponent would be someone who thought there was vagueness in non-linguistic reality: in the world itself.

Thus, on Russell's view, there could be a fact of the matter, in the world, that my colleague has exactly n number of hairs on his head. The problem is that the concept of baldness has vague boundaries, which means that for some number n, of hairs on one's head, it is not clear that the concept of baldness is correctly applicable. Some people are clearly bald and some people are clearly not bald, but there is a class of people for whom it is uncertain whether or not they are bald. We could, therefore, know everything about the state of such a person's head except whether they are bald.

Having argued that it is words, not the world, that are vague, Russell then produces a more general diagnosis of vagueness. It is representations that are vague and concepts are a kind of representation (a linguistic representation of non-linguistic reality). Things such as tables and chairs and men's heads are neither vague nor precise. They just *are*. Both vagueness and precision concern the relation between the representation and the thing represented. Further, they can come in degrees. One concept can be more vague than another without either being wholly vague or precise. Once

one considers such things, one can see just how many of our concepts are vague. Thus, we may only be able to know vaguely how many hairs are on my colleague's head because what counts as a hair might only be vaguely defined. What about a hair that is only just breaking through the surface of the skin? What about velutinous hairs? What about a hair transplant or wig?

The vagueness of language is inevitable. Our concepts are derived from the evidence of our senses and the knowledge derived from the senses is vague. Russell then gives a definition of vagueness, which is a perilous task as it appears to be an attempt to make vagueness precise. His strategy avoids this. Instead, he offers a precise definition of accuracy, then says that vagueness is when something is not accurate.

Timothy Williamson, whose own book is now the best detailed source on this subject, says that Russell's paper "set the agenda for most subsequent work" (1994: 52). Williamson identifies ten theses in the paper: (1) only representations are vague; (2) all language is vague; (3) there is higher-order vagueness; (4) vagueness invalidates classical logic; (5) vagueness is not generality; (6) accuracy is isomorphism; (7) precision is one-one correlation; (8) meaning is a special case of representing; (9) precision diminishes probability, and (10) vagueness is a natural phenomenon. Williamson discusses and criticises each thesis in turn so I would direct readers there, after reading this paper, if they want a detailed critique.

Reflection on philosophical problems has convinced me that a much larger number than I used to think, or than is generally thought, are connected with the principles of symbolism, that is to say, with the relation between what means and what is meant. In dealing with highly abstract matters it is much easier to grasp the symbols (usually words) than it is to grasp what they stand for. The result of this is that almost all thinking that purports to be philosophical or logical consists in attributing to the world the properties of language. Since language really occurs, it obviously has all the properties common to all occurrences, and to that extent the metaphysic based upon linguistic considerations may not be erroneous. But language has many properties which are not shared by things in general, and when these properties intrude into our metaphysic it becomes altogether misleading. I do not think that the study of the principles of symbolism will yield any *positive* results in metaphysics, but I do think it will yield a great many negative results by enabling us to avoid fallacious inferences from symbols to things. The influence of symbolism on philosophy is mainly unconscious; if it were conscious it would do less harm. By studying the principles of symbolism we can learn not to be unconsciously influenced by language, and in this way can escape a host of erroneous notions.

Vagueness, which is my topic tonight,[1] illustrates these remarks. You will no doubt think that, in the words of the poet: "Who speaks of vagueness should himself be vague." I propose to prove that all language is vague and that therefore my language is vague, but I do not wish this conclusion to be one that you could derive without the help of the syllogism. I shall be as little vague as I know how to be if I am to employ the English language. You all know that I invented a special language with a view to avoiding vagueness, but unfortunately it is unsuited for public occasions. I shall therefore, though regretfully, address you in English, and whatever vagueness is to be found in my words must be attributed to our ancestors for not having been predominantly interested in logic.

There is a certain tendency in those who have realized that words are vague to infer that things also are vague. We hear a great deal about the flux and the continuum and the unanalysability of the Universe, and it is often suggested that as our language becomes more precise, it becomes less adapted to represent the primitive chaos out of which man is supposed to have evolved the cosmos. This seems to me precisely a case of the fallacy of verbalism – the fallacy that consists in mistaking the properties of words for the properties of things. Vagueness and precision alike are characteristics which can only belong to a representation, of which language is an example. They have to do with the relation between a representation and that which it represents. Apart from representation, whether cognitive or mechanical, there can be no such thing as vagueness or precision; things are what they are, and there is an end of it. Nothing is more or less what it is, or to a certain extent possessed of the properties which it possesses. Idealism has produced habits of confusion even in the minds of those who think that they have rejected it. Ever since Kant there has been a tendency in philosophy to confuse knowledge with what is known. It is thought that there must be some kind of identity between the knower and the known, and hence the knower infers that the known also is muddle-headed. All this identity of knower and known, and all this supposed intimacy of the relation of knowing, seems to me a delusion. Knowing is an occurrence having a certain relation to some other occurrence, or groups of occurrences, or characteristic of a group of occurrences, which constitutes what is said to be known. When knowledge is vague, this does not apply to the knowing as an occurrence; as an occurrence it is incapable of being either vague or precise, just as all other occurrences are. Vagueness in a cognitive occurrence is a characteristic of its relation to that which is known, not a characteristic of the occurrence in itself.

1 Read before the Jowett Society, Oxford.

Let us consider the various ways in which common words are vague, and let us begin with such a word as "red". It is perfectly obvious, since colours form a continuum, that there are shades of colour concerning which we shall be in doubt whether to call them red or not, not because we are ignorant of the meaning of the word "red", but because it is a word the extent of whose application is essentially doubtful. This, of course, is the answer to the old puzzle about the man who went bald. It is supposed that at first he was not bald, that he lost his hairs one by one, and that in the end he was bald; therefore, it is argued, there must have been one hair the loss of which converted him into a bald man. This, of course, is absurd. Baldness is a vague conception; some men are certainly bald, some are certainly not bald, while between them there are men of whom it is not true to say they must either be bald or not bald. The law of excluded middle is true when precise symbols are employed, but it is not true when symbols are vague, as, in fact, all symbols are. All words describing sensible qualities have the same kind of vagueness which belongs to the word "red". This vagueness exists also, though in a lesser degree, in the quantitative words which science has tried hardest to make precise, such as a metre or a second. I am not going to invoke Einstein for the purpose of making these words vague. The metre, for example, is defined as the distance between two marks on a certain rod in Paris, when that rod is at a certain temperature. Now the marks are not points, but patches of a finite size, so that the distance between them is not a precise conception. Moreover, temperature cannot be measured with more than a certain degree of accuracy, and the temperature of a rod is never quite uniform. For all these reasons the conception of a metre is lacking in precision. The same applies to a second. The second is defined by relation to the rotation of the earth, but the earth is not a rigid body, and two parts of the earth's surface do not take exactly the same time to rotate; moreover all observations have a margin of error. There are some occurrences of which we can say that they take less than a second to happen, and others of which we can say that they take more, but between the two there will be a number of occurrences of which we believe that they do not all last equally long, but of none of which we can say whether they last more or less than a second. Therefore, when we say an occurrence lasts a second, all that it is worth while to mean is that no possible accuracy of observation will show whether it lasts more or less than a second.

Now let us take proper names. I pass by the irrelevant fact that the same proper name often belongs to many people. I once knew a man called Ebenezer Wilkes Smith, and I decline to believe that anybody else ever had this name. You might say, therefore, that here at last we have discovered an unambiguous symbol. This, however, would be a mistake. Mr. Ebenezer Wilkes Smith was born, and being born is a gradual process. It would seem

natural to suppose that the name was not attributable before birth; if so, there was doubt, while birth was taking place, whether the name was attributable or not. If it be said that the name was attributable before birth, the ambiguity is even more obvious, since no one can decide how long before birth the name became attributable. Death also is a process: even when it is what is called instantaneous, death must occupy a finite time. If you continue to apply the name to the corpse, there must gradually come a stage in decomposition when the name ceases to be attributable, but no one can say precisely when this stage has been reached. The fact is that all words are attributable without doubt over a certain area, but become questionable within a penumbra, outside which they are again certainly not attributable. Someone might seek to obtain precision in the use of words by saying that no word is to be applied in the penumbra, but unfortunately the penumbra itself is not accurately definable, and all the vaguenesses which apply to the primary use of words apply also when we try to fix a limit to their indubitable applicability. This has a reason in our physiological constitution. Stimuli which for various reasons we believe to be different produce in us indistinguishable sensations. It is not clear whether the sensations are really different like their stimuli and only our power to discriminate between sensations is deficient, or whether the sensations themselves are sometimes identical in relevant respects even when the stimuli differ in relevant respects. This is a kind of question which the theory of quanta at some much later stage in its development may be able to answer, but for the present it may be left in doubt. For our purpose it is not the vital question. What is clear is that the knowledge that we can obtain through our sensations is not as finegrained as the stimuli to those sensations. We cannot see with the naked eye the difference between two glasses of water of which one is wholesome while the other is full of typhoid bacilli. In this case a microscope enables us to see the difference, but in the absence of a microscope the difference is only inferred from the differing effects of things which are sensibly indistinguishable. It is this fact that things which our senses do not distinguish produce different effects – as, for example, one glass of water gives you typhoid while the other does not – that has led us to regard the knowledge derived from the senses as vague. And the vagueness of the knowledge derived from the senses infects all words in the definition of which there is a sensible element. This includes all words which contain geographical or chronological constituents, such as "Julius Caesar", "the twentieth century", or "the solar system".

There remains a more abstract class of words: first, words which apply to all parts of time and space, such as "matter" or "causality"; secondly, the words of pure logic. I shall leave out of discussion the first class of words, since all of them raise great difficulties, and I can scarcely imagine a human

being who would deny that they are all more or less vague. I come therefore to the words of pure logic, words such as "or" and "not". Are these words also vague or have they a precise meaning?

