Joshua Rust

John Searle and *The Construction of Social Reality*

CONTINUUM STUDIES IN AMERICAN PHILOSOPHY

JOHN SEARLE AND *THE CONSTRUCTION OF SOCIAL REALITY*

Continuum Studies in American Philosophy:
Series Editor: James Fieser, University of Tennessee at Martin

Dorothy G. Rogers, *America's First Women Philosophers*
Thom Brooks and Fabian Freyenhagen (eds), *The Legacy of John Rawls*
James A. Marcum, *Thomas Kuhn's Revolution*
Joshua Rust, *John Searle and* The Construction of Social Reality
Eve Gaudet, *Quine on Meaning*
David Roberts, *Virtue Ethics: Dewey and MacIntyre*
Douglas McDermid, *The Varieties of Pragmatism*
G. Elijah Dann, *After Rorty*
Timothy Mosteller, *Relativism in Contemporary American Philosophy*
Elizabeth F. Cooke, *Peirce's Pragmatic Theory of Inquiry*

JOHN SEARLE AND *THE CONSTRUCTION OF SOCIAL REALITY*

JOSHUA RUST

continuum
LONDON • NEW YORK

Continuum
The Tower Building, 11 York Road, London SE1 7NX
15 East 26th Street, New York, NY 10010

www.continuumbooks.com

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

ISBN: HB: 0-8264-8586-3

Typeset by Aarontype Limited, Easton, Bristol
Printed and bound in Great Britain by MPG Books Ltd, Bodmin, Cornwall

To my parents, Val and Diane Rust

Contents

Acknowledgements

This book is based on a dissertation written at the University of California, Riverside. I am indebted to my committee members whose contribution cannot be understated. Eric Schwitzgebel was something of an ideal chairperson in terms of his congeniality and enthusiasm for the project, his patience and diligence in reviewing multiple drafts, and his extensive understanding of John Searle's corpus. I explicitly and inexplicitly rely heavily on Larry Wright's work on explanation, and the orientation of this book owes much to his long-running Wittgenstein reading group, which is itself something of an institution at UC Riverside. Gary Watson's remarkable ability to cut to the heart of an issue was especially vital in transforming the dissertation, and the final chapter in particular, into its present form.

My family, and my parents in particular, were unerring in, not just enduring, but lovingly supporting what at times was a difficult and tedious intellectual process. Special thanks to Marijana Weiner, who was generous in providing much-needed editorial and emotional support. Finally, I am grateful for the explicit and tacit contributions of a number of colleagues and friends including Paul Hoffman, Chris Yeomans, Jessica Ludescher, Rachel Stevens, and Jenn Figg.

This page intentionally left blank

Introduction: "X Counts as Y in C"

In 1996 Toy Biz, the manufacturer of Marvel Comic's popular X-men action figures, sued US Customs Service in the Court of International Trade. Toy Biz successfully argued that the playthings should be classified as *toys* not *dolls*. According to Customs' classification, dolls purport to be human, toys do not. If the figures are not deemed to represent humans they would be subject to only a 6.8 per cent import duty instead of the higher 12 per cent for dolls.

On the one hand, the X-men seem human. The US government argued that the figures should be classified as humans, and thus dolls, because each character had a "distinctive individual personality". As for their super-human traits, the defense argued that, for example, Wolverine, who has a set of one-foot-long retractable claws on each hand, is simply "a man with prosthetic hands". However, it must be conceded that the ability to manipulate fire, shapeshift, or control weather systems at will, sharply distinguishes the X-men from ordinary human beings. In January of 2003, Judge Judith Barzilay declared, following the plaintiff's argument, that the X-men figures appeared to be "nonhuman creatures" due to "their extraordinary and unnatural . . . powers". The figures were thus found to merit the reclassification sought by Toy Biz.

One fan laments that the reclassification "is almost unthinkable. . . . Marvel's super heroes are supposed to be as human as you or I. They live in New York. They have families and go to work. And now they're no longer human?"[1] Indeed, since its inception in 1963 the comic book has tended to use the X-men, depicted as being almost universally feared and despised by those in the mainstream, to explicitly allegorize race relations. To those who

follow the comic book, the reclassification from doll to toy—from human to non-human—is not without irony.

The doll status of the X-men figures is a good example of what John Searle, in *The Construction of Social Reality* (*CSR*), calls an institutional fact. The rules that constitute institutional facts can be characterized according to the formula, "X counts as Y in context C," where X is a brute fact and Y is an institutional fact. In this manuscript I will refer to the "X counts as Y in C" formula as the "constitutive formula". Searle intends the formula to convey the sense in which an institutional fact Y is embodied or manifest in, but cannot be reduced to, a brute fact X. Using Searle's formula, playthings that purport to be human (X) count as dolls (Y) within the jurisdiction of US Customs (C), and those that do not purport to be human (X) count as toys (Y). It also underscores the sense in which institutional facts can be traced back to our collective acceptances. Moreover, institutional facts often implicate certain rights and obligations (they have a "status-function"), so that the reclassification of the X-men gives Toy Biz the right to pay the lower import duty.

Another example of an institutional fact is the wooden tally. Developed economies need a means to track debt. In medieval Europe one common means was the wooden tally. This consisted of a hazelwood stick on which was inscribed the date, the amount owed, as well as the debtor's name. The stick, along with this information, was split into two pieces, starting at about two inches from the bottom. The longer half—the "stock"—was retained by the creditor, whereas the shorter half—the "stub"—was kept by the debtor. If there was any question as to the size of the debt, the two halves could be put back together again. This helped guard against the possibility of fraud. When the debt was repaid the tally would then be destroyed. The stub (X) counts as an indication that I owe money to a creditor (Y) in medieval Europe (C). However, outside this context the stub (X) is not in itself an indication of debt-owed (Y).

Dolls, wooden tallies, or—Searle's archetypical example—money, cannot be reduced to the physical properties that underlie them: "a dollar" is not *just* the paper and ink out of which it is physically constituted. Nevertheless a dollar must be constructed of something, be it green paper and ink or metal. In claiming that *all* institutional facts—the US Customs' distinction between toys and dolls, indications of debt, money, language, marriage, football games—can be characterized according to the constitutive formula, Searle is claiming that an institutional fact Y is always founded on some brute fact X.

My intention is not to disagree with Searle on this point. It may be the case, as Searle contends, that for any institutional fact there is some constitutive, underlying brute fact to which I can point. Others dispute this and argue that some institutional facts do not seem to have a basis in some brute fact X.[2] My principal aim, however, is not to falsify Searle's account by way of counterexamples.

My concern runs somewhat deeper: *disagreement* presupposes that I am in the first place clear about what Searle is trying to convey with the constitutive formula. I am not clear.

Nor is Searle particularly helpful when it comes to the framing of his own insights. The constitutive formula is a crucial part of the answer to the questions Searle asks himself at the beginning of his book: "How are institutional facts possible? And what exactly is the structure of such facts?" (*CSR*, p. 2) But while Searle determines the structure of institutional facts to be "X counts as Y in C," what does he mean when he asks about how these facts are possible? Is he providing a foundational ontology of social reality, as Bertrand Russell's atomism attempted to identify the logical structure of brute reality? Or is he proffering a kind of mnemonic by which inquiry into institutional reality might proceed? Even though it is clear that Searle has said something interesting and important, there remain metaphilosophical questions about the significance of those claims.

Chapter 1—Searle's Institutional Atomisms

It is clear that the constitutive formula tells us something interesting about the nature of institutional reality. But there remains a question as to *how* it is interesting. Which puzzle does Searle intend to solve in asking the question, how are institutional facts possible? There may be an analogy between Searle's project and that of the atomists. Perhaps Searle's formula outlines the most general contours of institutional reality in somewhat the same way the atomists attempted to use logic to lay bare the structure of brute reality. This chapter fleshes out the comparison, noting points where the analogy breaks down. The almost stifling self-consciousness with which the atomists formulated the doctrine of philosophical analysis gives us a portrait of how we might understand the significance of the constitutive formula as an answer to Searle's own question.

Chapter 2—First Criticism of Institutional Atomism

The analogy between Searle and the atomists allows me to marshal part of an extensive body of criticism, originally directed against the atomists, against institutional analysis. I appeal to an argument originally advanced by John Wisdom and J.O. Urmson, who claim that there are principled reasons to think that it is impossible to complete the analysis of a given institution. I advance this argument by looking at difficulties that arise in attempting to characterize the institution of money.

Chapter 3—Second Criticism of Institutional Atomism

I argue that Searle, even by his own terms, has no basis by which to uphold the constitutive formula as the logical structure of institutional reality.

If these criticisms are convincing, we are again in the position of needing to ask what Searle hopes to have accomplished when he

asserts that "X counts as Y in C". How else might we understand the constitutive formula if not by means of an analogy with the atomists? Using groundwork established in Chapter 4, I take up this question in Chapters 5 and 6. Suggesting that Searle has advanced an ideal type, I will argue that he can avoid these objections.

Chapter 4—Kuhn, Weber, and Instruments of Inquiry

In Chapter 4 I set aside explicit discussion of Searle's view in order to present Max Weber's concept of the ideal type. I use Kuhn's notion of a paradigm as means of introducing the ideal type. This chapter begins with a sketch of Thomas Kuhn's view of inquiry in the physical sciences. I then chart some of the ways in which Max Weber's view of inquiry in the social sciences complements and anticipates Kuhn's depiction.

Both Weber and Kuhn characterize paradigms and ideal types as tools of inquiry, which give rise to puzzles and crises. I look at a number of responses, outlined by Kuhn and Weber, that the social and natural sciences have recourse to in the event of crisis.

Inquiry, I suggest, can proceed linearly, when there is a dominant paradigm or ideal type, or conjunctively, when there are multiple paradigms or ideal types in play. Regarding the latter possibility, Weber contends that there are no principled reasons why a researcher should not expect to employ several, incommensurable ideal types in order to understand a given phenomenon. Following Weber I suggest that reality is complex and so we can only expect so much from any one of our abstractions.

My exposition of Weber will help in my attempt to recharacterize the significance of Searle's constitutive formula in light of the atomist objections.

Why discuss Weber in the first place? Searle writes that since he takes himself to be addressing what "might be thought of as

problems in the foundations of the social sciences, one might suppose they would have been addressed and solved already in the various social sciences, and in particular by the great founders of the social sciences in the nineteenth century and the early parts of the twentieth century" (*CSR*, p. xii). Suggesting that the constitutive formula is an ideal type is interesting and provocative because it has the effect of locating Searle's examination of social reality *under* the umbrella of one of the founders of the social sciences, namely Weber.

The final chapters of the book reconnect my discussion of Weber to Searle's project. We can distinguish the constitutive formula itself ("X counts as Y in C") from the explication of a particular institution by means of the constitutive formula (green pieces of paper count as money). Chapter 5 argues that the latter are ideal types whereas Chapter 6 makes the more ambitious claim that the constitutive formula itself is an ideal type.

Chapter 5—Searle and the Ideal Type: Applications of the Constitutive Formula

In this chapter I argue that we should not expect the constitutive formula to help the researcher generate canonical articulations of our institutions. To make this claim I build off my Chapter 2 discussion of money. Searle holds that green pieces of paper (X) count as media of exchange (Y). A number of economists and sociologists have formulated alternatives to this neoclassical account of money: according to the chartalist account, green pieces of paper (X) count as an indication of debt-owed (Y). I argue that the chartalists and the neoclassicalists are not engaged in a factual dispute, but are rather advancing incommensurable ideal types. They are not making empirical claims but are rather advancing proposals for how a particular research program might proceed. If this is correct then *both* of these views can coexist.

Moreover, because both accounts of money can be expressed in terms of the "X counts as Y in C" formula, this suggests that the constitutive formula will not represent our institutions in an unambiguous, fully explicit way. This evokes Wisdom's objection, which I discuss in Chapter 2. Wisdom argues that a complete analysis of an institution is not in principle possible. Bringing Searle's remarks about money under the rubric of the ideal type sidesteps the force of Wisdom's objection. It does so, not by denying his insight, but by reevaluating the atomist's hyperbolic criteria for success. Because the ideal type brings us back to the actual conditions by which inquiry proceeds and succeeds, we need not be worried about the possibility of not being able to characterize a given institution exhaustively.