Words such as "or" and "not" might seem, at first sight, to have a perfectly precise meaning: "*p* or *q*" is true when *p* is true, true when *q* is true, and false when both are false. But the trouble is that this involves the notions of "true" and "false"; and it will be found, I think, that all the concepts of logic involve these notions, directly or indirectly. Now "true" and "false" can only have a *precise* meaning when the symbols employed – words, perceptions, images, or what not – are themselves precise. We have seen that, in practice, this is not the case. It follows that every proposition that can be framed in practice has a certain degree of vagueness; that is to say, there is not one definite fact necessary and sufficient for its truth, but a certain region of possible facts, any one of which would make it true. And this region is itself ill-defined: we cannot assign to it a definite boundary. This is the difference between vagueness and generality. A proposition involving a general concept – e.g. "This is a man" – will be verified by a number of facts, such as "This" being Brown or Jones or Robinson. But if "man" were a precise idea, the set of possible facts that would verify "this is a man" would be quite definite. Since, however, the conception "man" is more or less vague, it is possible to discover prehistoric specimens concerning which there is not, even in theory, a definite answer to the question, "Is this a man?" As applied to such specimens, the proposition "this is a man" is neither definitely true nor definitely false. Since all non-logical words have this kind of vagueness, it follows that the conceptions of truth and falsehood, as applied to propositions composed of or containing non-logical words, are themselves more or less vague. Since propositions containing non-logical words are the substructure on which logical propositions are built, it follows that logical propositions also, so far as we can know them, become vague through the vagueness of "truth" and "falsehood". We can see an ideal of precision, to which we can approximate indefinitely; but we cannot attain this ideal. Logical words, like the rest, when used by human beings, share the vagueness of all other words. There is, however, less vagueness about logical words than about the words of daily life, because logical words apply essentially to symbols, and may be conceived as applying rather to possible than to actual symbols. We are capable of imagining what a precise symbolism would be, though we cannot actually construct such a symbolism. Hence we are able to *imagine* a precise meaning for such words as "or" and "not". We can, in fact, see precisely what they would mean if our symbolism were precise. All traditional logic habitually assumes that precise symbols are being employed. It is therefore not applicable to this terrestrial life, but only to an imagined celestial existence. Where, however,

this celestial existence would differ from ours, so far as logic is concerned, would be not in the nature of what is known, but only in the accuracy of our knowledge. Therefore, if the hypothesis of a precise symbolism enables us to draw any inferences as to what is symbolized, there is no reason to distrust such inferences merely on the ground that our actual symbolism is not precise. We are able to conceive precision; indeed, if we could not do so, we could not conceive vagueness, which is merely the contrary of precision. This is one reason why logic takes us nearer to heaven than most other studies. On this point I agree with Plato. But those who dislike logic will, I fear, find my heaven disappointing.

It is now time to tackle the definition of vagueness. Vagueness, though it applies primarily to what is cognitive, is a conception applicable to every kind of representation – for example, a photograph, or a barograph. But before defining vagueness it is necessary to define accuracy. One of the most easily intelligible definitions of accuracy is as follows: One structure is an accurate representation of another when the words describing the one will also describe the other by being given new meanings. For example, "Brutus killed Caesar" has the same structure as "Plato loved Socrates", because both can be represented by the symbol "*xRy*", by giving suitable meanings to *x* and *R* and *y*. But this definition, though easy to understand, does not give the essence of the matter, since the introduction of words describing the two systems is irrelevant. The exact definition is as follows: One system of terms related in various ways is an accurate representation of another system of terms related in various other ways if there is a one-one relation of the terms of the one to the terms of the other, and likewise a one-one relation of the relations of the one to the relations of the other, such that, when two or more terms in the one system have a relation belonging to that system, the corresponding terms of the other system have the corresponding relation belonging to the other system. Maps, charts, photographs, catalogues, etc. all come within this definition in so far as they are accurate.

Per contra, a representation is *vague* when the relation of the representing system to the represented system is not one-one, but one-many. For example, a photograph which is so smudged that it might equally represent Brown or Jones or Robinson is vague. A small-scale map is usually vaguer than a large-scale map, because it does not show all the turns and twists of the roads, rivers, etc. so that various slightly different courses are compatible with the representation that it gives. Vagueness, clearly, is a matter of degree, depending upon the extent of the possible differences between different systems represented by the same representation. Accuracy, on the contrary, is an ideal limit.

Passing from representation in general to the kinds of representation that are specially interesting to the logician, the representing system will

consist of words, perceptions, thoughts, or something of the kind, and the would-be one-one relation between the representing system and the represented system will be *meaning*. In an accurate language, meaning would be a one-one relation; no word would have two meanings, and no two words would have the same meaning. In actual languages, as we have seen, meaning is one-many. (It happens often that two words have the same meaning, but this is easily avoided, and can be assumed not to happen without injuring the argument.) That is to say, there is not only one object that a word means, and not only one possible fact that will verify a proposition. The fact that meaning is a one-many relation is the precise statement of the fact that all language is more or less vague. There is, however, a complication about language as a method of representing a system, namely that words which mean relations are not themselves relations, but just as substantial or unsubstantial as other words.[2] In this respect a map, for instance, is superior to language, since the fact that one place is to the west of another is represented by the fact that the corresponding place on the map is to the left of the other; that is to say, a relation is represented by a relation. But in language this is not the case. Certain relations of higher order are represented by relations, in accordance with the rules of syntax. For example, "*A* precedes *B*" and "*B* precedes *A*" have different meanings, because the order of the words is an essential part of the meaning of the sentence. But this does not hold of elementary relations; the word "precedes", though it means a relation, is not a relation. I believe that this simple fact is at the bottom of the hopeless muddle which has prevailed in *all* schools of philosophy as to the nature of relations. It would, however, take me too far from my present theme to pursue this line of thought.

It may be said: How do you know that all knowledge is vague, and what does it matter if it is? The case which I took before, of two glasses of water, one of which is wholesome while the other gives you typhoid, will illustrate both points. Without calling in the microscope it is obvious that you cannot distinguish the wholesome glass of water from the one that will give you typhoid, just as, without calling in the telescope it is obvious that what you see of a man who is 200 yards away is vague compared to what you see of a man who is 2 feet away; that is to say, many men who look quite different when seen close at hand look indistinguishable at a distance, while men who look different at a distance never look indistinguishable when seen close at hand. Therefore, according to the definition, there is less vagueness in the near appearance than in the distant one. There is still less

2 A word is a class of series, and both classes and series are logical fictions. See *Analysis of Mind*, ch. x; *Introduction to Mathematical Philosophy*, ch. xvii.

vagueness about the appearance under the microscope. It is perfectly ordinary facts of this kind that prove the vagueness of most of our knowledge, and lead us to infer the vagueness of all of it.

It would be a great mistake to suppose that vague knowledge must be false. On the contrary, a vague belief has a much better chance of being true than a precise one, because there are more possible facts that would verify it. If I believe that so-and-so is tall, I am more likely to be right than if I believe that his height is between 6 ft. 2 in. and 6 ft. 3 in. In regard to beliefs and propositions, though not in regard to single words, we can distinguish between accuracy and precision. A belief is *precise* when only one fact would verify it; it is *accurate* when it is both precise and true. Precision diminishes the likelihood of truth, but often increases the pragmatic value of a belief if it is true – for example, in the case of the water that contained the typhoid bacilli. Science is perpetually trying to substitute more precise beliefs for vague ones; this makes it harder for a scientific proposition to be true than for the vague beliefs of uneducated persons to be true, but makes scientific truth better worth having if it can be obtained.

Vagueness in our knowledge is, I believe, merely a particular case of a general law of physics, namely the law that what may be called the appearances of a thing at different places are less and less differentiated as we get further away from the thing. When I speak of "appearances" I am speaking of something purely physical – the sort of thing, in fact, that, if it is visual, can be photographed. From a close-up photograph it is possible to infer a photograph of the same object at a distance, while the contrary inference is much more precarious. That is to say, there is a one-many relation between distant and close-up appearances. Therefore the distant appearance, regarded as a representation of the close-up appearance, is vague according to our definition. I think all vagueness in language and thought is essentially analogous to this vagueness which may exist in a photograph. My own belief is that most of the problems of epistemology, in so far as they are genuine, are really problems of physics and physiology; moreover, I believe that physiology is only a complicated branch of physics. The habit of treating knowledge as something mysterious and wonderful seems to me unfortunate. People do not say that a barometer "knows" when it is going to rain; but I doubt if there is any essential difference in this respect between the barometer and the meteorologist who observes it. There is only one philosophical theory which seems to me in a position to ignore physics, and that is solipsism. If you are willing to believe that nothing exists except what you directly experience, no other person can prove that you are wrong, and probably no valid arguments exist against your view. But if you are going to allow any inferences from what you directly experience to other entities, then physics supplies the safest form of such inferences. And I believe that

(apart from illegitimate problems derived from misunderstood symbolism) physics, in its modern forms, supplies materials for answers to all philosophical problems that are capable of being answered, except the one problem raised by solipsism, namely: Is there any valid inference ever from an entity experienced to one inferred? On this problem, I see no refutation of the sceptical position. But the sceptical philosophy is so short as to be uninteresting; therefore it is natural for a person who has learnt to philosophise to work out other alternatives, even if there is no very good ground for regarding them as preferable.

19

PHYSICS AND NEUTRAL MONISM

1927

This selection appeared as a chapter of Russell's book *The Analysis of Matter* (1927). Though Russell's Hegelianism was long since rejected, the position he advocates here, in relation to mind and matter, has all the appearance of a Hegelian synthesis. Materialism and idealism, on their own, lead to contradictions. Only neutral monism can rid us of such contradictions and, Russell argues further, modern physics supports such a position.

The version Russell presents is idiosyncratic, however. Mental and physical events are said to be compresent. Events are compresent when they have at least some shared spatio-temporal location or, as Russell says in an earlier chapter (Russell 1927: ch. 28), "they overlap in space-time". But compresence of events does not seem to entail identity, notwithstanding Russell's discussion of logical and spatio-temporal interpenetration, so it could be argued that it does not establish monism. Mental and physical event identity is the usual way of characterising monism. But consider the two events of a sphere rotating and simultaneously increasing in heat. These two events evidently have identical spatio-temporal locations so they are compresent. But they are clearly not identical events in any credible sense, even though the same substance is the subject of the two events. One event essentially involves a change in temperature while the other does not. If these two events can be compresent without being identical, why can't mental and physical events be so? And if they need not be identical, monism – hence neutral monism – is not established.

Note also that, in this selection, Russell further pursues his empiricist programme of constructing the physical world out of events. A position of event neutral monism seems particularly to accord with his reduction of the category of substance.

As this paper is also a contribution to the philosophy of mind, it is worth drawing attention to the following passage that has a remarkable similarity to some more recent, celebrated contributions

> to that field: "It is obvious that a man who can see knows things which a blind man cannot know; but a blind man can know the whole of physics. Thus the knowledge which other men have and he has not is not part of physics". Compare this quotation with some more recent claims in the philosophy of mind (Nagel 1974; Jackson 1986).

In this chapter, I wish to define the outcome of our analysis in regard to the old controversy between materialism and idealism, and to make it clear wherein our theory differs from both. So long as the views set forth in previous chapters [not included] are supposed to be either materialistic or idealistic, they will seem to involve inconsistencies, since some seem to tend in the one direction, some in the other. For example, when I say that my percepts are in my head, I shall be thought materialistic; when I say that my head consists of my percepts and other similar events, I shall be thought idealistic. Yet the former statement is a logical consequence of the latter.

Both materialism and idealism have been guilty, unconsciously and in spite of explicit disavowals, of a confusion in their imaginative picture of matter. They have thought of the matter in the external world as being represented by their percepts when they see and touch, whereas these percepts are really part of the matter of the percipient's brain. By examining our percepts it is possible – so I have contended – to infer certain formal mathematical properties of external matter, though the inference is not demonstrative or certain. But by examining our percepts we obtain knowledge which is not purely formal as to the matter of our brains. This knowledge, it is true, is fragmentary, but so far as it goes it has merits surpassing those of the knowledge given by physics.

The usual view would be that by psychology we acquire knowledge of our "minds", but that the only way to acquire knowledge of our brains is to have them examined by a physiologist, usually after we are dead, which seems somewhat unsatisfactory. I should say that what the physiologist sees when he looks at a brain is part of his own brain, not part of the brain he is examining. The feeling of paradox about this view comes, I should say, from wrong views of space. It is true that what we see is not located where our percept of our own brain would be located if we could see our own brain; but this is a question of perceptual space, not of the space of physics. The space of physics is connected with causation in a manner which compels us to hold that our percepts are in our brains, if we accept the causal theory of perception, as I think we are bound to do. To say that two events have no spatio-temporal separation is to say that they are compresent; to say that

they have a small separation is to say that they are connected by causal chains all of which are short. The percept must therefore be nearer to the sense-organ than to the physical object, nearer to the nerve than to the sense-organ, and nearer to the cerebral end of the nerve than to the other end. This is inevitable, unless we are going to say that the percept is not in space-time at all. It is usual to hold that "mental" events are in time but not in space; let us ask ourselves whether there is any ground for this view as regards percepts.

The question whether percepts are located in physical space is the same as the question of their causal connection with physical events. If they can be effects and causes of physical events, we are bound to give them a position in physical space-time in so far as interval is concerned, since interval was defined in causal terms. But the real question is as to "compresence" in the sense of Chapter XXVIII [not included]. Can a mental event be compresent with a physical event? If yes, then a mental event has a position in the space-time order; if no, then it has no such position. This, therefore, is the crucial question.

When I maintain that a percept and a physical event can be compresent, I am not maintaining that a percept can have to a piece of matter the sort of relation which another piece of matter would have. The relation of compresence is between a percept and a physical *event*, and physical events are not to be confounded with pieces of matter. A piece of matter is a logical structure composed of events; the causal laws of the events concerned, and the abstract logical properties of their spatio-temporal relations, are more or less known, but their intrinsic character is not known. Percepts fit into the same causal scheme as physical events, and are not known to have any intrinsic character which physical events cannot have, since we do not know of any intrinsic character which could be incompatible with the logical properties that physics assigns to physical events. There is therefore no ground for the view that percepts cannot be physical events, or for supposing that they are never compresent with other physical events.

The fact that mental events admittedly have temporal relations has much force, now that time and space are so much less distinct than they were. It has become difficult to hold that mental events, though in time, are not in space. The fact that their relations to each other can be viewed as only temporal is a fact which they share with any set of events forming the biography of one piece of matter. Relatively to axes moving with the percipient's brain, the interval between two percepts of his which are not compresent should always be temporal, if his percepts are in his head. But the interval between simultaneous percepts of different percipients is of a different kind; and their whole causal environment is such as to make us call

this interval space-like. I conclude, then, that there is no good ground for excluding percepts from the physical world, but several strong reasons for including them. The difficulties that have been supposed to stand in the way seem to me to be entirely due to wrong views as to the physical world, and more particularly as to physical space. The wrong views as to physical space have been encouraged by the notion that the primary qualities are objective, which has been held imaginatively by many men who would have emphatically repudiated it so far as their explicit thought was concerned.

I hold, therefore, that two simultaneous percepts of one percipient have the relation of compresence out of which spatio-temporal order arises. It is almost irresistible to go a step further, and say that any two simultaneous perceived contents of a mind are compresent, so that all our conscious mental states are in our heads. I see as little reason against this extension as against the view that percepts can be compresent. A percept differs from another mental state, I should say, only in the nature of its causal relation to an external stimulus. Some relation of this kind no doubt always exists, but with other mental states the relation may be more indirect, or may be only to some state of the body, more particularly the brain. "Unconscious" mental states will be events compresent with certain other mental states, but not having those effects which constitute what is called awareness of a mental state. However, I have no wish to go further into psychology than is necessary, and I will pursue this topic no longer, but return to matters of more concern to physics.

The point which concerns the philosophy of matter is that the events out of which we have been constructing the physical world are very different from matter as traditionally conceived. Matter was expected to be impenetrable and indestructible. The matter that we construct is impenetrable as a result of definition: the matter in a place is all the events that are there, and consequently no other event or piece of matter can be there. This is a tautology, not a physical fact; one might as well argue that London is impenetrable because nobody can live in it except one of its inhabitants. Indestructibility, on the other hand, is an empirical property, believed to be approximately but not exactly possessed by matter. I mean by indestructibility, not conservation of mass, which is known to be only approximate, but conservation of electrons and protons. At present it is not known whether an electron and a proton sometimes enter into a suicide pact or not,[1] but

1 It is thought highly probable that they do. See Dr Jeans, "Recent Developments of Cosmical Physics", *Nature*, 4, December 1926.

there is certainly no known reason why electrons and protons should be indestructible.

Electrons and protons, however, are not the stuff of the physical world: they are elaborate logical structures composed of events, and ultimately of particulars, in the sense of Chapter XXVII [not included]. As to what the events are that compose the physical world, they are, in the first place, percepts, and then whatever can be inferred from percepts by the methods considered in Part II [not included]. But on various inferential grounds we are led to the view that a percept in which we cannot perceive a structure nevertheless often has a structure, i.e. that the apparently simple is often complex. We cannot therefore treat the *minimum visible* as a particular, for both physical and psychological facts may lead us to attribute a structure to it – not merely a structure in general, but such and such a structure.

Events are neither impenetrable nor indestructible. Space-time is constructed by means of co-punctuality, which is the same thing as spatio-temporal interpenetration. Perhaps it is not unnecessary to explain that spatio-temporal interpenetration is quite a different thing from logical interpenetration, though it may be suspected that some philosophers have been led to favour the latter as a result of the arguments for the former. We are accustomed to imagining that numerical diversity involves spatio-temporal separation; hence we tend to think that, if two diverse entities are in one place, they cannot be wholly diverse, but must be also in some sense one. It is this combination that is supposed to constitute logical interpenetration. For my part, I do not think that logical interpenetration can be defined without obvious self-contradiction; Bergson, who advocates it, does not define it. The only author I know of who has dealt seriously with its difficulties is Bradley, in whom, quite consistently, it led to a thorough-going monism, combined with the avowal that, in the end, all truth is self-contradictory. I should myself regard this latter result as a refutation of the logic from which it follows. Therefore, while I respect Bradley more than any other advocate of interpenetration, he seems to me, in virtue of his ability, to have done more than any other philosopher to disprove the kind of system which he advocated. However that may be, the spatio-temporal interpenetration which is used in constructing space-time order is quite different from logical interpenetration. Philosophers have been slaves of space and time in the imaginative application of their logic. This is partly due to Euler's diagrams and the notion that the traditional *A, E, I, O* were elementary forms of propositions and the confounding of "x is a β" with "all α's are β's". All this led to a confusion between classes and individuals, and to the inference that individuals can interpenetrate because classes can overlap. I do not suggest explicit confusions of this sort, but only that

traditional elementary logic, taught in youth, is an almost fatal barrier to clear thinking in later years, unless much time is spent in acquiring a new technique.

On the question of the material out of which the physical world is constructed, the views advocated in this volume have, perhaps, more affinity with idealism than with materialism. What are called "mental" events, if we have been right, are part of the material of the physical world, and what is in our heads is the mind (with additions) rather than what the physiologist sees through his microscope. It is true that we have not suggested that all reality is mental. The positive arguments in favour of such a view, whether Berkeleyan or German, appear to me fallacious. The sceptical argument of the phenomenalists, that, whatever else there may be, we cannot know it, is much more worthy of respect. There are, in fact, if we have been right, three grades of certainty. The highest grade belongs to my own percepts; the second grade to the percepts of other people; the third to events which are not percepts of anybody. It is to be observed, however, that the second grade belongs only to the percepts of those who can communicate with me, directly or indirectly, and of those who are known to be closely analogous to people who can communicate with me. The percepts of minds, if such there be, which are not related to mine by communication – e.g. minds in other planets – can have, at best, only the third grade of certainty, that, namely, which belongs to the apparently lifeless physical world.

The events which are not perceived by any person who can communicate with me, supposing they have been rightly inferred, have a causal connection with percepts, and are inferred by means of this connection. Much is known about their structure, but nothing about their quality.