Chapter 6—Searle and the Ideal Type: the Constitutive Formula and the Status-function

In this final chapter I take aim at the constitutive formula itself, and not just particular applications of it. I argue that, just as the claim "green pieces of paper (X) count as a medium of exchange (Y)" is an ideal type, the formula "X counts as Y in C" is itself an ideal type. In this way, since ideal types highlight and suppress aspects of institutional reality, and the constitutive formula is an ideal type, we should expect that there are additional ideal types that uncover characteristics of institutional reality left unturned by Searle's formula. To this end, if the constitutive formula identifies a certain "normative component" indicative of institutional reality, I compare Searle's account of social reality with other models of normativity, including Aristotle's conception of the *phronimos*. I conclude, then, that Searle and the Aristotelians have articulated different ideal types, and so have formulated different instruments that attend inquiry.

Notes

1. Neil King, "Fans Howl in Protest as Judge Decides X-Men Aren't Human," *Wall Street Journal Online*, January 20, 2003.
2. Barry Smith and John Searle, "The Construction of Social Reality: An Exchange," in *John Searle's Ideas About Social Reality: Extensions, Criticisms, and Reconstructions*, eds David R. Koepsell and Laurence S. Moss (Malden, MA: Blackwell Publishing, 2003), p. 287.

1

Searle's Institutional Atomism

Overview of Searle's *Construction*

In *The Construction of Social Reality*, John Searle puzzles over the possibility and structure of a social world which, though real, also seems to be a product of intersubjective agreement. By "social reality" Searle has in mind institutions like money, property, marriage, government, football, and cocktail parties. In virtue of what, does some physical, brute event count as, say, a marriage or a game of football? The task that frames the *Construction* is, in Searle's words, "to assimilate social reality to our basic ontology of physics, chemistry, and biology" (*CSR*, p. 41). Searle imagines a "continuous line that goes from molecules and mountains to screwdrivers, levers, and beautiful sunsets, and then to legislatures, money, and nation-states" (*CSR*, p. 41). One end of this line tapers off in brute facts, perhaps the objects of the physical sciences, whereas the other end extends into the realm of institutional facts, such as money and nations, that exist only in virtue of human agreement.

Searle readily vacillates between a *negative* and a *positive* formulation of the notion of a brute fact. Starting with the positive formulation, Searle sometimes seems to commit himself to a tough "scientific metaphysics"[1] when he writes that "most of our metaphysics is derived from physics.... We live in a world made entirely of physical particles in fields of force" (*CSR*, pp. 6–7). However, when he goes on to characterize the aspects of these brute phenomena that are important for his articulation of institutional reality he emphasizes that "the features of the world I

described in characterizing our fundamental ontology, e.g., mountains and molecules, *exist independently of our representations of them*" (*CSR*, p. 9). Here, brute facts are defined negatively, as those facts that do *not* exist relative to the intentionality of observers. They are those things "out in the world," "intrinsic to nature" (*CSR*, p. 9), that would continue to exist without the presence of humans beings (*CSR*, p. 11).

Searle seems to underscore the negative formulation of brute facts. In the last three chapters of the *Construction*, Searle calls the thesis that there exists a world or reality that is independent of our representations, "external realism" (*CSR*, p. 150). He defends the thesis of external realism against those who might deny the existence of brute facts, who might deny that there are features of the universe that exist independently of us. Despite the claim that he founds institutional reality on an ontology informed by science, he only requires the milder, negative claim—the negative claim offers sufficient contrast to draw out the defining feature of institutional facts. To characterize brute facts strictly in terms of a metaphysics derived from science may, for Searle's purposes, be saying too much.[2] The minimalist, negative characterization of the notion of a brute fact enjoys the advantage of being profoundly uncontentious. Thus, it is an ideal base from which to build a theory of institutional facts.

The claim that institutional facts exist relative to our representations of them ultimately amounts to the claim that we impose a function on a brute fact that hitherto had no such function (*CSR*, p. 14). This imposition of a function onto a brute fact is expressed by the formula "X counts as Y in context C", which captures the basic form of what Searle calls a "constitutive rule". Searle sometimes paraphrases "X counts as Y in C" as "Y is imposed on X in context C". These two variants of the constitutive formula are equivalent in Searle's view. As noted in the introduction, I will call "X counts as Y in C" the "constitutive formula". "Y" designates the institutional fact, such as money. "X" refers to the underlying

brute fact on which the institutional fact is instantiated. The context (C) is most broadly conceived in terms of the overall system of agreement within which X is recognized as Y; however, C more often takes the form of perspicuous conditions by which X can count as Y (i.e., a promise only counts as a marriage vow in the presence of a judge or religious leader). The formula tells us that the institutional fact is logically dependent on brute facts (*CSR*, p. 56).

I illustrated institutional facts by appeal to Judge Barzilay's verdict on the status of the X-men action figures in the introduction. Only playthings that purport to be human count as dolls, not toys. For Searle, the archetypal cases of institutional facts in the *Construction* are games, money and marriage. In *Speech Acts* Searle writes:

> It is only given the institution of marriage that certain forms of behavior constitute Mr Smith's marrying Miss Jones. Similarly, it is only given the institution of baseball that certain movements by certain men constitute the Dodgers' beating the Giants 3 to 2 in eleven innings. And, at an even simpler level, it is only given the institution of money that I now have a five dollar bill in my hand. Take away the institution and all I have is a piece of paper with various gray and green markings.[3]

In these examples, the behavior of Mr Smith and Miss Jones, the movements of the baseball players, and the piece of paper with gray and green markings, are what the X term designates by Searle's formula. They are descriptions of characteristically institutional phenomena in terms of brute objects. The Y term designates the institutional fact: marriage, a baseball game, and money.

Institutional facts transpire through collective agreement or acceptance. Some X cannot be a medium of exchange simply in virtue of *my* deciding it is so. The collective intentionality which underlies institutional facts enables cooperative behavior but, in the case of institutional facts, is a condition for the norms and standards to which participants are subject.

Institutional facts are *iterated*. That which is designated by the Y term can serve as an X term for a higher-order institutional fact. Searle provides the following account of a marriage ceremony (*CSR*, pp. 82–3):

Marriage ceremony	Y
↑ (counts as)	↑
Entering into a contract	X/Y
↑	↑
Making a promise/commissive Speech Act (C)	X/Y
↑	↑
Uttering a Speech Act	X/Y
↑	↑
Such and such noises ("I do")	X

It is nevertheless the case that all institutional facts are ultimately constructed upon some brute fact, in this case the sound "I do".

With respect to the Y term, Searle sometimes speaks of a "status-function" rather than an institutional fact.[4] The status-function is thus "marked by the Y term in the formula 'X counts as Y'" (*CSR*, pp. 55, 63, 72, 112, 121). The status-function, or that which is denoted by the Y term, flags two crucial aspects: first, all institutional facts or status-functions are "agentive functions" that, second, involve or implicate rights and obligations.[5]

Beginning with the notion of an "agentive function": these are functions that are assigned in a way that is crucially dependent on our immediate purposes or activities. A rock is only a paperweight as part of a set of purposes that we have. Likewise, if certain green pieces of paper are a medium of exchange, that function only persists in virtue of the fact that we use them as such. Finally, what it is to be President of the United States is to have the agentive function of, among other things, signing or vetoing Congressional bills. We do not *discover* the functions which are constitutive

of something's being a paperweight, money, or the President in the same way that we discover that, for example, the function of a heart is to pump blood. The heart, then, has a "nonagentive function".

Paperweights, money, and the Presidency have (been assigned) agentive functions, but only money and the Presidency are institutional facts. What distinguishes a paperweight from money or the Presidency? That is, what distinguishes noninstitutional agentive functions from institutional agentive functions? Searle writes that money and the Presidency are different from paperweights in that the former are the product of the *collective* intentional imposition of agentive functions (*CSR*, p. 38). Moreover, because language is the product of collective intentionality,[6] Searle recommends that institutional agentive functions (money) have language as a constitutive element in a way that noninstitutional agentive functions (paperweight) do not (*CSR*, pp. 23, 37). As we will see in Chapter 3, these answers are not entirely satisfactory even by Searle's own lights. Moreover, even if it were the case that institutional facts were dependent on collective imposition and language, this account seems anyway to miss the important difference between the two types of case.

An important hint to a better answer to this question can be found late in the *Construction*, where Searle writes that the "content" or "primitive structure" of (ongoing) status-functions "involves" one of two "conventional power modes": enablements or requirements (*CSR*, p. 104), which Searle sometimes paraphrases in terms of rights and obligations (*CSR*, p. 100, 83, 103, 109). Status-functions first, are agentive functions and, second, have rights and obligations as their content. It is in the relationship between these two aspects of the status-function that we can most helpfully distinguish institutional agentive functions from noninstitutional agentive functions. The former either confer or directly implicate rights and obligations in a way that noninstitutional agentive functions do not. Whatever the paperweight does

for me, it does not give the bearer a new *right* to engage in certain behavior. This can be easily seen by looking at cases of breakdown.

Just as a heart might be *dysfunctional* if it does not pump blood, a feather is dysfunctional as a paperweight in a way that a rock normally would not be. If one of the functions of the President is to sign or veto Congressional bills, a President who skirts this obligation is not just dysfunctional, but is *remiss*. The crucial difference between institutional agentive functions and noninstitutional agentive functions is that the former implicate the possibility of remission. Looking at the institution of promise keeping, when someone fails to keep a promise,

> there is something wrong in a way that is different from the way there is something wrong with the man who stumbles when he walks; that is, there is a socially created normative component in the institutional structure, and this is accounted for only by the fact that the institutional structure is a structure of rules. ... It is precisely because of the rule that making a promise counts as undertaking an obligation that we recognize that certain kinds of behavior within the institution of promising are acceptable and certain other kinds are remiss (*CSR*, p. 146).

One's heart or one's walking may be arrhythmic, and so dysfunctional. But only institutional, agentive functions implicate rights and obligations, and with them the further possibility of remission. That institutional facts depend on the collective agreement of rules highlights the normative component indicative of institutional phenomena. Keeping in mind the way in which they are related, there are two aspects of institutional facts (Y) or status-functions: such facts, first, consist in agentive functions that, second, involve rights and obligations, and so the possibility of, not just dysfunction, but remission.

Institutional facts, for Searle, are agentive functions imposed by collective intentionality on an X that cannot perform that function

in virtue of its structure alone. Above I suggested that institutional facts can be contrasted with noninstitutional agentive functions in that the former implicate rights and obligations. Searle also distinguishes *institutional* facts from *social* facts. I likewise want to suggest that institutional facts can be contrasted with facts that involve collective intentionality—noninstitutional *social facts*—by appeal to the possibility of rights and obligations. To the extent that all social facts involve the imposition of agentive functions, the following will only reiterate the above claims while keeping an eye fixed on the particular case of social facts.

Any fact involving collective intentionality is a social fact (*CSR*, pp. 26, 38). Institutional facts are a subclass of social facts, so we might distinguish noninstitutional social facts from institutional social facts. With this in mind, I will contrast institutional facts (institutional social facts) from social facts (noninstitutional social facts). Searle argues that while both institutional and social facts may implicate the collective imposition of agentive functions, institutional facts are such that an imposed agentive function cannot solely be a result of X's physical features, but rather in virtue of continued human cooperation and recognition that X counts as Y. A wall which surrounds a city is a social fact, because it performs the function of keeping intruders out in a way that is dependent on the structure of X. As the wall crumbles over time it becomes an institutional fact if the line of rocks is recognized as a kind of boundary line by both inhabitants and noninhabitants alike. This function is in some sense symbolic, and does not rely on the physical structure of X alone.

Again, just as the possibility of remission distinguishes institutional agentive functions from noninstitutional agentive functions, the same mark can be used to distinguish social facts from institutional facts. The capacity to impose functions on such objects is a symptom of a more important shift indicated by the terms we would use to characterize functional breakdown. Searle argues that the "content" of the status-function is articulated in terms of

the distribution of rights and obligations on participants (*CSR*, p. 104). An outsider who steps over the stone boundary marker is *remiss* in a way ancestors who had once climbed the wall could not be; or else, if they were remiss, the wall was already performing a function that goes beyond its physical structure, but we cannot divine *this* from the structure of X alone. Consider the case where a country has marked its more remote borders with small stakes. These stakes, like the stones, have the *status* of a boundary marker. In order to further prevent illegal immigration, imagine that the country eventually builds a wall over the stakes. Migrants who then scale the wall are thereby *not* remiss. We cannot identify a status-function simply by looking at the X on which it is imposed.

Facts about the structure of X, while telling, do not seem to be the important difference between social and institutional facts. We should rather look to the normative content which underlies Y to characterize the difference between social and institutional facts. In the final chapter, I will return to this point in detail, presenting a positive account of the two varieties of normativity indicative of social and institutional facts.