While, on the question of the stuff of the world, the theory of the foregoing pages has certain affinities with idealism – namely, that mental events are part of that stuff, and that the rest of the stuff resembles them more than it resembles traditional billiard-balls – the position advocated as regards scientific laws has more affinity with materialism than with idealism. Inference from one event to another, where possible, seems only to acquire exactness when it can be stated in terms of the laws of physics. There are psychological laws, physiological laws, and chemical laws, which cannot at present be reduced to physical laws. But none of them is exact and without exceptions; they state tendencies and averages rather than mathematical laws governing minimum events. Take, for example, the psychological laws of memory. We cannot say: At 12.55 G.M.T. on such and such a day, *A* will remember the event *e* – unless, indeed, we are in a position to remind him of it at that moment. The known laws of memory belong to an early stage of science – earlier than Kepler's laws or Boyle's law. We can say

that, if *A* and *B* have been experienced together, the recurrence of *A tends* to cause a recollection of *B*, but we cannot say that it is sure to do so, or that it will do so in one assignable class of cases and not in another. One supposes that, to obtain an exact causal theory of memory, it would be necessary to know more about the structure of the brain. The ideal to be aimed at would be something like the physical explanation of fluorescence, which is a phenomenon in many ways analogous to memory. So far as causal laws go, therefore, physics seems to be supreme among the sciences, not only as against other sciences of matter, but also as against the sciences that deal with life and mind.

There is, however, one important limitation to this. We need to know in what physical circumstances such-and-such a percept will arise, and we must not neglect the more intimate qualitative knowledge which we possess concerning mental events. There will thus remain a certain sphere which will be outside physics. To take a simple instance: physics might, ideally, be able to predict that at such a time my eye would receive a stimulus of a certain sort; it might be able to trace the physical properties of the resulting events in the eye and the brain, one of which is, in fact, a visual percept; but it could not itself give us the knowledge that one of them is a visual percept. It is obvious that a man who can see knows things which a blind man cannot know; but a blind man can know the whole of physics. Thus the knowledge which other men have and he has not is not part of physics.

Although there is thus a sphere excluded from physics, yet physics, together with a "dictionary", gives, apparently, all causal knowledge. One supposes that, given the physical characteristics of the events in my head, the "dictionary" gives the "mental" events in my head. This is by no means a matter of course. The whole of the foregoing theory of physics might be true without entailing this consequence. So far as physics can show, it might be possible for different groups of events having the same structure to have the same part in causal series. That is to say, given the physical causal laws, and given enough knowledge of an initial group of events to determine the purely physical properties of their effects, it might nevertheless be the case that these effects could be qualitatively of different sorts. If that were so, physical determinism would not entail psychological determinism, since, given two percepts of identical structure but diverse quality, we could not tell which would result from a stimulus known only as to its physical, i.e. structural, properties. This is an unavoidable consequence of the abstractness of physics. If physics is concerned only with structure, it cannot, *per se*, warrant inferences to any but the structural properties of events. Now it may be a fact that (e.g.) the structure of visual percepts is very different

from that of tactual percepts; but I do not think such differences could be established with sufficient strictness and generality to enable us to say that such-and-such a stimulus must produce a visual percept, while such another must produce a tactual percept.

On this matter, we must, I think, appeal to evidence which is partly psychological. We do know, as a matter of fact, that we can, in normal circumstances, more or less infer the percept from the stimulus. If this were not the case, speaking and writing would be useless. When the lessons are read, the congregation can follow the words in their own Bibles. The differences in their "thoughts" meanwhile can be connected causally, at least in part, with differences in their past experience, and these are supposed to make themselves effective by causing differences in the structure of brains. All this seems sufficiently probable to be worth taking seriously; but it lies outside physics, and does not follow from the causal autonomy of physics, supposing this to be established even for human bodies. It will be observed that what we are now considering is the converse of what is required for the inference from perception to physics. What is wanted there is that, given the percept, we should be able to infer, at least partially, the structure of the stimulus – or at any rate that this should be possible when a sufficient number of percepts are given. What we want now is that, given the *structure* of the stimulus (which is all that physics can give), we should be able to infer the *quality* of the percept – with the same limitations as before. Whether this is the case or not, is a question lying outside physics; but there is reason to think that it is the case.

The aim of physics, consciously or unconsciously, has always been to discover what we may call the causal skeleton of the world. It is perhaps surprising that there should be such a skeleton, but physics seems to prove that there is, particularly when taken in conjunction with the evidence that percepts are determined by the physical character of their stimuli. There is reason – though not quite conclusive reason – for regarding physics as causally dominant, in the sense that, given the physical structure of the world, the qualities of its events, in so far as we are acquainted with them, can be inferred by means of correlations. We have thus in effect a psycho-cerebral parallelism, although the interpretation to be put upon it is not the usual one. We suppose that, given sufficient knowledge, we could infer the qualities of the events in our heads from their physical properties. This is what is really meant when it is said, loosely, that the state of the mind can be inferred from the state of the brain. Although I think that this is probably true, I am less anxious to assert it than to assert, what seems to me much more certain, that its truth does not follow from the causal autonomy of physics or from physical determinism as applied to all matter, including

that of living bodies. This latter result flows from the abstractness of physics, and belongs to the philosophy of physics. The other proposition, if true, cannot be established by considering physics alone, but only by a study of percepts for their own sakes, which belongs to psychology. Physics studies percepts only in their cognitive aspect; their other aspects lie outside its purview.

Even if we reject the view that the quality of events in our heads can be inferred from their structure, the view that physical determinism applies to human bodies brings us very near to what is most disliked in materialism. Physics may be unable to tell us what we shall hear or see or "think", but it can, on the view advocated in these pages, tell us what we shall say or write, where we shall go, whether we shall commit murder or theft, and so on. For all these are bodily movements, and thus come within the scope of physical laws. We are often asked to concede that the beauties of poetry or music cannot result from physical laws. I should concede that the beauty does not result from physics, since beauty depends in part upon intrinsic quality; if it were, as some writers on æsthetics contend, solely a matter of form, it would come within the scope of physics, but I think these writers do not realize what an abstract affair form really is. I should concede also that the *thoughts* of Shakespeare or Bach do not come within the scope of physics. But their thoughts are of no importance to us: their whole social efficacy depended upon certain black marks which they made on white paper. Now there seems no reason to suppose that physics does not apply to the making of these marks, which was a movement of matter just as truly as the revolution of the earth in its orbit. In any case, it is undeniable that the socially important part of their thought had a one–one relation to certain purely physical events, namely the occurrence of the black marks on the white paper. And no one can doubt that the causes of our emotions when we read Shakespeare or hear Bach are purely physical. Thus we cannot escape from the universality of physical causation.

This, however, is perhaps not quite the last word on the subject. We have seen that, on the basis of physics itself, there may be limits to physical determinism. We know of no laws as to when a quantum transaction will take place or a radio-active atom will break down. We know fairly well what will happen *if* anything happens, and we know statistical averages, which suffice to determine macroscopic phenomena. But if mind and brain are causally interconnected, very small cerebral differences must be correlated with noticeable mental differences. Thus we are perhaps forced to descend into the region of quantum transactions, and to desert the macroscopic level where statistical averages obtain. Perhaps the electron jumps when it likes; perhaps the minute phenomena in the brain which make all

the difference to mental phenomena belong to the region where physical laws no longer determine definitely what must happen. This, of course, is merely a speculative possibility; but it interposes a veto upon materialistic dogmatism. It may be that the progress of physics will decide the matter one way or other; for the present, as in so many other matters, the philosopher must be content to await the progress of science.

20

LANGUAGE AND METAPHYSICS

1940

Russell famously attacked ordinary language philosophy in "The cult of 'common usage' " (Russell 1953). In this selection, a chapter from *An Inquiry into Meaning and Truth* (1940), Russell offers a more sober and moderate assessment of the relation between the world and language: "with sufficient caution, the properties of language may help us understand the structure of the world". Russell puts this into practice and shows how the existence of universals can be inferred from language, using arguments that already will be familiar to readers of Part III. Hence, we cannot claim, as some "modern philosophers" do, that we can know about language but nothing else: "complete metaphysical agnosticism is not compatible with the maintenance of linguistic propositions".

Just over a decade later, Russell became embroiled in his final philosophical crusade, against "ordinary language" philosophy. This should not be regarded as at all in conflict with his earlier view that language can help us know the world. While his former view could be summarised as the claim that something about the world can be discovered through language alone, he regarded the ordinary language movement as claiming that everything it is possible to discover can be discovered through language alone. Furthermore, he saw it as the claim not just that language sufficed, but that "ordinary", non-philosophical language sufficed. For the full details of this controversy, see *Papers* 11, Part XI.

In the present chapter I propose to consider whether anything, and, if so, what, can be inferred from the structure of language as to the structure of the world. There has been a tendency, especially among logical positivists, to treat language as an independent realm, which can be studied without regard to non-linguistic occurrences. To some extent, and in a limited field, this separation of language from other facts is possible; the detached study of logical syntax has undoubtedly yielded valuable results. But I think it is easy to exaggerate what can be achieved by syntax alone. There is, I think,

a discoverable relation between the structure of sentences and the structure of the occurrences to which the sentences refer. I do not think the structure of non-verbal facts is wholly unknowable, and I believe that, with sufficient caution, the properties of language may help us to understand the structure of the world.

With regard to the relation of words to non-verbal facts, most philosophers can be divided into three broad types:

(*a*) Those who infer properties of the world from properties of language. These are a very distinguished party; they include Parmenides, Plato, Spinoza, Leibniz, Hegel, and Bradley.

(*b*) Those who maintain that knowledge is only of words. Among these are the Nominalists and some of the Logical Positivists.

(*c*) Those who maintain that there is knowledge not expressible in words, and use words to tell us what this knowledge is. These include the mystics, Bergson, and Wittgenstein; also certain aspects of Hegel and Bradley.

Of these three parties, the third can be dismissed as self-contradictory. The second comes to grief on the empirical fact that we can know what words occur in a sentence, and that this is not a verbal fact, although it is indispensable to the verbalists. If, therefore, we are confined to the above three alternatives, we must make the best of the first.

We may divide our problem into two parts: first, what is implied by the correspondence theory of truth, in the measure in which we have accepted this theory? Second, is there anything in the world corresponding to the distinction between different parts of speech, as this appears in a logical language?