Again, in highlighting certain parts of the *Construction* over others, I do not believe that Searle and I are in disagreement over this point. Indeed, in contrasting institutional facts from social facts, Searle sometimes explicitly appeals to the deontology of rights and obligations as the defining marker between such facts: "Animals running in a pack can have all the consciousness and collective intentionality they need. They can even have hierarchies and a dominant male; they can cooperate in the hunt, share their food, and even have pair bonding." These activities are social facts, but are not institutional facts. "Why not? Because [an animal] cannot represent to himself the relevant deontic phenomena" (*CSR*, p. 70). The possibility of rights and obligations, rather than the structure of X (sharing behavior, or even behavior such as exchanging a dollar bill for food), is what really matters in identifying institutional phenomena.

Framing the Construction

The question which concerns this book might be described as metaphilosophical. While it is clear that the constitutive formula highlights important features of social reality, Searle offers the reader little guidance in establishing its philosophical significance.

Searle does say, for example, that he is concerned with what "might be thought of as problems in the foundations of the social sciences" (*CSR*, p. xii), that he aims to excavate the "logical structure of institutional facts" (*CSR*, pp. 31, 56, 90, 94, 104–10, 112, 191).[7] Searle is interested in characterizing social reality in terms of a "Foundational Ontology" (*CSR*, pp. 3, 5–7, 13, 47),[8] and explicitly describes the investigation as "ontological, i.e., about how social facts exist" (*CSR*, p. 5): Searle attempts "to develop a general theory of the ontology of social facts and social institutions" (*CSR*, p. xii). In *Speech Acts*, which prefigures the *Construction*, Searle contends that "Every institutional fact is underlain by a rule of the form 'X counts as Y in C' ".[9] Searle says that he is conducting an analysis of our institutions (*CSR*, pp. 56, 79, 90, 100, 149), and that "the analysis of the structure of institutional facts reveal that they are logically dependent on brute facts" (*CSR*, p. 56).

These remarks are suggestive, but their significance is not transparent. Indeed, only after an extended clarificatory dialogue with Searle did Herbert Dreyfus begin to feel as though he thought he understood the nature of Searle's project: "I now understand that, when Searle analyzes the role of propositional representations in constituting actions and institutional facts, he is doing logical analysis ...".[10] Dreyfus' characterization suggests that Searle's project might find an analogue in that of the atomists and positivists. Ian Hacking explicitly says this much when he writes:

> Searle uses the word "construction" in a rather literal way. Chapter 1 is called "The Building Blocks of Social Reality".

> He has the bricks-and-mortar connotation of "construction" in mind. He writes in a philosophical tradition that includes Bertrand Russell ("logical construction") and Rudolf Carnap (Aufbau).[11]

Just as Carnap and Russell outline the terms by which we ought to analyze brute facts, Hacking suggests that Searle is outlining the terms by which we ought to analyze institutional facts. Perhaps, if Hacking is right, a comparison with the atomists or positivists will help us get clearer about the sort of endeavor in which Searle is engaged. "Construction" is, for Searle, Carnap, and Russell, logical or conceptual. The formula, according to Hacking, summarizes a much more detailed construction that includes collective intentionality:

> Searle assembles (his word) his building blocks by the end of Chapter 1. The ideas are three in number, namely: (1) functions are imposed on entities; (2) there is collective intentionality; (3) there is a distinction between constitutive and regulative rules. Constitutive rules (3) make possible some activity, while regulative rules tell how to conduct it, once the activity is recognized or engaged in. Constitutive rules are what matter for Searle, for he is concerned with what makes possible the existence of institutions, not how they are regulated.[12]

Searle's logical construction includes more than the claim that "X counts as Y in C". Following Hacking, we see that collective intentionality and the imposition of functions are part of the logical construction. Brute facts and collective intentionality are part of the ontology of institutional reality.

Prompted by Searle's own characterizations of the *Construction*, as well as Hacking's and Dreyfus' articulation of the text, I suggest that there is an analogy between philosophical analysis and the kind of project Searle understands himself to have undertaken. Just as the atomists are interested in articulating the constitutive

elements of brute reality, Searle is interested in describing the constitutive elements of social and institutional reality. The constitutive formula offers a skeletal outline of what such an institutional analysis would look like.

Comparing Searle with Russell and the atomists is more promising than comparing him to the positivists. The atomists were not apologetic about the fact that they were interested in metaphysics, whereas Carnap and the positivists were more exclusively concerned with the meaning of our words. Like Russell, Searle boldly characterizes his project as "ontological":

> Since our investigation is ontological, i.e., about how social facts exist, we need to figure out how social reality fits into our overall ontology, i.e., how the existence of social facts relates to other things that exist. We will have to make some substantive presuppositions about *how the world is in fact* in order that we can even pose the question we are trying to answer (*CSR*, pp. 5–6).

Searle is not primarily occupied with the meaning of the word "institution," "rule," etc. (or what they should mean); nor is he directly concerned with the meaning of words which refer to particular institutions, such as "marriage" or "money," except insofar as these concepts relate to the ontology itself. Russell thus appears to be a natural point of comparison, if only because the vocabulary, if not the substance, coheres more readily with that of Searle's. Later, I will review criticisms that will be equally applicable to an atomist *or* positivist spin on the *Construction*.

I will begin by articulating what atomism is, or better, how the atomists understood the significance of their own project. Following this discussion, we will be in a position to criticize both the atomists and, to the extent that a parallel can be drawn, Searle. The atomists and positivists were self-critical enough to provide the means for such criticism.

I do not believe that Searle's project must be understood by means of an analogy with the atomists. Nevertheless, I believe the following comparison may come close to articulating Searle's actual metaphilosophical views; it does the best job of clarifying a number of puzzling statements that can be found in the *Construction*. However, given the near absence of metaphilosophical reflection in the text, it is difficult to know for sure. In any case, the atomist reading of Searle's text provides a useful foil against which to advance an alternative reading, to be proposed later.

Logical atomism and analysis

J.O. Urmson, in *Philosophical Analysis*, explores the relationship between logical atomism and the program of analysis in which the atomists were engaged. Analysis is a technique by which we redescribe unwanted, misleading terms in less contentious ways. Russell's "On Denoting" offers the pivotal example, where terms that seem to refer to nonexistent entities (the King of France) are cast off as descriptions.[13] "The King of France is bald" is analyzed in the following way: "There is one and only one thing which is the King of France and whatever is the King of France is bald." The task is to see if we can understand how this sentence can make sense without postulating a world filled with imaginary beasts, places, people, etc.; having done this, the worst that can happen is we have said something false. The new formulation, then, has a structure more appropriate to the form of fact.

Logical atomism is Russell's metaphysical doctrine. Russell's ontology is a familiar and relatively simple one: reality consists of particulars and their characteristics and relations. An atomic fact consists of a particular that possesses a characteristic or stands in some relation to other particulars. To describe this reality we need a vocabulary of proper names which stand for particulars, and a set of predicates to describe the qualities or relationships of those

particulars. Thus, any statement about brute facts should take the form "$R(n)$," where n is a proper name that designates some component or particular, and R designates some characteristic or relation. If R is a relation, at least two particulars must be designated.

One concern is that even if one accepts the program of analysis offered in "On Denoting," it is not clear that one is driven to the metaphysical commitments of logical atomism. There is a gap between analysis and logical atomism that can be seen most clearly when we start considering what kinds of expressions are taken to be misleading. In "On Denoting" Russell reflects on sentences about unicorns and the King of France. However, in *The Philosophy of Logical Atomism*, "desk,"[14] rather than "the King of France," becomes the object of investigation. It is striking that many who might worry that "the King of France" might move us to postulate nonexistent entities would *not* have the same worries about desks; the urgency to analyze desks is not there in the same way that it might be for unicorns. Indeed sentences like "the desk is old" is precisely the sort of normal, unproblematic case against which "the King of France is old" might appear puzzling in the first place.

Urmson's two types of analysis: problem and resolution

Even if expressions like "the King of France" should be clarified, it is by no means clear that "desk" stands in need of the same treatment. Urmson uses the difference between the two cases as a basis for distinguishing between two types of analyses. On one hand, the paraphrasing of "the King of France" into a description is an example of "same-level analysis" which "involves only logical and not metaphysical progress".[15] Elsewhere Urmson writes that same-level analysis "was intended to bring out the logical form of the proposition, concealed as it was behind a misleading grammatical form" and "it solved logical puzzles rather than

giving metaphysical insight".[16] On the other hand, the paraphrasing of "desk" into a system of simples and their relations, attempts to make metaphysical progress by "replacing logical constructions by basic realities".[17] Logical atomism depends on a kind of analysis that provides, not only linguistic guidance, but also metaphysical insight and progress. Urmson calls this, "new-level analysis". It does this, like same-level analysis, by correcting misleading syntax.

Urmson's distinction between different types of analysis is helpful, but also misleading. Urmson recommends that same-level analysis does not tell us anything about the world. This cannot be right; the whole point of paraphrasing "the King of France" was to ensure that some entities are not counted among the real. The improvement of the form of the statement is intimately tied to convictions about the way the world is. I am using the word "metaphysics" to flag the fact that analysis takes for granted the world is a certain way; "the present King of France" does not refer to any entity. Like new-level analysis, same-level analysis also gets the metaphysics right, if only by ensuring that no occult entities are admitted. If this is right, new-level and same-level analysis are more similar than Urmson's discussion would suggest.

However, "more similar" is not the same as "identical," and Urmson has drawn our attention to a useful distinction, even if it tends to eclipse important similarities. New-level analysis does make more adventurous metaphysical claims than same-level analysis. Given that nouns typically do refer to something, same-level analysis ensures that we are not stirred to think that "the King of France," "the unicorn," "the round square" must also refer to something—that there must in some sense be a King of France. Russell successfully bars, through same-level analysis, the evocation of certain entities. To show how the King of France does not refer to anything is not a metaphysically neutral insight, but is significantly less contentious than the positive claim that the universe fundamentally consists of an indefinite number of small,

fleeting particulars in the form of sense data. Many would agree with the former claim, but be surprised by the latter; thus, there remains a difference between same-level analysis and new-level analysis, even if it is not the case that same-level analysis is metaphysically indifferent. My deflationary reading of the difference between same-level and new-level analysis returns the gulf to a gap, and gives us reason to recharacterize the difference between logical atomism and analysis.

Distinguishing metaphysics and methodology

Having become clearer on the difference between same- and new-level analysis, we can return to the task of characterizing the gap between logical atomism and analysis. If Urmson mischaracterizes the difference between same-level and new-level analysis, then in both cases we may distinguish between a metaphysical component (i.e., logical atomism) and a methodological component (i.e., analysis). The metaphysical or ontological component establishes the terms of inquiry, what is kept fixed or held fast. The methodological component is concerned with articulating what the world in fact looks like according to those terms. I would like to proceed as if this distinction were unproblematic.[18]

The distinction between metaphysics and methodology is helpful in that it implies a natural division of labor—there are those philosophers who analyze the world and those who justify or provide the ontological underpinnings of those analyses. Indeed, especially among the atomists, the distinction between metaphysics and methodology is a widely accepted one and yet, in practice there is a certain amount of slippage between the two tasks.

For example, in Russell's articles in *The Monist*, and especially "Excurses into Metaphysics: What There Is," the title alone is enough to suggest that Russell is primarily concerned with metaphysics and not methodology. However, the document is rife with methodological suggestions about how one might analyze desks,

chairs, and even phantoms or hallucinations in terms of sense-data. But these concrete suggestions about how one might go about analyzing the world are ultimately extra-metaphysical. According to Russell, metaphysics consists of particulars and their relations; if we want an account of the world, we must analyze it in these terms. It is by virtue of a different and distinguishable investigation that we might find out what those particulars and relations, in fact, are. Thus, in this text, Russell's attempt at analysis seems to serve largely illustrative or pedagogical purposes.

"On Denoting," on the other hand, is primarily but not exclusively concerned with methodology. Given a barebones metaphysic and a system of analysis, how can we analyze "the King of France" in particular, and descriptions in general? The metaphysics (i.e., that "the King of France" does not refer to anything) is presupposed by Russell's largely methodological investigation. Attempts to analyze, say, nations or individuals are methodological pursuits.

Analysis—the methodological endeavor—attempts to map the particular features of our world whereas the metaphysical endeavor attempts to articulate and justify the logical limitations of such a map. Methodology is concerned with contingent features, whereas the metaphysical is concerned with necessary ones—what is held fast. As we engage in analysis, we learn something about how our world is actually constituted; subsequent applications of analysis allow us to come to terms with our world as it is given to us. We come to find out, for example, that the world consists, not just of particulars standing in certain relations to each other, but of sense-data, colors and even desks, people, and nations. A suitable metaphysics, then, enables, or rather authenticates, forms of empirical investigation.

The distinction between metaphysics and methodology dovetails with what Locke calls the "under-labourer" conception of philosophy.[19] In this view, methodological tasks could easily be outsourced to other disciplines. The aim of metaphysics is not so

much to make epistemological progress, but to clear potential impediments to such progress. There is, nevertheless, an important point of intersection between the methodology and the metaphysics. Methodological investigations that cannot be made to work are an important source of objections to the metaphysical investigation. Unanalyzable objects might, for example, serve as counterexamples.

Methodological and metaphysical tasks: Searle and the atomists

I would like now to return to the *Construction*. Just as we can distinguish the atomists' methodological task from their metaphysical task, we can distinguish a methodological task and a metaphysical task in the *Construction*.

Methodology: institutional analysis

The most obvious case of methodology is a social scientific research program implied by the constitutive formula. If institutions can be articulated in terms of iterations of the constitutive formula, we might call this methodological program "institutional analysis". Institutional analysis is methodological, in that it outlines a procedure the social scientist follows to depict the structure of institutional reality; the aim is to discover the formal structure of the institution of, say, money, or marriage.

Institutional analysis is framed by a philosophical, or ontological project. According to the atomist analogy, the success of the philosophical task provides a mandate for the institutional analyst. The social scientist should attempt to codify our institutions according to the constitutive formula, because the formula has been shown to get the ontology right.

Like the ontological account, institutional analysis is analogous to the empirical research program championed by Russell. Just as Russell attempts to describe a desk in terms of a system of correlated particulars, the institutional analyst provides the following (simplified) description of the institution of a marriage ceremony: such and such noises count as entering into a marriage in virtue of certain collective acceptances (*CSR*, p. 83).

The methodological program implied by the constitutive formula need not be seen as competitive with accounts of money or marriage that do not explicitly cite the formula. Just as Russell takes himself to be clarifying what we already mean when we refer to the King of France or desks, Searle likewise is articulating that which must be presupposed by any social scientific investigation. Like Russell, Searle might agree that accounts of institutional reality that explicitly employ the constitutive formula avoid potential confusions; such researchers are not likely to confuse money with the brute stuff on which it is instantiated. For example, the move to a "cashless society" is less significant than writers of op-ed articles in the popular and financial press would have us believe.[20] We have only exchanged one brute fact (bits of paper) for another (data bits); otherwise, money itself is not importantly different.

Neither Russell nor Searle seriously attempts to analyze, respectively, a desk or the institution of money; this is presumably the task of other philosophers, scientists, or social scientists. Russell and Searle might optimistically be described as project managers. In bracketing methodological concerns, Searle appears to fall back on an under-labourer conception of philosophy. Thomas Osborne describes the *Construction* as "a terrific philosophical work" but contends that "it is quite literally *indifferent* as sociology".[21] He presses Searle to articulate the "ways he himself might be *expecting* his book to affect the day-to-day practices of the social sciences".[22] Searle implicitly underscores the atomist divide between metaphysics and methodology in his response: "The

truth is, I had no expectations whatever. I found the problems so hard and working them out so consuming, that I never gave any thought to what consequences my book might have for the actual practice of working social scientists".[23]

According to the under-labourer conception, the social scientist does not need the philosopher any more than the scientist needs *The Philosophy of Logical Atomism* in order to be able to analyze things in terms of their component parts. Indeed, on this point, both Searle and Osborne appear to agree that the *Construction*'s project (metaphysics) is distinguishable from that of the social scientist's (methodology). For Russell's and Searle's purposes, it is sufficient merely to gesture at what empirical or institutional descriptions might look like.

Metaphysics: Searle's building blocks

Searle's particulars are Hacking's building blocks: the world contains brute phenomena, including collective intentionality. Searle exploits the causal connection between intentionality and the status-function Y, in order to analyze a constitutive rule in terms of two types of brute phenomena: brute fact X and intentionality. Y is entailed by the analysis because of intentionality's special causal properties:

> Even natural phenomena, such as rivers and trees, can be assigned functions, and thus assessed as good or bad, depending on what functions we choose to assign to them and how well they serve those functions. This is the feature of intentionality I am calling "the assignment—or imposition—of function" (*CSR*, p. 14).

Because institutional facts are assigned functions, and intentionality has the special feature of being able to impose such functions, Searle can tell us of what institutions are constituted—namely,

brute facts and collective intentionality. This answers the driving question of the *Construction*: how can institutions exist in a world that consists entirely of brute facts (including intentionality)?

Is Searle a reductionist about institutional facts, as the analogy with the atomist would suggest? In the *Construction* Searle comes across as ambivalent about reductionism, as evidenced by Hacking's attempt to articulate his view:

> A comparison with and contrast to Carnap is useful. Carnap is often called a "physicalist". He wanted to "reduce" propositions about experience and much else to propositions about the physical world. Whatever "reduce" might mean in logic or some philosophical pastiche of natural science, Searle most strongly does not want to reduce the social to the physical.... He does think that an honest philosopher ought to show how certain facts describable in terms of the basic ontology are both necessary and sufficient for the existence of social reality.[24]

Searle clearly states that "satisfying the X term, 'such and such bits of paper,' is not by itself sufficient for being money" (*CSR*, p. 44). More forcefully, Searle affirms that "We are not trying to reduce the concept 'money' to noninstitutional concepts" (*CSR*, p. 53) and status-functions or "deontic phenomena are not reducible to something more primitive and simple" (*CSR*, p. 70). And yet, he writes that "one of the aims of this book is to show ... how the world of institutions is part of the 'physical' world" (*CSR*, p. 120). Indeed, sometimes Searle comes across as a brazen reductionist: "In our toughest metaphysical moods we want to ask 'But is an X really a Y?' For example, are these bits of paper really *money*? ... Surely when you get down to brass tacks, these are not real facts" (*CSR*, p. 45). In an interview, Searle says of the *Construction*:

> I start with what we know about the world: the world consists of entities described by physics and chemistry. I start with the fact that we're products of evolutionary biology, we're biological

> beasts. Then I ask, how is it possible in a world consisting entirely of brute facts, of physical particles and fields of force, how is it possible to have consciousness, intentionality, money, property, marriage, and so on?[25]

There is a sense in which Searle is both a reductionist and an anti-reductionist. Searle needs the appeal to the collective imposition of constitutive rules on brute phenomena in order to provide logical conditions for, say, money in particular and institutions in general: such and such bits of paper are not sufficient to be money, but such and such bits of paper, which are seen to function as a medium for exchange, *is* sufficient. Searle is a nonreductionist because he does not think that the Y term can be reduced to the X term alone (money is not *just* green bits of paper). But he is a reductionist in that he appears to think that the status-function denoted by the Y term is just the collective intentional imposition on a given brute fact. In this way, institutional facts are characterized in terms of two kinds of brute facts: whatever is designated by the X term and the collective intentionality which causes or creates a status-function. Searle can exploit the causal connection between function and intentionality, so the status-function Y is further reduced to intentionality.[26] Institutional reality can be understood in terms of a world that consists entirely of brute facts and conscious biological beasts.

Just as the atomists analyze brute objects into particulars and their relations, Searle analyzes institutional facts into brute facts and intentional impositions. These are Searle's building blocks. Obviously, Searle's building blocks or atomic facts are not Russell's. From Russell's point of view, Searle's building blocks are the results of analysis, not the basis of it. Searle may or may not share Russell's view: and while Searle almost certainly would not see those physical particles further reduced to sense-data, the tone of the discussion indicates a deep sympathy with the broad outlines of atomist outlook.

Whether Searle's atomic facts consist of higher-order phenomena such as people, intentionality, collective intentionality, actions, and speech acts, or lower-order phenomena such as atoms and fields of force, or even sense-data, it seems clear that a comparison with the logical atomists not only preserves, but also highlights and clarifies many of the remarks found in the *Construction*. This alone is enough to recommend the atomist lens as a tool to help us uncover the significance of the X counts as Y in C formula.

I would like to conclude this section by noting that Margaret Gilbert identifies a way of thinking about social phenomena that finds reflection in Searle's own writing. Gilbert's *On Social Facts* aims at articulating what she takes to be a key notion of any investigation into the social: "plural subjecthood". This term, which we will have occasion to review in Chapter 6, implicates a version of collective- or we-intentionality[27] that underlies Searle's own account of institutional phenomena (of which the plural subject is a variety). Gilbert goes on to explicitly and helpfully articulate the outlines of the metaphilosophical framework I am attributing to Searle in this chapter.

Gilbert emphasizes that if philosophy is to concern itself with questions of the human sciences (i.e., the degree to which the human sciences can be scientific), it must *first* inquire into the nature of social phenomena: "Presumably in a logical ordering ontology precedes methodology. Weber and Durkheim saw things this way, and acted accordingly."[28] Gilbert's observation is striking as it gives a kind of urgency to the task of articulating which "social phenomena" are presupposed by any inquiry into the social facts. Her own comments recall Searle's claim that he is concerned with questions about "the foundations of the social sciences" (*CSR*, p. xii).

Using the concept of a plural subject, she discovers that Weber's and Durkheim's basic ontologies lack "a special aptness for the label 'social,' " unlike her notion of plural subjecthood. She concludes that, "Indeed, [the notion of the plural subject] carves

nature at her joints".[29] She thus is able to offer an "effective, unitary account of those sciences aptly referred to as the social sciences".[30]

Gilbert's picture, and particularly her split between ontology (metaphysics) and methodology, nicely tracks the analogy with the logical atomist, who holds that we need to determine the terms of inquiry (simples and their logical relations) before we can be clear about the status of objects like tables and chairs. It is this picture that I am attributing to Searle under the atomist framing of his "X counts as Y in C" formula.

A difficulty with the atomist reading of the constitutive formula

I have suggested that the *Construction*'s attempt to outline the logical structure of institutional reality in terms of the constitutive formula is importantly analogous to the atomist's attempt to outline brute reality in terms of set theoretic constructions. Searle gives us two necessary conditions for something's being an institutional fact: there must be some (1) brute fact, X, on which some (2) agentive function is imposed by collective intentionality.

There is a difficulty with this reading, which may require that we further soften the analogy between Searle and the atomist. To see this we have to look, not at the *Construction* itself (1995), but to Searle's reply to Barry Smith's critique of the "X counts as Y in C" formula (2003). Smith believes that he has discovered counterexamples to Searle's formula, and the ontology which it entails; there are Y's which do not appear to have a corresponding X—a promise (Y) remains standing even after the words are uttered (X). There are, contrary to the constitutive formula, "free standing Y terms". In that way, it does not appear to be the case that "all institutional forms of human culture . . . must always have the structure X counts as Y in C . . ." (*CSR*, p. 40).

Searle's response to Smith's counterexample is surprising: "I think [Smith] thinks that I am trying to answer his various questions about necessity. I am not. He compares this problem to the old positivist attempts to reduce analytic propositions to logical truths."[31] Here Searle explicitly distances his project from that of the atomists. It is still the case that both Russell and Searle take themselves to be doing ontology, but Searle, unlike Russell, does not take himself to be formulating a set of conditions presupposed by any account of institutional reality: thus X and intentional imposition are not necessarily necessary components of institutional facts as previously suggested.

Searle goes on to recommend that the status-function, or Y term, is the important part of the formula. Searle responds by accusing Smith of misunderstanding both the nature and the focus of the project: Searle is not interested in articulating conditions, and is anyway more interested in the Y term than the formula in its entirety.

This does dodge the force of Smith's counterexamples, but raises more questions than it answers. First, if the nature of Searle's project was unclear before, it is all the more so. The *Construction* purports to show how institutional reality is compatible in a world that consists entirely of brute facts, including intentionality. As we have seen, not just Smith, but Dreyfus and Hacking take Searle to be engaged in something like logical construction, in the tradition of Russell and Carnap. If Searle is outlining the ontology of socially created reality (*CSR*, p. 13), the nature of his project remains deeply unclear if he does not take himself to be providing conditions in terms of lower-order phenomena. This project becomes especially baffling against repeated claims that he has delineated institutional reality's *logical* structure, demonstrating the "Logical Priority of Brute Facts over Institutional Facts" (*CSR*, p. 34). "The analysis of the structure of institutional facts," Searle writes, "reveal that they are logically dependent on brute facts" (*CSR*, p. 56).