As regards "correspondence", we have been led to the belief that, when a proposition is true, it is true in virtue of one or more occurrences which are called its "verifiers". If it is a proposition containing no variable, it cannot have more than one verifier. We may confine ourselves to this case, since it involves the whole of the problem with which we are concerned. We have thus to inquire whether, given a sentence (supposed true) which contains no variable, we can infer anything as to the structure of the verifier from that of the sentence. In this inquiry we shall presuppose a logical language.

Consider first a group of sentences which all contain a certain name (or a synonym for it). These sentences all have something in common. Can we say that their verifiers also have something in common?

Here we must distinguish according to the kind of name concerned. If *W* is a complete group of qualities, such as we considered in the last chapter [not included], and we form a number of judgments of perception, such as "*W* is red", "*W* is round", "*W* is bright", etc., these all have one single

verifier, namely *W*. But if I make a number of true statements concerning a given shade of colour *C*, they all have different verifiers. These all have a common part *C*, just as the statements have a common part "*C*". It will be seen that here, as in the last chapter [not included], we are led to a view which, syntactically, is scarcely distinguishable from the subject-predicate view, from which it differs only in that it regards the "subject" as a bundle of compresent qualities. We may state what has just been said as follows: given a number of subject-predicate sentences expressing judgments of perception, such as "this is red", if they all have the same subject they all have the same verifier, which is what the subject designates; if they all have the same predicate, the verifiers all have a common part, which is what the predicate designates.

This theory is not applicable to such a sentence as "*A* is to the left of *B*", where "*A*" and "*B*" are names for two parts of my visual field. So far as "*A*" and "*B*" are concerned, we considered this sentence sufficiently in the last chapter. What I now wish to examine is the question: what, if anything, is common to the verifiers of a number of different sentences of the form "*A* is to the left of *B*"?

The question involved is the old question of "universals". We might have investigated this question in connection with predicates – say "red is a colour", or "high C is a sound". But since we have explained the more apparently obvious subject-predicate sentence – e.g. "this is red" – as really not subject-predicate sentences, we shall find it more convenient to discuss "universals" in connection with relations.

Sentences – except object-words used in an exclamatory manner – require words other than names. Such words, generically, we call "relation-words", including predicates as words for monadic relations. The definition, as explained in Chapter VI [not included], is syntactical: a "name" is a word which can occur significantly in an atomic sentence of any form; a "relation-word" is one which can occur in some atomic sentences, but only in such as contain the appropriate number of names.

It is generally agreed that language requires relation-words; the question at issue is: "what does this imply as regards the verifiers of sentences?" A "universal" may be defined as "the meaning (if any) of a relation-word". Such words as "if" and "or" have no meaning in isolation, and it may be that the same is true of relation-words.

It may be suggested (erroneously, as I think and shall try to prove) that we need not assume universals, but only a set of stimuli to the making of one of a set of similar noises. The matter is, however, not quite straightforward. A defender of universals, if attacked, might begin in this way: "you say that two cats, because they are similar, stimulate the utterance of two

similar noises which are both instances of the word 'cat'. But the cats must be *really* similar to each other, and so must the noises. And if they are *really* similar, it is impossible that 'similarity' should be just a word. It is a word which you utter on certain occasions, namely, when there *is* similarity. Your tricks and devices", he will say, "may seem to dispose of other universals, but only by putting all the work on to this one remaining universal, similarity; of that you cannot get rid, and therefore you might as well admit all the rest."

The question of universals is difficult, not only to decide, but to formulate. Let us consider "*A* is to the left of *B*". Places in the momentary visual field, as we have seen, are absolute, and are defined by relation to the centre of the field of vision. They may be defined by the two relations right-and-left, up-and-down; these relations, at any rate, suffice for topological purposes. In order to study momentary visual space, it is necessary to keep the eyes motionless and attend to things near the periphery as well as in the centre of the field of vision. If we are not deliberately keeping our eyes motionless, we shall look directly at whatever we notice; the natural way to examine a series of places is to look at each in turn. But if we want to study what we can see at one moment, this method will not do, since a given physical object, as a visual datum, is different when it is seen directly and when it is far from the centre of the field. In fact, however, this makes very little difference. We cannot escape from the fact that visual positions form a two-dimensional series, and that such series demand dyadic asymmetrical relations. The view we take as to colours makes no difference in this respect.

It seems that there is no escape from admitting relations as parts of the non-linguistic constitution of the world; similarity, and perhaps also asymmetrical relations, cannot be explained away, like "or" and "not", as belonging only to speech. Such words as "before" and "above", just as truly as proper names, "mean" something which occurs in objects of perception. It follows that there is a valid form of analysis which is not that of whole and part. We can perceive *A-before-B* as a whole, but if we perceived it *only* as a whole we should not know whether we had seen it or *B-before-A*. The whole-and-part analysis of the datum *A-before-B* yields only *A* and *B*, and leaves out "before". In a logical language, therefore, there will be *some* distinctions of parts of speech which correspond to objective distinctions.

Let us examine once more the question whether asymmetrical relations are needed as well as similarity; and let us take, for the purpose, "*A* is above *B*", where "*A*" and "*B*" are proper names of events. We shall suppose that we perceive that *A* is above *B*. Now it is clear, to begin with a trivial point, that we do not need the word "below" as well as the word "above"; either alone suffices. I shall therefore assume that our language contains no word "below". The whole percept, *A-above-B*, resembles other percepts

C-above-D, *E-above-F*, etc., in a manner which makes us call them all facts of vertical order. So far, we do not need a concept "above"; we may have merely a group of similar occurrences, all called "vertical orders", i.e. all causing a noise similar to "above". So far, we can do with only similarity.

But now we must consider asymmetry. When you say "*A* is above *B*", how does your hearer know that you have not said "*B is above A*"? In exactly the same way as *you* know that *A* is above *B*; he perceives that the noise "*A*" precedes the noise "*B*".

Thus the vital matter is the distinction between *A-first-and-then-B*, *B-first-and-then-A*; or, in writing, between *AB* and *BA*. Consider, then, the two following shapes: *AB* and *BA*. I want to make it clear that I am speaking of just these, not of others like them. Let S_1 be the proper name of the first shape, S_2 that of the second; let A_1, A_2 be the proper names of the two *A*'s, and B_1, B_2 of the two *B*'s. Then S_1, S_2 each consist of two parts, and one part of S_1 is closely similar to one part of S_2, while the other part is closely similar to the other part. Moreover, the ordering relation is the same in both cases. Nevertheless, the two wholes are not very similar. Perhaps asymmetry could be explained in this way: given a number of *A*'s and a number of *B*'s, arranged in pairs, the resulting wholes fall into two classes, members of the same class being closely similar to each other, while members of different classes are very dissimilar. If we give the proper names S_3, S_4 to the following two shapes: *AB* and *BA*, then it is obvious that S_1 and S_3 are very similar, and so are S_2 and S_4, but S_1 and S_3 are not very similar to S_2 and S_4. (Observe that, in describing S_1 and S_2, we shall have to say: S_1 consists of A_1 before B_1, S_2 consists of B_2 before A_2.) Perhaps in this way it is possible to explain asymmetry in terms of similarity, though the explanation is not very satisfactory.

Assuming that we can, in the above manner or in some other, get rid of all universals except similarity, it remains to be considered whether similarity itself could be explained away.

We will consider this in the simplest possible case. Two patches of red (not necessarily of exactly the same shade) are similar, and so are two instances of the word "red". Let us suppose that we are being shown a number of coloured discs and asked to name their colours – say in a test for colour-blindness. We are shown two red discs in succession, and each time we say "red". We have been saying that, in the primary language, similar stimuli produce similar reactions; our theory of meaning has been based on this. In our case, the two discs are similar, and the two utterances of the word "red" are similar. Are we saying the *same* thing about the discs and about the utterances when we say the discs are similar and when we say the utterances are similar? or are we only saying similar things? In the former case, similarity is a true universal; in the latter case, not. The difficulty, in

the latter case, is the endless regress; but are we sure that this difficulty is inseparable? We shall say, if we adopt this alternative: if *A* and *B* are perceived to be similar, and *C* and *D* are also perceived to be similar, that means that *AB* is a whole of a certain kind and *CD* is a whole of the same kind; i.e. since we do not want to define the kind by a universal, *AB* and *CD* are *similar* wholes. I do not see how we are to avoid an endless regress of the vicious kind if we attempt to explain similarity in this way.

I conclude, therefore, though with hesitation, that there are universals, and not merely general words. Similarity, at least, will have to be admitted; and in that case it seems hardly worth while to adopt elaborate devices for the exclusion of other universals.

It should be observed that the above argument only proves the necessity of the word "similar", not of the word "similarity".

Some propositions containing the word "similarity" can be replaced by equivalent propositions containing the word "similar", while other cannot. These latter need not be admitted. Suppose, for example, I say "similarity exists". If "exists" means what it does when I say "the President of the United States exists", my statement is nonsense. What I can mean may, to begin with, be expressed in the statement: "there are occurrences which require for their verbal description sentences of the form '*a* is similar to *b*' ". But this linguistic fact seems to imply a fact about the occurrences described, namely the sort of fact that is asserted when I say "*a* is similar to *b*". When I say "similarity exists", it is this fact about the world, not a fact about language, that I mean to assert. The word "yellow" is necessary because there are yellow things; the word "similar" is necessary because there are pairs of similar things. And the similarity of two things is as truly a non-linguistic fact as the yellowness of one thing.

We have arrived, in this chapter, at a result which has been, in a sense, the goal of all our discussions. The result I have in mind is this: that complete metaphysical agnosticism is not compatible with the maintenance of linguistic propositions. Some modern philosophers hold that we know much about language, but nothing about anything else. This view forgets that language is an empirical phenomenon like another, and that a man who is metaphysically agnostic must deny that he knows when he uses a word. For my part, I believe that, partly by means of the study of syntax, we can arrive at considerable knowledge concerning the structure of the world.

21

THE PRINCIPLE OF INDIVIDUATION

1947

The final paper in this collection is also one of the most difficult and abstract, though Russell begins with a characteristically concise summary of the problem: "what is meant by a 'particular'?". Is it something constituted by its qualities or is it ultimate and indefinable? Russell dislikes the latter view, which has all the difficulties of the traditional conception of substance. Instead, Russell tries to overcome the problems of the first view and traces a notion of a particular as a complex of compresent quality instances. The upshot is that we have no need for proper names. Names of qualities and words for compresence and for spatial and temporal relations are all that is needed for a complete description of the world. This account brings together a number of the ideas that already have appeared in earlier papers.