Second, the shift in emphasis from the constitutive formula to the status-function seems unnecessary, even by Searle's own lights. Searle's follow-up remarks to Smith, such as "one could state the thesis of the whole book without" appeal to the constitutive formula, radically and bizarrely underplay the force of his own insight. A much less rash response is readily available to Searle—one that not only preserves, but strengthens, the integrity and centrality of the constitutive formula.

Despite undue focus on money-type examples, when it comes to the imposition of a Y on some brute fact X, Searle is principally interested in *agents*, as opposed to things like green bits of paper. After pointing out that status-functions can be imposed on people, objects, and events, he goes on to emphasize that "the category of people, including groups, is fundamental in the sense that the imposition of status-functions on objects and events works only in relation to people" (*CSR*, p. 97). Elsewhere, he writes that "Social objects are always constituted by social acts; and in a sense, *the object is just the continuous possibility of the activity*. A twenty dollar bill, for example, is a standing possibility of paying for something" (*CSR*, p. 36), or "What we think of as social *objects*, such as governments, money, and universities, are in fact placeholders for patterns of *activities*" (*CSR*, p. 57).

So long as we are talking about people and their actions, which for Searle are anyway "fundamental," that which is specified by the Y term is always *embodied*, in strict accordance with the "X counts as Y in C" formula. Here, the rights and obligations specified by the Y terms obviously cannot be reduced to the underlying organism, but are nevertheless manifest in the behavior of this brute, biological X. Institutional facts such as money or promises, including facts denoted by free standing Y terms, are derivative of the fact that the status-function must be imposed on persons. Recourse to this modest reformulation sustains our interest in vast swaths of the *Construction*, as the constitutive formula is not cast off as *merely* pedagogical. Searle is too easily startled by

Smith's counterexample. If the constitutive formula is a ladder, there are many more rungs yet to climb before we should consider kicking it away.

If Searle's response to Smith is reconcilable with the *Construction*, the analogy with the atomists must be further thinned. Perhaps, unlike the atomist, Searle is not interested in articulating necessary conditions for the possibility of institutional reality. While it seems possible that Searle may have misrepresented his own views in order to avoid Smith's counterexample it is clear that he is, like the atomist, interested in depicting the (logical?) structure or ontology of institutional reality by some other means.

For my purposes, even this minimal comparison with the atomists will be sufficient to drive the objections of the next two chapters. Except in the final chapter, where I attend to just the Y term (the status-function), I will continue to address Searle's "X counts as Y in C" formula. Because that discussion does not leverage the possibility of free standing Y terms, difficulties with the constitutive formula are easily reconfigured to apply to the status-function.

Notes

1. Newton Garver, "Philosophy as Grammar," in *The Cambridge Companion to Wittgenstein*, eds Hans Sluga and David G. Stern (Cambridge: Cambridge University Press, 1996), p. 158.
2. It seems clear that Searle does not need to embrace the positive articulation of brute fact in order to get his account of social reality off the ground, but it still may be the case that he endorses the stronger account anyway.
3. John R. Searle, *Speech Acts: An Essay in the Philosophy of Language* (Cambridge: Cambridge University Press, 1969), p. 51.
4. Like institutional facts, status-functions are iterated so that the rights and obligations of, e.g., a bail enforcement officer build upon the rights and obligations conferred to certain private citizens of the United States of America to make a citizen's arrest. Here there are at least three nested levels of status-function, wherein more specific enablements and

requirements are conferred to a member of the lower level: citizens of the United States of America—those citizens empowered to make a citizen's arrest—bail enforcement officers.

5. One exception is honorific status-functions, which do not confer any rights or obligations on the bearer.
6. Language, like football or marriage, is itself an institutional fact. Just as, through a process of collective agreement, running into the end zone under certain conditions counts as a touchdown, the utterance of certain sounds under certain conditions constitutes a language-game of one sort or another: giving orders, describing something, reporting, speculating, making a joke, etc.
7. John R. Searle, *Mind, Language, and Society: Philosophy in the Real World*, 1st ed. (New York, NY: Basic Books, 1998), pp. 123–4.
8. See also John R. Searle, *Consciousness and Language* (New York: Cambridge University Press, 2002), p. 138.
9. Searle, *Speech Acts: An Essay in the Philosophy of Language*, pp. 51–2.
10. Hubert L. Dreyfus, "The Primacy of Phenomenology over Logical Analysis," *Philosophical Topics* 27, no. 2 (1999): p. 2.
11. Ian Hacking, "Searle, Reality and the Social," *History of the Human Sciences* 10, no. 4 (1997): p. 85.
12. Ibid., p. 88.
13. Bertrand Russell, "On Denoting," in *Classics of Analytic Philosophy*, ed. Robert R. Ammerman (Indianapolis: Hackett, 1990).
14. Bertrand Russell, *The Philosophy of Logical Atomism*, ed. David Francis Pears (LaSalle, Ill.: Open Court, 1985), p. 144.
15. J. O. Urmson, *Philosophical Analysis; Its Development between the Two World Wars* (Oxford: Clarendon Press, 1956), p. 40.
16. Ibid., p. 39.
17. Ibid., p. 41.
18. I use the distinction between metaphysics and methodology because the atomists did. There are less caustic ways to characterize what I take to be correct about the distinction: for example, Bas van Fraassen in *The Empirical Stance* distinguishes "epistemic stances" (metaphysics), which are strategies or policies for generating factual beliefs, from "facts" (the product of methodology), which are claims about the nature of the world. Unfortunately the characterization of "metaphysics" tends to harden or entrench an otherwise promising epistemic stance. Indeed, van Fraassen's rechacterization of the empiricist tradition anticipates the recharacterization of Searle's institutional atomism that I will undertake in

Chapters 4, 5, and 6. See Bas C. van Fraassen, *The Empirical Stance* (New Haven, CT: Yale University Press, 2002).

19. Peter Winch, *The Idea of a Social Science and Its Relation to Philosophy*, 2nd ed. (London: Routledge, 1990), pp. 3–5.
20. John N. Smithin, "Introduction," in *What Is Money?* ed. John N. Smithin (London; New York: Routledge, 2000), p. 1.
21. Thomas Osborne, "The Limits of Ontology," *History of the Human Sciences* 10, no. 4 (1997): p. 98.
22. Ibid., p. 102.
23. John R. Searle, "Reply to Thomas Osborne," *History of the Human Sciences* 10, no. 4 (1997): p. 109.
24. Hacking, "Searle, Reality and the Social," p. 89.
25. Faigenbaum in Barry Smith, *John Searle* (Cambridge; New York: Cambridge University Press, 2003), pp. 16–17.
26. I should note here that Searle writes that explicit intentionality can be a proxy for less objectionable phenomena; Searle is free to evoke the "Background" instead of collective intentionality in order to explain how status-functions get created. I will return to this important point in Chapter 3 (*CSR*, pp. 5, 13, 126, 129, 142).
27. However, Gilbert, unlike Searle, thinks that we-intentionality can be analyzed into configurations of I-intentions.
28. Margaret Gilbert, *On Social Facts* (Princeton, NJ: Princeton University Press, 1992), p. 440.
29. Ibid., p. 442.
30. Ibid., p. 22.
31. John Searle and Barry Smith, "The Construction of Social Reality: An Exchange," in *John Searle's Ideas About Social Reality: Extensions, Criticisms, and Reconstructions*, eds David R. Koepsell and Laurence S. Moss (Malden, MA: Blackwell Publishing, 2003), p. 300.

2

First Criticism of Searle's Institutional Atomism—Methodology

To the extent that the *Construction* is understood by way of an analogy with atomism, criticisms directed against the atomist can be reformulated against the constitutive formula. The first criticism, the subject of this chapter, concerns the methodological program, and questions the possibility of using the constitutive formula to depict institutional reality. The second and perhaps more important criticism concerns the possibility of Searle's metaphysical program; this will be the subject of the next chapter.

Even in rejecting the atomist's propensity towards metaphysics, the positivists held onto the promise that logical analysis outlines a viable methodology or research program. Perhaps the constitutive formula gives the researcher the means to represent the deep structure of a given institution independent of accepting Searle's metaphysical picture.

An objection originally formulated by John Wisdom and later taken up by J.O. Urmson in *Philosophical Analysis* attacks both the atomists and the positivists for formulating the broad outlines of a research program that cannot be undertaken. Roughly, these authors give principled reasons for thinking that philosophical analysis is impossible because any given analysis cannot be completed. I will argue that institutional analysis proves equally intractable.

Wisdom and the impossibility of philosophical analysis

John Wisdom points out that if "England declared war" (in 1939) cannot mean "All Englishmen declared war," so that "England"

is analyzed in terms of "all Englishmen," then it might be analyzed as follows: "A great many people in England are hurrying to make shells and tanks and guns and gas, and a message has been sent by the man who manages these things for Englishmen to the man who does the same for the people they are going to fight, to say that now they are going to fight and so on."[1]

This is presumably a logically adequate definition, but it is only so in virtue of the qualification, "and so on". Wisdom compares the employment of the ellipsis to the following analysis of a chair: "A chair is a set of sensations *chairishly* related".[2] These quasi-analyses, by the analyst's own lights, evade the very question asked. The analyst's response to both of these questions is to have "chairishly" defined, and "and so on" enumerated.

There are times, of course, when it is appropriate to et ceteraize the analysis. Analysis—the methodological endeavor—attempts to map the particular features of our world, whereas the metaphysical endeavor attempts to articulate and justify the logical limitations of such a map. If the project is not philosophical analysis, per se, but rather, metaphysics, then these gestures are adequate. For the purposes of specifying metaphysical conditions, we just need a rough picture of what the analysis of "England declared war" or "chair" might look like, to make sure that the metaphysics and methodology are in accord with one another (consider, again, Russell's analysis of a desk).[3]

However, when we are engaging in methodology per se these rough gestures no longer serve this purpose. Perhaps, however, they serve another purpose. Elliptical expressions instructively tell us the sort of things and relations that ultimately constitute declaring war or chairs, as if to reassure us that a thorough analysis of the term will eventually be forthcoming. The purpose of the et cetera clause, then, is not to collude in the establishment of an ontology, but to help gesture at the outlines of a more intensive methodological investigation—it is a pedagogical device.

However, if we wish to avoid the triviality of the "etc.," then we are led to the other horn of the analyst's dilemma. Wisdom worries: "The only possible definition of 'chairishly related' will involve the expression 'and so on' "; it will involve an infinite disjunction of conjunctions of statements about sensations. And it will be complained that it is just this infinity which is mysterious".[4] Similarly, if we attempt to spell out the expression "and so on" with respect to the analysis of declaring war in terms of individuals and their relations, we are left with an infinite (or at least intractably large) disjunction of conjunctions of statements about individuals. If this is right, the pedagogical gesture implied in "and so on" is misleading: an analysis will not be forthcoming, because there are principled reasons for thinking that it is impossible to complete the analysis.

Why does the analysis of a declaration of war or chair lead to an intractably large number of statements about individuals for Wisdom and Urmson? The question the atomists are trying to answer is "*How* can England declare war (in 1939)?" Let us rather consider the following question: "How can the United States declare war (today)?"

There are a large number of necessary conditions which must obtain for the United States to declare war. Article I, Section 8 of the United States Constitution states that "The Congress shall have power . . . To declare War." Formally, Congress must pass a resolution empowering the President to wage war; but a whole host of conditions must be in place in order for this authorization to count as a declaration of war. Following the atomist program, each one of these terms—Congress, President, resolution, etc.—would have to be further analyzed into their components. For example it would be noted that one branch of the legislature consists of 100 individuals elected in accordance with the 17th amendment of the Constitution, so that each member is at least thirty years of age, has been a citizen for nine years, and is a

resident of the state which he or she represents. New terms are introduced: Constitution, amendment, state, etc. Atomism, in its most strident form, would maintain that a complete articulation of what it is for the United States to declare war would involve further explication in terms of, perhaps, individual sensations and their relations. Then we will have made some progress in answering the question, "What is it for the United States to declare war?"

Unfortunately, this analysis, even if it could be undertaken, is not yet complete. In analyzing the statement, we would have to compound the already difficult task of specifying the conjunction of individuals and their actions which are, in fact, individually necessary and jointly sufficient for the United States declaring war, with the impossible task of specifying all of the alternative ways that the United States *might* have declared war, but did not. The War Powers Resolution, for example, gives the President the authority to use force in certain circumstances, although it stipulates that the President must obtain congressional approval if the deployment is to extend over 60 days; even under the strictly formal sense of "declare," this might be thought of as a standing declaration applicable to specified conditions. In the case of defending the United States against an immediate threat, the President does not need congressional authorization to wage war, without stretching the formal sense of the term too much. In this case, the President is authorized both to declare and wage war.

But if we are interested in articulating how we in fact understand a particular institution, it will not do to exclude the number of informal ways by which the United States might be said to declare war. While such usage might raise eyebrows (consider those who insist that a tomato is a fruit because it develops from the ovary of a flower), it takes a particular context to rise to the level of being *incorrect*. In *Youngstown Sheet and Tube Co. v. Sawyer*, one of several cases that establish executive powers in making war, Justice Jackson writes:

> in the absence of either a congressional grant or denial of authority, [the President] can only rely on his own independent powers, but there is a zone of twilight in which he and Congress may have concurrent authority, or in which its distribution is uncertain. Therefore, congressional inertia ... may sometimes ... enable, if not invite, measures on independent presidential responsibility.[5]

Because "theatre of war" is an "expanding concept",[6] formal authorization is not necessary: congressional silence is a means by which a President's military action can be justified. In this case, congressional silence functions as a kind of declaration of war, if declarations of war are simply that which "enable, if not invite, measures of independent presidential responsibility". Moreover, if a President were to use force against Congress's wishes, Jackson's analysis only demands that the "Supreme Court would 'scrutinize with caution' whether the Commander-in-Chief possessed that power" after the fact;[7] whatever the subsequent findings of the Supreme Court, there is still a sense in which war was declared.

As we move away from the formal usage, the number of ways in which the United States could be said to have declared war increases indefinitely. Urmson imagines much less plausible but possible means by which England (or the United States) could be said to have declared war.

> There is an indefinite range of things that people might have done, any set of which would have counted as England declaring war. There might have been a revolution in 1938 and the war might have been declared by a mass meeting or a revolutionary junta; and this would count as ... declaring war. This did not happen; but we can understand the statement "England declared war" without knowing that it did not happen.[8]

When Urmson points out that "we can understand the statement," he is pointing to our ordinary competence with these

words; but more generally, this linguistic competence presupposed a kind of general understanding which underlies our ability, not just to talk, but act in appropriate, purposive ways.

Each one of these ways by which the United States could be said to have declared war would involve a large conjunction of statements, like the standard case of Congress passing a resolution authorizing a President to wage war, which are individually necessary and jointly sufficient. Explicit congressional authorization is not a sufficient condition for the declaration of war; Wisdom and Urmson point out that it is not even necessary. If Urmson is right, and "an indefinitely long list of alternative ways of declaring war could not be completed even in theory," then we are left with what Wisdom called an "infinite disjunction of conjunctions of statements about sensations".[9,10]

If philosophical analysis proceeds by formalizing our widely shared intuitions about what a desk is, or what it is for the United States to declare war, then the task proves impossible: it is always possible to imagine new situations not covered by the current definition. The statement, "The United States declared war" is perfectly intelligible even to someone who did not understand that formally the Constitution accords this authority to Congress; or at least there is a perfectly ordinary way of understanding this statement which does not require this knowledge. If Wisdom and Urmson are right, an infinite number of things the United States could have done is compatible with what a person normally understands by these words. If this understanding is the guide by which inquirers uncover and articulate the form of fact, then a given analysis seems impossible.

Wisdom and Urmson are drawing our attention to a disturbing feature of the atomist (or, for that matter, the positivist) view: they demand from articulateness what it cannot deliver, so that if nothing is to be left unstated about what it is for the United States to declare war, we will be saddled with an unmanageable infinitude. In the face of this infinitude we may choose to cut the analysis

short, so that we answer the question "What is it for the United States to declare war?" with "Congress explicitly passes such-and-such resolution." This avoids the infinite disjunction of conjunctions that worries Wisdom, but it is incompatible with our understanding of what we mean by "The United States declaring war". However, we understand that congressional action is neither sufficient nor necessary to the United States declaring war. While the alterative ways by which the United States could declare war are progressively implausible articulations of what we could mean by this statement, we do not want to unconditionally exclude these possibilities.

That the atomist had been expecting something impossible from our ability to characterize what we mean by a given proposition has been shown in three ways. Given what the atomists demand, if we cut off a given analysis with an etc-clause, we have provided a logically adequate analysis, but we are left with a trivial articulation. But if we try to spell out the etc-clause, we are left with an infinitude—completing the analysis proves to be impossible. Finally, if we try to cut off the infinitude prematurely, we end up making claims that are incompatible with our understanding of the ways in which the United States might declare war.

Wisdom and the impossibility of institutional analysis

To the extent that Searle's views align with the atomists, Wisdom's objection represents a serious obstacle to the possibility of institutional analysis. I will show how Wisdom's objection can be tailored to the case of institutional analysis by considering two examples: golf's "loose impediment" rule and money.

Loose impediments in golf

The *Rules of Golf*, published by the USGA, stipulates 34 rules that govern the game of golf.[11] The rules have changed over time, as

reflected in the fact that the book is updated every four years. For example, in 2004 a serious breach of etiquette became explicit grounds for disqualification. Rule 23 defines "loose impediments" as "Natural objects such as stones, leaves, twigs, branches and the like, dung, worms, and insects and casts or heaps made by them, provided they are not fixed or growing, and not solidly embedded and do not adhere to the ball." A loose impediment may be removed, so long as the ball is not moved in the process. A loose impediment may not be removed if the ball is in a hazard or already in motion. In the 1999 Phoenix Open, Tiger Woods appealed to the rule in order that a 1,000-pound boulder be removed as a loose impediment. The tour official ruled that since the boulder was not "solidly embedded" it does indeed *count as* a loose impediment.[12]

It is not possible to anticipate all the situations in which something would or would not count as a loose impediment. In *Rules of Golf* itself, the rule is clarified by appeal to numerous "decisions" or illustrations. So, for example, we are told that a dead snake is a loose impediment while a live snake is an outside agency (interestingly, worms and insects count as loose impediments). An ant hill, a fallen tree not attached to its stump, and the gravel on a road are loose impediments, while an orange peel which has adhered to the ball, and dew or frost in the line of putt are not.

According to Wisdom, we can always imagine alternative factual situations that are not covered by a previous understanding of the claim, "United States declared war". Analysis proves intractable. If the aim of institutional analysis is to provide the means to describe institutional phenomena without remainder, it must likewise fall. Consider the case of loose impediments described above. While it may have been clear to tour officials, upon being confronted with the case, that gravel, dead snakes, and ant hills, count as loose impediments, specifying what a loose impediment is in advance is an impossible task. Hence the stir caused by Woods' call to move a half-ton boulder on account of the rule.

We should expect that in the future, rule 23 will be pushed in equally surprising and novel ways. This is a vivid illustration of Wisdom's objection. The best we can expect from our articulations is a rough statement in the form of a rule buttressed by a list of exemplary cases—this is precisely what the *Rules of Golf* gives us.

Money

I have looked at how Wisdom's objection has played out with respect to a rule of the game of golf: we can imagine new and novel ways the "loose impediment" rule in golf might be implemented. Searle's archetypical institutional fact is money. Money, for Searle, is defined as a medium of exchange. A number of different brute objects (X) can come to count as money or a medium of exchange (Y). Much like declaring war, the analysis invariably implicates many other institutions, each of which will require its own analysis. For example, certain kinds of money must be printed by the Bureau of Printing and Engraving (although there are contexts in which counterfeits are nevertheless knowingly circulated and accepted as media of exchange). These institutions will in turn implicate other institutions, and so on. Even if we determine that something being a green piece of paper with such-and-such markings is a necessary condition of what it is to be money, a full articulation of the other conditions seems to be a nearly impossible task.

The problem is, following Wisdom's objection, green bits of paper do not even seem to be *necessary* to something's being money: an indefinite number of other objects could also come to count as a medium of exchange. Searle himself talks about green bits of paper, gold coins, electronic data, Marlboro cigarette packs, and cowrie shells as different things that might count as money. Even if we confine our attention to paper money, there is no limit to the arrangements of colors and markings we could imagine to count as money. Moreover, anthropologists have identified

many more such objects: woodpecker scalps, goats, pig tusks, porpoise teeth, wampum, feathers, carved stones, potlatch coppers, dog teeth, and ceremonial axes. Indeed, if money is just a medium of exchange, in a barter situation *any* tradable object counts as money.

Looking at money in particular, there may be a way out of Wisdom's difficulty. If institutional analysis requires that we explicitly specify every X that could count as money, it is easy to see how the analysis would quickly become unwieldy. One way out of this problem is to propose a definition that would identify what all these different X's have in common. This might save us the trouble of having to list out every different X. Perhaps, then, bits of paper, gold coins, electronic data, Marlboro packs, and woodpecker scalps share the property of being *media of exchange*; and any object that falls under this definition counts as money. In that way the institutional analysis seems to have become tractable again.

This suggestion is problematic. If "media of exchange" is a definition, it seems to be a definition of the institutional fact *money* (Y) and not the brute facts (X) of which money consists: money *is* a medium of exchange whereas bits of green paper *count as* a medium of exchange. This suggestion is consistent with Searle's usage in the *Construction*. It would be a mistake to identify the green piece of paper as a medium of exchange in the same way it would be a mistake to identify the paper as money—while X counts as Y, it is not the case that X alone can be identified with Y. Searle has given us some indication of what all money has in common, not what all the objects that might count as money have in common. Simply saying that money is a medium of exchange does not get us any closer to resolving Wisdom's objection: we are still left with an infinite disjunction of conjunctions when we try to articulate the X term.

We are left with the task of specifying what the continuum of X's have in common *without* being able to appeal to the very status-function that make shells, bits of paper, and woodpecker skulls

significant in the first place. Is there any physical property we can point to which these objects have in common? To be useful as a medium of exchange the brute fact ideally must be divisible, rare, easily distinguishable from objects which are not to serve as money, storable, durable, and portable.[13] Indeed, it is to these sort of criteria that anthropologists would appeal in explaining why so many societies would tend to employ gold and silver as a standard unit of exchange.[14] But while these criteria serve to exclude certain brute facts—common river pebbles, planets, trees, mountains, bodies of water—from counting as media of exchange, these guidelines are nevertheless repudiated by the historical record. Cattle are not easily divisible. It is not clear that woodpecker scalps are easily distinguishable from the scalps of other small birds.[15] Packs of cigarettes are not particularly durable. Finally, with respect to portability, the Yapanese Islanders in the Western Caroline Islands of Micronesia used large stone disks as money, some of which had a diameter of at least four meters and weighed more than nine metric tons.[16] If these usability criteria help account for why societies tended to adopt precious metals, and ultimately paper, as media of exchange, it is not clear that these principles can also be relied on to cut through Wisdom's infinitude. Every rule would require additional rules to account for actual and imagined deviations. Such a system of rules is bound to become intractably large. Any attempt to systematize the kinds of brute fact that would count as, say, a loose impediment runs into an identical problem.

In looking at money, we have seen that the X term of the constitutive formula can be characterized indefinitely. I would now like to suggest that this is also true of that which is denoted by the Y term.

What is money? What does money do? Searle writes that "For money, the statuses have typically been imposed on bits of metal and paper, and the function are those of serving as a medium of exchange, repository of value, etc." (*CSR*, p. 83). Searle's answer

can be directly traced back to the orthodoxy of contemporary economics textbooks. Consider this representative passage:

> *Money* is the standard object used in exchanging goods and services. In short, money is the *medium of exchange.* In a system of monetary exchange, people trade money for goods when they purchase something and trade goods for money when they sell something, but they do not trade goods directly for other goods.[17]

The passage above hints at a familiar story or history which explains and/or justifies[18] the claim that some objects (X) *count as* money (Y), defined as a medium of exchange (Y).