Russell makes a statement that will please contemporary realist metaphysicians: "to allow grammar to dictate our metaphysics is now generally recognized to be dangerous". This has a different emphasis from the argument of the previous selection, though, as explained there, it cannot be identified as an inconsistency in Russell's position.

I shall be concerned in what follows with a very old problem, much discussed by the scholastics, but still, in our day, far from being definitively solved. The problem, in its broadest and simplest terms, is this: "How shall we define the diversity which makes us count objects as two in a census?" We may put the same problem in words that look different, e.g. "what is meant by a 'particular'?" or "what sort of objects can have proper names?"

Three views have been influentially advocated.

First, a particular is constituted by qualities: when all its qualities have been enumerated, it is fully defined. This is the view of Leibniz.

Second, a particular is defined by its spatio-temporal position. This is the view of Thomas Aquinas as regards material substances.

Third, numerical diversity is ultimate and indefinable. This, I think, would be the view of most modern empiricists if they took the trouble to have a definite view.

Before we discuss the question itself, there are some preliminaries by which the problem may be simplified.

The second of the above three theories is reducible to either the first or the third according to the way in which it is interpreted. If we take a Newtonian view, according to which there actually are points, then two different points are exactly alike in all their qualities, and their diversity must be that bare numerical diversity contemplated in the third theory. If, on the other hand, we take – as every one now does – a relational view of space, the second theory will have to say: "if *A* and *B* differ in spatio-temporal position, then *A* and *B* are two." But here there are difficulties. Suppose *A* is a shade of colour: it may occur in a number of places and yet be only one. Therefore our *A* and *B* must not be qualities, or, if they are, they must be qualities that never recur. If they are not qualities or bundles of qualities, they must be particulars of the sort contemplated in our third theory; if they are qualities or bundles of qualities, it is the first of our three theories that we are adopting. Our second theory, therefore, may be ignored.

Various considerations, of which the theory of relativity has been the most influential, have led to the conclusion that, if there are particulars, they are not persistent entities which move, but events, as limited in temporal duration as in spatial extent. It is therefore events, not particles, that we have to consider in seeking a theory of particulars. I assume – though this is not essential – that a single event may last for a finite time and occupy a finite volume; and that two events may overlap both in space and time. If "events" are to be satisfactory in interpreting common sense and physics, they must have certain further characteristics. If *A* and *B* are events, and *A* wholly precedes *B*, *A* and *B* must be diverse. If *A, B, C* are events, and *A* wholly precedes *B* and *B* wholly precedes *C*, then *A* must wholly precede *C*. We have to consider what is the simplest way of defining the class "events" so that events may have these properties. Provided these properties attach to the members of one class, we are at liberty to define the class as we choose. But we shall, of course, wish to minimize the element of unverifiable assumption that will be involved.

We have to ask ourselves: what is meant by an "instance"? Take some definite shade of colour, which we will call "*C*". Let us assume that it is a shade of one of the colours of the rainbow, so that it occurs wherever there is a rainbow or a solar spectrum. On each occasion of its occurrence, we say that there is an "instance" of *C*. Is each instance an unanalysable particular, of which *C* is a quality? Or is each instance a complex of qualities of which *C* is one? The former is the third of the above theories; the latter is the first.

There are difficulties in either view. Taking first the view that an instance of *C* is an unanalysable particular, we find that we encounter all the familiar difficulties connected with the traditional notion of "substance". The particular cannot be defined or recognized or known; it is something serving the merely grammatical purpose of providing the subject in a subject–predicate sentence such as "this is red". And to allow grammar to dictate our metaphysic is now generally recognized to be dangerous.

There are other difficulties. For instance, is it logically necessary that an instance of redness should be red? One would naturally say yes, but a little reflection will bring hesitation. A particular, on the theory that we are considering, is not *defined* by its predicates; it is, in fact, not *defined* at all. It is supposed to be simple, and, if so, whatever predicates it may have cannot be logically necessary. If *A* is a particular which is an instance of redness, we shall have to say that this is a purely contingent fact, and that *A* might just as well be an instance of hardness, or loudness, or what you will. This seems quite contrary to what we have in mind when we speak of an "instance".

It is difficult to see how something so unknowable as such a particular would have to be can be required for the interpretation of empirical knowledge. The notion of a substance as a peg on which to hang predicates is repugnant, but the theory that we have been considering cannot avoid its objectionable features. I conclude, therefore, that we must, if possible, find some other way of defining "events".

But when we abandon particulars in the sense which we have just decided to reject, we are faced with the difficulty of finding something that will not be repeated. A simple quality, such as the shade of colour *C*, cannot be expected to occur only once. We shall seek to escape this difficulty by considering a "complex" of qualities. What I mean will be most easily understood if stated in psychological terms. If I see something and at the same time hear something else, my visual and auditory experiences have a relation which I call "compresence". If at the same moment I am remembering something that happened yesterday and anticipating with dread a forthcoming visit to the dentist, my remembering and anticipating are also "compresent" with my seeing and hearing. We can go on to form the whole group of my present experiences. That is to say, given any group of experiences which are all compresent, if I can find any other experience which is compresent with all of them I add it to the group, and I go on until there is no further experience which is compresent with each and all of the members of the group. I thus arrive at a group having the two properties: (*a*) that any two members of the group are compresent, (*b*) that nothing outside the group is compresent with every member of the group. Such a group I shall call a "complex of compresence".

Such a complex I suppose to consist of constituents most of which, in the natural course of events, may be expected to be members of many other complexes. The shade of colour *C*, we supposed, recurs every time anybody sees a rainbow distinctly. My recollection may be qualitatively indistinguishable from a recollection that I had yesterday. My apprehension of dental pain may be just what I felt before my last visit to the dentist. All these items of the complex of compresence may occur frequently, and are not essentially dated. That is to say, if *A* is one of them, and *A* precedes (or follows) *B*, we have no reason to suppose that *A* and *B* are not identical.

Have we any reason, either logical or empirical, to believe that a complex of compresence, as a whole, cannot be repeated? Let us, in the first place, confine ourselves to one person's experience. My visual field is very complex, though probably not infinitely complex. Every time I move my eyes, the visual qualities connected with a given object which remains visible undergo changes: what I see out of the corner of my eyes looks different from what is in the centre of my field of vision. If it is true, as some maintain, that my memory is coloured by my whole past experience, then it follows logically that my recollections cannot be exactly similar on two different occasions; even if we reject this doctrine, such exact similarity seems very improbable.

From such considerations I think we ought to conclude that the exact repetition of my total momentary experience, which is what, in this connection, I call a "complex of compresence", is not logically impossible, but is empirically so exceedingly improbable that we may assume its non-occurrence. In that case, a complex of compresence will, so far as one person's experience is concerned, have the formal properties required of "events", i.e. if *A, B, C* are complexes of compresence, then if *A* wholly precedes *B*, *A* and *B* are not identical; and if *B* also wholly precedes *C*, then *A* wholly precedes *C*. We thus have the requisites for defining the time-order in one person's experience.

This, however, is only part, and not the most difficult part, of what we have to accomplish. We have to extend space–time order beyond one person's experience to the experiences of different people and to the physical world. In regard to the physical world, especially, this is very difficult.

So long as we confine ourselves to one person's experience, we need only concern ourselves with time. But now we have also to take account of space. That is to say, we have to find a definition of "events" which shall insure that each event has, not merely a unique temporal position, but a unique spatio–temporal position.

So long as we confine ourselves to experiences, there is no fresh difficulty of a serious kind. It may be taken as virtually certain, on empirical grounds, that my visual field, whenever my eyes are open, is not exactly similar to

that of any one else. If *A* and *B* are looking simultaneously at the same scene, there are differences of perspective; if they change places, *A* will not see exactly what *B* was seeing, because of differences of eyesight, changes of lighting meanwhile, and so on. In short, the reasons for supposing that no total momentary experience of *A* is ever exactly like some total momentary experience of *B* are of the same kind as the reasons for supposing that no two total momentary experiences of *A* are ever exactly alike.

This being granted, we can establish a spatial order among percipients by means of the laws of perspective, provided there is any physical object that all the percipients concerned are perceiving. If there is not, a process by means of intermediate links can reach the same result. There are of course complications and difficulties, but they are not such as concern our subject at all closely, and we may safely ignore them.

What can be said about the purely physical world is hypothetical, since physics gives no information except as to structure. But there are reasons for supposing that, at every place in physical space–time, there is at every moment a multiplicity of occurrences, just as there is in a mind. "Compresence", which I take to have a merely ostensive definition, appears in psychology as "simultaneity in one experience", but in physics as "overlapping in space–time". If, as I maintain, my thoughts are in my head, it is obvious that these are different aspects of one relation. However, this identification is inessential to my present argument.

When I look at the stars on a clear night, each star that I see has an effect on me, and has an effect on the eye before it has an effect on the mind. It follows that, at the surface of the eye, something causally connected with each visible star is happening. The same considerations apply to ordinary objects seen in daylight. At this moment I can see white pages covered with writing, some books, an oval table, innumerable chimneys, green trees, clouds, and blue sky. I can see these things because there is a chain of physical causation from them to my eyes and thence to the brain. It follows that what is going on at the surface of my eye is as complex as my visual field, in fact as complex as the whole of what I can see. This complexity must be physical, not merely physiological or psychological; the optic nerve could not make the complex responses that it does make except under the influence of equally complex stimuli. We must hold that, wherever the light of a certain star penetrates, something connected with that star is happening. Therefore in a place where a telescope photographs many millions of stars, many millions of things must be happening, each connected with its own star. These things are only "experienced" in places where there is a recording nervous system, but that they happen in other places also can be shown by cameras and dictaphones. There is therefore no difficulty of principle in constructing "complexes of compresence", where there are no percipients,

on the same principles as we employed in dealing with momentary experiences.