In the textbook account, monetary exchange is explicitly contrasted with barter. A society which uses some good or subset of goods as a kind of intermediary can more efficiently trade goods than a society that does not. For example, there must be a coincidence of wants: under a barter system an individual with a surplus of some good must find an individual with a deficiency of the same good, and who also has a surplus of something the first individual wants. There must also be the further coincidence of timing. A society that has an intermediate commodity allows the individual to sell the good, and perhaps later, with a different individual, purchase something needed. If money is a medium of exchange (Y), then these objects (X) *count as* a primitive form of money. As the system becomes more efficient and goods become more saleable, the number of intermediate goods that serve as a medium of exchange declines until there is perhaps one such good. Historically this process is said to culminate with the use of precious metals as the dominant medium of exchange, perhaps in the form of coins (which overcome the inconvenience of having to weigh the metal in every transaction). Such objects are easily divisible, have high inherent worth, and are easily manageable. According to the orthodoxy, the invisible hand of the market is responsible for these and other economic innovations:

> Market forces will also lead mints to issue coins of standardized weights and fineness, so coins will be issued in standard dollar amounts. Any mint that issued non-standardized coins would impose additional inconvenience on its customers and have to charge a lower minting fee to compensate them. It would therefore find it difficult to survive against competitors who issued standardized coins.[19]

Developments in efficiency continue with the introduction of paper money (X) that initially represents deposits of gold or coins. As a medium of exchange, paper facilitates economic transaction (there is no need to carry and protect large caches of coinage) and enables new banking businesses that specialize in the storing and handling of these deposits. Because the overall redemption of these deposits is normally quite low (after new deposits are subtracted out), bankers may freely lend out and collect interest on the gold that is not required for day-to-day transactions. In an advanced economy, market forces will drive financial instruments (i.e., the transferring of debts and assets of other firms and individuals) to replace gold as a standard redemption medium.

A common subplot in the story occurs with the introduction of paper money. At this point, market forces, which have thus far proved largely benevolent, may be interfered with or undermined as governments and banks acquire unprecedented control over the money supply.[20]

Recently a number of economists and social scientists have started to question the orthodox model of economic interaction.[21] I will now survey one such account, which Randall Wray calls the "chartalist view".

The force of orthodox or real analysis derives from an intuitive portrait of the kinds of transactions that might take place in a hypothetical village fair. According to the origins account, barter—the exchange of goods—is the principal economic action. In Searle's terms, objects like shells, coins, and paper (X) count as

media of exchange (Y). Other functions money serves—a unit of account or a store of value—are regarded as secondary, inessential and perhaps derivative.

According to the chartalist, setting aside the model of the village fair, Wray and others look towards a different institution on which to model money. They point out anthropological research does not support the orthodox genealogy—"There is no evidence of barter-based markets (outside trivial prisoner-of-war cases)."[22] They use such evidence to dispute, for example, the textbook claim that most societies eventually gravitate toward gold and silver coins. Contrary to the textbook account, Wray points out that, for example, early coin denominations were much too high to serve the purpose attributed to them by the orthodox view:

> For example, the most common denomination of the earliest electrum coins would have had a purchasing power of about ten sheep. They might have sufficed for wholesale trade of large merchants, but they could not have been used in day-to-day retail trade. It is also quite unlikely that coins would have been invented to facilitate trade, as the Phoenicians, and other peoples with sophisticated trade managed without coins for many centuries.[23]

What function (Y), then, does money serve? The evidence points away from money as a vehicle of the village fair, and to its role in early protectionist, taxation, or tort schemes. Wray draws our attention to the "price lists"[24] or ornate and fixed compensation schedules[25] which can be found on the carefully protected clay tablets of early societies. These lists existed for a variety of different reasons: they specified debt owed for protection and services provided by the sovereign authority, or debt owed for various injuries incurred on another (a practice that helped prevent blood feuds). The names or descriptions of types of individuals inscribed on these price lists, or indications of debt in general, are money according to Wray. These names or descriptions distribute the

rights and obligations indicative of institutional facts; just my owning a dollar bill indicates that I have the right to purchase certain items (*CSR*, p. 105). Money, in both cases, provides a means to distribute certain kinds of rights and obligations. The compensation schedules were not hammered out through an ongoing process of bargaining by the affected parties, much less by the "continuously transacting economic agents" of the orthodox account. Generally, these price lists were determined in advance by the existing sovereign authority, public assembly or custom. The compensation was listed first in terms of a percentage of production, then in standardized units of perhaps wheat or barley, and eventually in terms of coins.

According to the chartalist view, what objects (X) are to count as money? Whatever object is used to tabulate *debt* of particular individuals counts as money. Money, then, is defined as an indication of debt-owed that takes the form of promissory notes and lists. These lists specify who owes (either in the form of a description—"all subjects owe . . ."—or by name), to whom (the king, the injured), and how much (units of grain); they might also record other information, such as the date or the seal of the receiver. In early societies, then, names and descriptions inscribed on the huge clay tablets protected in temple vaults are what count as money.

In the Introduction, I used the hazelwood tally as an example of an institutional fact. In medieval Europe the tally was used to calculate debt: the tally consisted of a hazelwood stick on which was inscribed the debtor's name and the date. The stick was partitioned in two from about an inch from the bottom. The longer half was kept by the creditor and the shorter half was kept by the debtor. Because the two halves of the hazelwood stick could be put back together, this ensured that the amount of debt could not be tampered with. In Searle's terms the hazelwood stick (X) *counts as* an indication of debt-owed or credit (Y), depending on which half of the tally one possesses. Insofar as these indicators can be

identified with money it turns out that the hazelwood stick (X) also counts as money (Y).

Since private parties could also be owed money, there existed the possibility of debt reconciliation and transfer between multiple parties. If I owed 5 units of grain to the king, but was owed 10 units of grain by another party, I could reconcile and transfer accounts so that the other party owed 5 units of grain to the king and 5 units of grain to me. As smaller clay tablets were circulated, debts owed and loans made could be more easily reconciled. Wray goes on to suggest that the medieval fairs and markets on which the orthodox model is based were in fact "developed to act as clearing houses, allowing merchants to settle their mutual debts and credits without the use of a single coin. While textbooks say that these fairs were early markets, the retail trade probably originated as a sideline to the clearing-house trade".[26]

As mentioned, the crown had a fixed compensation schedule, often in the form of annual taxes. The crown also needs to make payments to the military and administrative bodies that support it. It may have done so in the form of coins, which—like clay tablets or tallies—are indications or tokens of the crown's debt to these parties. There is nothing special about coins in the chartalist genealogy; all of these objects are tokens, indications, or evidence of the crown's debt. Because there are subjects that also owe money to the crown, if a tax payer can acquire these tallies or coins from the public servants then the king's debt to the public servants is transferred to the tax payer, negating tax debt. How does the tax payer acquire tallies or coins from the public servants? By exchanging goods or services for these tokens. It is only at this point, as per the orthodox account, that money, which is normally an indication of debt-owed, also serves as a medium of exchange. For the chartalist, this function of money as a medium of exchange is an important but not necessary function of money: as in the case of feudal Europe sometimes the services provided to the crown (so that the crown is a debtor) was negated by debt incurred

by that *same* person in the form of taxes, rent, etc. This is, of course, the familiar story of the company store.

Wray sums up the chartalist view on the origins of money as follows:

> An inordinate focus of economists on precious metal coins and market exchange then appears to be misplaced. The key concept is debt, and specifically, the ability of the state to impose a tax debt on its subjects. Once it has done this, it can choose the form in which subjects can pay the tax. Certainly the government's tokens can also be used as a medium of exchange, but this *derives from* its ability to impose taxes, and is necessitated by the imposition of the tax (if one has a tax liability but is not a creditor of the crown, one must offer things for sale to obtain the crown's tokens).[27]

According to this origins-story, a radically different picture of the function of money begins to emerge. In this account, whatever object is used to tabulate debt counts as money. Given my purposes, I will focus on the following feature: objects (X) that count as money are not media of exchange (Y_1) but rather "evidence of debt" (Y_2).[28] That is, on this view, money is less a means of accretion, and more a means of compensation.

There are a number of different features that other authors emphasize about this particular account of money and the way in which it contrasts with the orthodox view. First, while there is great overlap between those objects which count as money in the orthodox view and those objects which count as money in the chartalist view, there are also divergences. While both accounts take gold coins to count as money, clay tablets or hazelwood tallies would not count as money according to the orthodoxy. According to the orthodoxy, but not the chartalist, potentially *any* object of barter is a medium of exchange, and thus money.[29]

Second, because money is an indication of debt-owed, its advent requires an institution large enough to create demand for its

tokens and tallies in the form of debt on others (taxes, compensation). Only then will public servants have confidence that their tokens and tallies will be redeemable in the form of goods and services. Contrary to free-market advocates, this story suggests both a much more prominent role for central banks and government institutions in the creation and regulation of money and concrete policy implications. Indeed, Wray thinks this picture implies that "the normal requirement is for a government deficit".[30]

Should Searle be concerned about methodological difficulties?

Following Wisdom's objection to philosophical analysis, I have argued that the difficulty is easily extendable to the case of institutional analysis; if something counts as money, it seems that there are an indefinite number of ways to recharacterize both the X and Y term. The various ways we might frame money is different from a dispute as to whether disputed territory X counts as India (Y_1) or China (Y_2).[31] In the case of a piece of gold, all participants may agree that it counts as money, but there remains a disagreement about how to characterize the significance of X among institutional analysts. Unlike the disputed territory, both the orthodox and the chartalist accounts seem *compatible* with the way in which X happens to be used.

This objection concerns methodology, and outlines a difficulty in applying the constitutive formula to say something about the contours of our actual institutions. In that way, it is not clear that Searle would be impressed with its force. It may be the case that, in fact, our institutions must be rendered in terms of a variety of overlapping characterizations, but Searle might point out that what is important is that each of these characterizations has the form "X counts as Y in C," and it is this form that is unavoidable. If this view is even coherent, perhaps Wisdom's methodological

difficulties can be bracketed—Searle is interested in institutional reality's logical structure and is quite prepared to leave it to the social scientists to contend with difficulties involved in saying something substantive about what our institutions actually look like. Again, we return to something like the under-labourer conception of philosophy characteristic of the atomist reading of the *Construction*.

Nevertheless, I think it is strange that the constitutive formula purports to outline the logical structure of institutional reality, and yet, except on the most general level, something like money does not appear to have a canonical structure. Organized society, on this view, is an infinitely complex amalgam of facts, all of which have the form "X counts as Y in C". This tension in articulability is surprising, and at least requires some explanation. To anticipate the next chapter, this tension is especially surprising if the constitutive formula is the structure, not just of institutional reality, but of the mental and linguistic representations that give rise to this reality (*CSR*, p. 90). On this view, our own intentional structure is now burdened with Wisdom's infinitude.

Notes

1. John Wisdom, "Metaphysics and Verification," *Mind* 47, no. 188 (1938): pp. 476–7.
2. Ibid., p. 478.
3. Wisdom frames the question as the possibility of analyzing "a declaration of war" as a positivist might: he appears to be strictly concerned with only the meaning of the terms. The atomist is likewise concerned with its meaning but is explicit in maintaining that its proper analysis tells us something about the world, what "England's declaring war" in fact denotes. The atomist and positivist would analyze the expression in the same way, but would articulate the significance of that analysis in different ways. For reasons that I have indicated above, namely that Searle does not seem to be concerned just with the meaning of the word "institution," I am more concerned with the atomist's articulation.

4. Wisdom, "Metaphysics and Verification," p. 478.
5. Justice Jackson in Henry Holzer, *Declaring War* (FrontPageMagazine.com, September 5 2002 [cited]); available from http://www.frontpagemag.com/Articles/Printable.asp?ID=2742.
6. Justice Black in ibid.([cited]).
7. Ibid.([cited]).
8. Urmson, *Philosophical Analysis; Its Development between the Two World Wars*, p. 152.
9. Wisdom, "Metaphysics and Verification," p. 478.
10. William Ramsey lists concision as a criterion of conceptual analysis. "It is assumed that overly complex and unwieldy definitions are defective or ad-hocish, even when no better definition is immediately available. If an analysis yields a definition that is highly disjunctive, heavily qualified, or involves a number of conditions, a common sentiment is the philosopher has not gotten it right yet. The proper definition should be short and simple". William Ramsey, "Prototypes and Conceptual Analysis," in *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, eds Michael R. DePaul and William Ramsey (Lanham, Md.: Rowman & Littlefield, 1998). If Wisdom and Urmson are correct, there are principled reasons for thinking that this criterion will always be violated.
11. United States Golf Association, *Decisions on the Rules of Golf, 2004–2005* (Far Hills, NJ: USGA, 2004).
12. Arnold Palmer, *Playing by the Rules: All the Rules of the Game, Complete with Memorable Rulings from Golf's Rich History* (New York: Pocket Books, 2002), pp. 184–5.
13. William J. Baumol and Alan S. Blinder, *Economics, Principles, and Policy* (New York: Harcourt Brace Jovanovich, 1979), p. 201.
14. In Chapter 5 I will suggest that the anthropological evidence for this claim is actually quite weak.
15. Baumol and Blinder, *Economics, Principles, and Policy*, p. 201.
16. Scott M. Fitzpatrick, "A Massive Undertaking: Examining Stone Money in Its Archaeological Context," *Antiquity* 76, no. 1 (2002): pp. 331–2.
17. Baumol and Blinder, *Economics, Principles, and Policy*, p. 200.
18. I will look more closely at the significance of such stories below.
19. Kevin Dowd, "Aristotle on Money," in *What Is Money?* ed. John N. Smithin (London; New York: Routledge, 2000), p. 144.