Abandoning speculations about the physical world, about which our knowledge is very limited, let us return to the world of experience. The view which I am suggesting, as preferable to the assumption of such wholly colourless particulars as points of space or particles of matter, may be expressed as follows:

There is a relation, which I call "compresence", which holds between two qualities when one person experiences them simultaneously – for example, between high *C* and vermilion when you hear one and see the other. We can form groups of qualities having the following two properties: (*a*) any two members of the group are compresent; (*b*) given anything not a member of the group, there is at least one member of the group with which it is not compresent. Any one such complete group of compresent qualities constitutes a single complex whole, defined when its constituents are given, but itself a unit, not a class. That is to say, it is something which exists, not merely because its constituents exist, but because, in virtue of being compresent, they constitute a single structure. One such structure may be called a "total momentary experience".

Total momentary experiences, as opposed to qualities, have time relations possessing the desired characteristics. I can see blue yesterday, red today, and blue again tomorrow. Therefore, so far as qualities are concerned, blue is before red and red is before blue, while blue, since it occurs yesterday and tomorrow, is before itself. We cannot therefore construct, out of qualities alone, such a relation as will generate a series. But out of total momentary experiences we can do this, provided no total momentary experience ever exactly recurs. That this does not happen is an empirical proposition, but, so far as our experience goes, a well-grounded one. I regard it as a merit in the above theory that it gets rid of what would otherwise be synthetic *a priori* knowledge. That, if *A* precedes *B*, *B* does not precede *A*, and that, if *A* precedes *B* and *B* precedes *C*, then *A* precedes *C*, are synthetic propositions; moreover, as we have just seen, they are not true if *A* and *B* and *C* are qualities. By making such statements (in so far as they are true) empirical generalizations, we overcome what would otherwise be a grave difficulty in the theory of knowledge.

I come back now to the conception of "instance". An "instance" of a quality, as I wish to use the word, is a complex of compresent qualities of which the quality in question is one. In some cases this view seems natural. An instance of "man" has other qualities besides humanity: he is white or black, French or English, wise or foolish, and so on. His passport enumerates enough of his characteristics to distinguish him from the rest of the human race. Each of these characteristics, presumably, exists in many other

instances. There are baby giraffes who have the height mentioned in his passport, and parrots who have the same birthday as he has. It is only the assemblage of qualities that makes the instance unique. Every man, in fact, is defined by such an assemblage of qualities, of which humanity is only one.

But when we come to points of space, instants of time, particles of matter, and such stock-in-trade of abstract science, we feel as if a particular could be a "mere" instance, differentiated from other instances by relations, not by qualities. To some degree, we think this of less abstract objects: we say "as like as two peas", suggesting that between two peas there are no qualitative differences. We think also that two patches of colour may be *merely* two, and may differ only numerically. This way of thinking, I maintain, is a mistake. I should say that, when the same shade of colour exists in two places, it is one, not two; there are, however, two complexes, in which the shade of colour is combined with the qualities that give position in the visual field. People have become so obsessed with the relativity of spatial position in physics that they have become oblivious of the absoluteness of spatial position in the visual field. At every moment, what is in the centre of my field of vision has a quality that may be called "centrality"; what is to the right is "dexter", what to the left "sinister", what above "superior", what below "inferior". These are *qualities* of the visual datum, not relations. It is the complex consisting of one such quality combined with a shade of colour that is distinct from the complex consisting of the same shade elsewhere. In short, the multiplicity of instances of a given shade of colour is formed exactly as the multiplicity of instances of humanity is formed, namely by the addition of other qualities.

As for points, instants, and particles, in so far as they are not logical fictions similar considerations apply. Take first instants. It will be found that what I call a "momentary total experience" has all the formal properties required of an "instant" in my biography. And it will be found that, where there is only matter, the "complex of compresence" may serve to define an instant of Einsteinian local time, or to define a "point-instant" in cosmic space–time. Points in perceptual space are defined without any trouble, since the qualities of up-and-down, right-or-left, in their various degrees, have already all the properties that we require of "points". It is indeed this fact, together with perception of depth, that has led us to place such emphasis on the spatial characteristics of the world.

I do not think "particles" can be dealt with quite in the above manner. In any case, they are no longer part of the fundamental apparatus of physics. They are, I should say, strings of events interconnected by the law of inertia. They are no longer indestructible, and have become merely convenient approximations.

I come now to a possible objection to the above theory, which was advanced by Arnauld against Leibniz. If a "particular" is really a complex of qualities, then the statement that such-and-such a particular has such-and-such a quality must, when true, be analytic; at least, so it would seem. Leibniz held (1) that every proposition has a subject and a predicate; (2) that a substance is defined by the total of its predicates; (3) that the soul is a substance. It followed that everything that can be truly said of a given soul consists in mentioning some predicate which is one of those that constitute the given soul. "Caesar", for example, was a collection of predicates, one of which was "crossing the Rubicon". He was therefore compelled by logic to cross the Rubicon, and there is no such thing as contingency or free will. Leibniz ought, on this point, to have agreed with Spinoza, but he chose not to, for reasons discreditable either to his intellect or to his moral character. The question is: Can I avoid agreeing with Spinoza without equal discredit?

What we have to consider is a subject–predicate proposition expressing a judgment of perception, such as "this is red". What is "this"? Clearly it is not my whole momentary experience; I am not saying "one of the qualities that I am at present experiencing is redness". The word "this" may be accompanied by a gesture, indicating that I mean what is in a certain direction, say the centre of my visual field. In that case, the core of what I am saying may be expressed by "centrality and redness overlap spatially in my present visual field". It is to be observed that, within the large complex of my total momentary experience, there are smaller complexes constituted by spatial compresence in perceptual space. Whatever quality I see in a certain direction has perceptual-spatial compresence with the visual quality constituting that direction. It would seem that the word "this", accompanied by a gesture, is equivalent to a description, viz. "what is occupying the centre of my visual field". To say that this description applies to redness is to say something which clearly is not analytic. But since it employs a description instead of a name, it is not quite what we set out to consider.

We were considering what sort of thing could have the formal properties that we attribute to "events". Such a thing must happen in only one time and place; it must not recur, either on another occasion or in another location. So far as time and *physical* space are concerned, these conditions are satisfied by the total "complex of compresence", whether this consists of my momentary experiences or of a full group of overlapping physical qualities. (I call such a group "full" when, if anything is added, the members will no longer be all compresent.) But when we come to consider perceptual space, we have to take account of another relation, that of compresence in perceptual space. This gives us a sub-group of my total momentary experiences, which may be roughly described as "all that I see over there". It is by means of such sub-groups that we define the particularity of a patch of colour

when I am simultaneously seeing another patch of just the same shade. Such sub-groups count as "events" when the space with which we are concerned is that of perception and not that of physics.

With these preliminaries, let us examine the question of proper names.

It seems preposterous to maintain that "Caesar crossed the Rubicon" is an analytic proposition. But if it is not, what do we mean by "Caesar"?

Taking Caesar as he was, without the limitations due to our ignorance, we may say that he was a series of events, each event being a momentary total experience. If we were to define "Caesar" by enumerating these events, the crossing of the Rubicon would have to come in our list, and "Caesar crossed the Rubicon" would be analytic. But in fact we do not define "Caesar" in this way, and we cannot do so, since we do not know all his experiences. What happens in fact is more like this: Certain series of experiences have certain characteristics which make us call such a series a "person". Every person has a number of characteristics that are peculiar to him; Caesar, for example, had the name "Julius Caesar". Suppose *P* is some property which has belonged to only one person; then we can say: "I give the name "*A*" to the person who had property "*P*". In this case, the name "*A*" is an abbreviation for "the person who had the property *P*". It is obvious that, if this person also had the property *Q*, the statement "*A* had the property *Q*" is not analytic unless *Q* is analytically a consequence of *P*.

This is all very well as regards a historical character, but how about somebody whom I know intimately, e.g. myself? How about such a statement as "I am hot"? This may, following our earlier analysis, be translated into "heat is one of the qualities that make up I-now". Here "I-now" may be taken as denoting the same complex that is denoted by "my total present momentary experience". But the question remains: how do I know what is denoted by "I-now"? What is denoted is continually changing; on no two occasions can the denotation be the same. But clearly the words "I-now" have in some sense a constant meaning; they are fixed elements in the language. We cannot say that, in the ordinary sense, "I-now" is a name, like "Julius Caesar", because to know what it denotes we must know when and by whom it is used. Nor has it any definable conceptual content, for that, equally, would not vary with each occasion when the phrase is used. Exactly the same problems arise in regard to the word "this".

But although "I-now" and "this" are not names in quite the ordinary sense, I incline to think that there is a sense in which they must count as names. A proper name, as opposed to a concealed description, can be given to the whole or to any part of what the speaker is at the moment experiencing. When our verbal inventiveness fails, we fall back on "this" for the part of our total momentary experience to which we are specially attending, and upon "I-now" for the total momentary experience. I maintain that I can

perceive a complex of compresent qualities without necessarily perceiving all the constituent qualities. I can give the name "this" to such a complex, and then, by attention, observe that redness is one of its constituent qualities. The resulting knowledge I express in the words "this is red", which, accordingly, is a judgment of analysis, but not, in the logical sense, an analytic judgment. A complex can be perceived without my being aware of all its parts; when, by attention, I become aware that it has such-and-such a part, this is a judgment of perception which analyses the whole, but is not analytic, because the whole was defined as "this", not as a complex of known parts.

The kind of thing I have in mind is the kind of thing that is emphasized by the *Gestalt* psychologists. Suppose I possessed a clock which showed not only hours and minutes, but the day of the month, the month of the year, and the year of the Christian era, and suppose that this clock were to function throughout my life. It would then never twice during my life present the same appearance. I might perceive that two appearances of it were different, without being able to say at once in what the difference consisted. Attention might lead me to say: "In this appearance the minute-hand is at the top; in that, it is at the bottom." Here "this" and "that" are merely names, and therefore nothing said about them can be logically analytic.

There is another way of escaping from the conclusion that judgments are analytic when in fact they are obviously empirical. Consider again our clock that never repeats itself. We can define a date unambiguously by means of this clock. Suppose that, when the clock indicates 10 h. 47 m. on June 15, 1947, I say "I am hot." This can be translated into: "Hotness is compresent with the appearance of the clock that is described as 10 h. 47 m. June 15, 1947." This is certainly not analytic.

One way of making clear the scope and purport of our discussion is to put it in terms of "minimum vocabularies". By a "minimum vocabulary" for some body of propositions I mean a collection of words in terms of which all the other words that are needed can be defined, provided none of the words in the minimum collection can be defined in terms of the others in the collection. Thus e.g. Peano showed that, given the vocabulary of logic, everything in arithmetic could be defined by means of the three terms "0", "number" and "successor". So we may ask: "What is, in principle, a minimum vocabulary for describing the world of my sensible experience?" We have to ask ourselves: Can I be content with names of qualities, and words for compresence and for spatial and temporal relations, or do I need also proper names? And in the latter case, what sorts of things will need proper names?

I have suggested that ordinary proper names, such as "Socrates", "France", or "the sun", apply to continuous portions of space–time which

happen to interest us, and that space–time is composed of "complexes of compresence", which themselves are composed of qualities. According to this theory, an "instance" of (say) a shade of colour is a complex of which that shade is a constituent. The colour itself exists wherever (as we should commonly say) there is something that has that colour. Any collection of compresent qualities may be called a "complex of compresence", but it is only a "complete" complex when it cannot be enlarged without ceasing to be a complex of compresence. Often a complete complex can be rendered definite by mentioning only some of its components; e.g. in the above case of the clock, the complex is determinate when we are told what appearance of the clock belongs to it. This is what makes dating convenient.

Subject–predicate propositions expressing judgments of perception occur in two ways. First, if a complex is rendered determinate when only some of its constituent qualities are assigned, we may state that this complex also has such-and-such other qualities; this is illustrated by the statement "I was hot when the clock said 10 h. 47 m."

Second, I may perceive a complex without being aware of all its parts; in that case, I may, by attention, arrive at a judgment of perception of the form "P is part of $\overline{W}$", where "$\overline{W}$" is the proper name of the perceived complex. If such judgments are admitted as irreducible, we need proper names for complexes. But it would seem that the need for such judgments only arises through ignorance, and that, with better knowledge, our whole $\overline{W}$ can always be described by means of its constituents. I think, therefore, though with some hesitation, that there is no theoretical need for proper names as opposed to names of qualities and of relations. Whatever is dated and located is complex, and the notion of simple "particulars" is a mistake.

BIBLIOGRAPHY

Argyll, Duke of (1866) *The Reign of Law*, London: Strahan.

Armstrong, D. (1978) *A Theory of Universals*, Cambridge: Cambridge University Press.

—— (1983) *What is a Law of Nature?*, Cambridge: Cambridge University Press.

—— (1989) *A Combinatorial Theory of Possibility*, Cambridge: Cambridge University Press.

—— (1997) *A World of States of Affairs*, Cambridge: Cambridge University Press.

Austin, J. L. (1962) *Sense and Sensibilia*, Oxford: Oxford University Press.

Bradley, F. H. (1893) *Appearance and Reality*, Oxford: Oxford University Press (1968 edn).

—— (1914) *Essays on Truth and Reality*, Oxford: Oxford University Press.

Cartwright, N. (1983) *How The Laws of Physics Lie*, Oxford: Oxford University Press.

—— (1999) *The Dappled World*, Cambridge: Cambridge University Press.

Daly, C. (1997) "Tropes", in D. H. Mellor and A. Oliver (eds) *Properties*, Oxford: Oxford University Press.

Demos, R. (1917) "A discussion of a certain type of negative proposition", *Mind*, 26: 188–96.

Dennett, D. (1991) *Consciousness Explained*, London: Penguin.

Dretske, F. (1977) "Laws of nature", *Philosophy of Science*, 44: 248–68.

Dummett, M. (1978) *Truth and Other Enigmas*, London: Duckworth.

Ellis, B. (2001) *Scientific Essentialism*, Cambridge: Cambridge University Press.

Frege, G. (1892) "Über sinn und bedeutung", *Zeitschrift für Philosophie und Philosophische Kritik*, 100: 25–50; trans. Max Black, "On sense and meaning" in Frege 1980.

—— (1980) *Translations from the Philosophical Writings of Gottlob Frege*, P. Geach and M. Black (eds), 3rd edn, Oxford: Blackwell.

Gopnik, A., Meltzoff, A. and Kuhl, P. (1999) *How Babies Think: The Science of Childhood*, London: Weidenfeld & Nicolson.

Griffin, N. (1991) *Russell's Idealist Apprenticeship*, Oxford: Clarendon Press.

Honderich, T. (1988) *A Theory of Determinism: The Mind, Neuroscience, and Life-hopes*, Oxford: Oxford University Press.

Hume, D. (1777) *An Enquiry Concerning Human Understanding*, L. A. Selby-Bigge (ed.), 3rd edn revised by P. H. Nidditch, Oxford: Clarendon Press, 1975.

Hylton, P. (1990) *Russell, Idealism, and the Emergence of Analytic Philosophy*, Oxford: Clarendon Press.

Jackson, F. (1986) "What Mary didn't know", *Journal of Philosophy*, 83: 291–5.

Kindinger, R. (ed.) (1965) *Philosophenbriefe aus der wissenschaftliche Korrespondenz von Alexius Meinong*, Graz: Akademische Druck-u. Verlagsanstalt.

Lewis, D. (1973) *Counterfactuals*, Oxford: Blackwell.

—— (1986) "A subjectivist's guide to objective chance", in *Philosophical Papers* II, Oxford: Oxford University Press.

Locke, J. (1690) *An Essay Concerning Human Understanding*, P. Nidditch (ed.). Oxford: Clarendon Press (1975 edn).

Lombardo, E. (2002) "Analogical versus discrete theories of possibility", *Australasian Journal of Philosophy*, 80: 307–20.

Lowe, E. J. (1980) "Sortal terms and natural laws", *American Philosophical Quarterly*, 17: 253–60.

—— (1982) "Laws, dispositions and sortal logic", *American Philosophical Quarterly*, 19: 41–50.

MacColl, H. (1905) "Symbolic reasoning (VI)", *Mind*, 14: 74–81.

—— (1905a) "Existential import", *Mind*, 14: 295–6.

Martin, C. B. (1994) "Dispositions and conditionals", *The Philosophical Quarterly*, 44: 1–8.

Miah, S. (1987) "The emergence of Russell's logical construction of physical objects", *Russell*, N.S. 7: 11–24.

Molnar, G. (1999) "Truthmakers for negative truths", *Australasian Journal of Philosophy*, 77: 72–86.

Nagel, T. (1974) "What is it like to be a bat?", *Philosophical Review*, 83: 435–50.

Patterson, W. A. (1996) "The logical structure of Russell's negative facts", *Russell*, N.S. 16: 45–66.

Pears, D. (ed.) (1972) *Russell's Logical Atomism*, London: Fontana.

—— (ed.) (1985) *The Philosophy of Logical Atomism*, La Salle, Illinois: Open Court.

Popper, K. R. (1934) *The Logic of Scientific Discovery*, trans., London: Hutchinson (1959).

Ramsey, F. P. (1925) "Universals", in D. H. Mellor (ed.), *Philosophical Papers*, Cambridge: Cambridge University Press (1990).

Russell, B. A. W. (1896) *German Social Democracy*, London: Longmans, Green, and Co.

—— (1897) *An Essay on the Foundations of Geometry*, Cambridge: Cambridge University Press; 2nd edn 1996, London: Routledge.

—— (1903) *The Principles of Mathematics*, London: George Allen & Unwin; paperback edn 1992, London: Routledge.

—— (1904) "Meinong's theory of complexes and assumptions", in *Papers* 4.

—— (1905) "On denoting", in *Papers* 4.

—— (1912) *The Problems of Philosophy*, Oxford: Oxford University Press.

—— (1914) *Our Knowledge of the External World: as a Field for Scientific Method in Philosophy*, London: Open Court.

—— (1918) "The philosophy of logical atomism", in *Papers* 8.

—— (1927) *The Philosophy of Matter*, London: George Allen & Unwin.

Russell, B. A. W. (1927a) *An Outline of Philosophy*, London: George Allen & Unwin.
—— (1940) *An Inquiry into Meaning and Truth*, London: George Allen & Unwin.
—— (1946) "My own philosophy", in *Papers* 11.
—— (1948) *Human Knowledge: Its Scope and Limits*, London: George Allen & Unwin.
—— (1948a) "Postulates of scientific inference", in *Papers* 11.
—— (1953) "The cult of 'common usage' ", in *Papers* 11.
—— (1955) "My debt to German learning", in *Papers* 11.
—— (1957) "Mr Strawson on referring", in *Papers* 11.
—— (1959) *My Philosophical Development*, London: George Allen & Unwin.
—— (1967) *The Autobiography of Bertrand Russell*, I (1872–1914), London: George Allen & Unwin.
—— (1968) *The Autobiography of Bertrand Russell*, II (1914–44), London: George Allen & Unwin.
—— (1973) *Essays in Analysis*, D. Lackey (ed.), London: George Allen & Unwin.
Russell, B. A. W. and Whitehead, A. N. (1910–13) *Principia Mathematica*, Cambridge: Cambridge University Press, 3 vols.
Skyrms, B. (1981) "Tractarian nominalism", *Philosophical Studies*, 40: 199–206; reprinted in Armstrong 1989.
Talmont-Kaminski, K. (1998–9) "Pears' two dogmas of Russell's logical atomism", *Russell*, N.S. 18: 117–25.
Tooley, M. (1977) "The nature of laws", *Canadian Journal of Philosophy*, 7: 667–98.
Tully, R. E. (1988) "Russell's neutral monism", *Russell*, N.S. 8: 209–24.
—— (1993–4) "Three studies of Russell's neutral monism", *Russell*, N.S. 13: 5–35 and 14: 185–202.
Van Cleve, J. (1985) "Three versions of the bundle theory", *Philosophical Studies*, 47: 95–107.
Williamson, T. (1994) *Vagueness*, London: Routledge.
Wittgenstein, L. (1921) *Tractatus Logico-Philosophicus*, trans. D. Pears and B. F. McGuinness (1961), London: Routledge & Kegan Paul.

INDEX