20. L. Randall Wray, "The Property Theory of Interest and Money," in *What Is Money?* ed. John N. Smithin (London; New York: Routledge, 2000), p. 42.
21. In fact these economists and social scientists have only recently started to resurrect a view that dates back to Keynes and earlier authors.
22. Wray, "The Property Theory of Interest and Money," p. 42.
23. Ibid., p. 45.
24. Geoffrey Ingham, "Modern Money," in *What Is Money?* ed. John N. Smithin (London; New York: Routledge, 2000), pp. 24–6.
25. Wray, "The Property Theory of Interest and Money," p. 43.
26. Ibid., p. 45.
27. Ibid., p. 47.
28. Ibid., p. 46.
29. It is precisely this consequence of the orthodox view that moves John Smithin to worry about proclamations such as "money is neutral" or "money is a veil". Ironically, "orthodox economics wanted to create a science that ignored money." "Underlying this perspective," Smithin writes, "is the view that economics deals fundamentally with the so-called 'real' exchange of goods and services, as opposed to the accumulation of financial resources." Smithin, "Introduction," p. 63.
30. Wray, "The Property Theory of Interest and Money," p. 59.
31. Smith and Searle, "The Construction of Social Reality: An Exchange," p. 293.

3

Second Criticism of Searle's Institutional Atomism—Metaphysics

Searle's program can be understood as importantly analogous to that of the atomists. The way in which the atomists assume that brute reality is specified by means of philosophical analysis resembles the way in which Searle supposes institutional reality can be depicted by means of the constitutive formula.

Is Searle in a position to argue that the constitutive formula represents the logical structure of institutional reality? Does Searle posit a structural isomorphism between the constitutive formula and institutional reality? Searle appears to argue that our confidence in the veracity of the constitutive formula is justified because institutional reality contains linguistic representations of the form X counts as Y in C.

But this claim may be incompatible with other aspects of Searle's own view. It is not clear how Searle is justified in holding that institutional reality has the structure posited by the constitutive formula. Perhaps our institutions have another structure, or that structure is such that it cannot be represented by any one model or schema.

John Austin and the relation between metaphysics and methodology

For the atomist there is a straightforward sense in which metaphysics justifies methodology: the philosophy of logical atomism vindicates the practice of philosophical analysis. The metaphysics

describes higher-order objects as logical constructions, arrangements of building blocks or atomic facts. The metaphysics assures us that our talk about high-order objects is shorthand for a series of atomic propositions that reflect the structure of these atomic facts. Worrisome propositions are eliminated in favor of other less problematic terms, just as "the King of France" is eliminated in favor of a description. The atomists generally agreed that the building blocks themselves were sense-data, which moves us closer to the realm of methodology or analysis, where we try to give flesh to the *a priori* frame. Why are the analyst's descriptions of the world less misleading than those of the ordinary language user? Because, citing the metaphysics, the analyst represents the facts better than does the ordinary language user; philosophical analysis employs a structure that is more similar to the form of fact. In this way the metaphysics justifies the analytical practice—without a conception of the way the world is, we would not be able to say that our analysis more precisely corresponds to that world.

For Searle, institutional analysis is justified by the structure of institutional reality, the form of institutional fact. To see this, it will be helpful to look more closely at Searle's defense of the correspondence theory of truth.

Searle's rejection of the idea that words mirror the form of fact

The last three chapters of the *Construction* are a matter of philosophical housekeeping (*CSR*, p. 199). In Chapters 7 and 8 Searle discusses realism, and in Chapter 9 he defends a version of the correspondence conception of truth, driving a wedge between the book's earlier and later chapters. For example, it is a tricky matter to establish a connection between the formal thesis of "external realism," discussed in Chapters 7 and 8, and the more robust conception of "brute fact" found in Chapters 1–6.[1] Likewise, in Chapter 9, the connection between Searle's defense of the

correspondence theory and the rest of the *Construction* is uncertain. Searle is candid about this: "My conception of social reality does not logically require the correspondence theory of truth—someone could reject the correspondence theory and still accept my analysis—but the overall picture I, in fact, hold proceeds by way of external realism through the correspondence theory to the structure of social reality" (*CSR*, pp. 199–200). Searle's remarks aid my own investigation.

Chapter 9 picks up from a conversation that began over 40 years ago in a debate between J.L. Austin and Peter Strawson,[2] and has since been taken up by those who champion the disquotation theory of truth. Disquotation, aligned with the Strawsonian position, is contrasted with Austin and the correspondence theory. While Searle *defends* Austinian correspondence, the strategy he employs recommends "reconciliation" as the better descriptor:[3] "my investigation at this point is a Wittgensteinian-style enterprise into the language games we play with these words, and its aim is to remove the false pictures that our misunderstandings of the language games engender" (*CSR*, pp. 209–10). Both correspondence and disquotation tend to provoke false pictures, albeit different ones: once we are clear as to what is *not* being said, there remains a sense in which it does not matter how we characterize "truth".

Searle's discussion concerns the status of statements like "*p* is true," where *p* is a statement or proposition about the world—say, "the coffee cup is full". Strawson, with qualifications, embraces the redundancy theory of truth, where " 'the coffee cup is full' is true" is just another way of saying "the coffee cup is full"; only, that the latter formulation is in general preferable because it does not contain the word "true," which fails to denote any property or relation. In that way, "is true" is "logically superfluous". Austin disagrees, suggesting that "true" in " 'the coffee cup is full' is true" *does* tell us something, not only about the world, but also about the statement "the coffee cup is full".

This brief sketch of the debate is necessary, first, to specify in what way it is *not* applicable to the present discussion, and second, to specify in what way it *is*. I am investigating the significance of institutional facts evoked by sentences like "The president signed the bill". This investigation concerns statements such as "the president signed the bill," and not statements like " 'The president signed the bill' is true". We *could* discuss this, and if we did, the Austin/Strawson debate on truth would be a good place to begin such an investigation.

However, there *is* a sense in which Searle's evocation of the Austin/Strawson debate furthers my discussion of the constitutive formula. It provides the seed for an objection *against* the suggestion that institutional analysis bears the requisite similarity to the atomists' version of philosophical analysis. When looking at statements, not of the form "*p* is true," but "*p*," what is striking is how much agreement there is between Austin and Strawson. Austin rejects the atomist model that the structure of "*p*" somehow exactly "pictures," "mirrors," or "is congruous with" the form of fact. Austin uses the word "conventional" to flag an emphatic break from the atomist model of structural isomorphism between word and world:

> The only essential point is this: that the correlation between the words (= sentences) and that type of situation, event, etc. . . . is *absolutely and purely* conventional. . . . There is no need whatsoever for the words used in making a true statement to "mirror" in any way, however indirect, any feature whatsoever of the situation or the event; a statement no more needs, in order to be true, to reproduce the "multiplicity," say, or the "structure" or "form" of the reality, than a word needs to be echoic or writing pictographic. To suppose that it does, is to fall once again into the error of reading back into the world the features of language.[4]

Strawson takes this, plus the redundancy of statements of the form "*p* is true," to be adequate reason to drop the notion of *correspondence* altogether: "the correspondence theory requires, not purification, but elimination".[5] Nevertheless, Strawson is not suggesting that our sentences do not say something about the world, and to the extent that they do, he would agree with Austin that our statements are "conventional".

More importantly, Searle—who is largely sympathetic to Austin's views on truth—does not argue against this point; there is good reason to think that, as far as "*p*" goes, Searle agrees with Austin in contending that our words are largely "conventional" in the way they correspond to brute reality. Searle is clear about this in the *Construction* when he writes that "The world divides up the way we divide it, and if we are ever inclined to think that our present way of dividing it is the right one, or is somehow inevitable, we can always imagine alternative systems of classification" (*CSR*, p. 160). It is this commitment to the possibility of alternative systems of classification that drives Searle's rejection, not of the correspondence theory per se, but of one of the misleading pictures that it tends to engender: "the picture that facts are complex objects or events and that truth consists of a kind of matching or isomorphism between the elements of the statement and the elements of the fact is absurd" (*CSR*, p. 205). So long as we are not seduced into thinking that there is no world, that nothing is given, Searle thinks this picture—that our words are conventional in Austin's sense—is largely right.

Here is the problem: I have relied on the atomist model of "mirroring" and "structural isomorphism" to help characterize the significance of the constitutive formula. According to Russell, the world consists of particulars and their relations; philosophical analysis gives us the tools to replicate this structure, so if we want an account of the world, we must analyze it in terms of these particulars. Likewise, because the constitutive formula mirrors, corresponds with, etc., the form of fact, institutional analysis is thus

justified. However, Searle, following Austin, explicitly embraces a form of conceptual relativism that is contrasted with the views of Russell and his followers. How, then, can we understand the constitutive formula through an extended comparison with the atomists? That is, how can our institutional utterances exhibit the "form of fact," if the meaning of our words are conventional in Austin's sense?

This question can be answered, and can be answered by *not* denying Searle's allegiance to Austin's version of correspondence. I believe that it can be answered in such a way that the comparison between the atomist's version of philosophical analysis and Searle's institutional analysis is advanced.

Defending the comparison between Searle and the atomists

The key to defending the comparison between philosophical analysis and institutional analysis is to see that the last three chapters of the *Construction* explicitly concern the relation in which our statements stand to *brute* fact, and cannot be taken to imply anything about how statements describe institutional reality. Searle's allegiance to Austin's suggestion, that words denote, not in virtue of the form of (brute) fact, but in virtue of conventionally assigned meaning, remains intact—but only insofar as it concerns statements about brute reality. Perhaps, for Searle, the descriptions of our institutions under the constitutive formula and institutional reality exhibit the same relation of "mirroring" or "structural isomorphism" that characterizes the atomist outlook.

Not only does the constitutive formula describe the structure of institutional reality, but it is structurally isomorphic with this reality. With the constitutive formula, Searle takes himself to be describing the logical structure of organized society. Searle does not just suggest that the constitutive formula is structurally isomorphic with institutional reality, but that it is a structural *feature* of that reality. Institutional facts are, by Searle's own lights,

language-dependent facts. Searle says, "it is a sufficient condition for a *fact* to be language dependent that two conditions be met. First, mental representations must be partly constitutive of the fact; and second, the representations in question must be language dependent" (*CSR*, p. 62). It follows readily that Searle thinks the institutional analyst who employs the constitutive formula, is describing institutional reality in a way which corresponds to that reality.

Genuine correspondence is then possible for the institutional analyst, even to an extent not available to the atomists themselves, because the facts the constitutive formula are meant to describe are themselves *linguistic*. Austin worries about the possibility of reading the structure of language back into the structure of the world. This error prompts Austin to recommend a weak kind of correspondence, in which the meaning of our words is assumed to be "*absolutely and purely* conventional". As far as our descriptions of *brute* facts go, Searle is in agreement with Austin. But the institutional analyst is immune to this criticism. Because institutional facts are themselves "conventional"[6] and thus "language-dependent," strong-correspondence of a kind sought after by the philosophical analyst turns out to be the prize of the institutional analyst. Searle's allegiance to Austin's conventionalism does not appear to threaten the comparison of Searle and the atomists, so long as we limit the comparison to descriptions of institutional facts.

CSR, p. 90: institutional reality and intentionality

Searle endorses a kind of strong correspondence between statements of the form "X counts as Y in C" and the constitutive rules that underlie our institutions.

The possibility of a strong correspondence between word and world is not the only reason why Searle wants to embed mental and linguistic representations into the very fabric of institutional

reality. Institutional facts are language-dependent. While indications that linguistic representations are constitutive of institutional reality are scattered throughout the *Construction* (*CSR*, pp. 37, 60, 62, 74), and are reiterated in later texts,[7] the strongest formulation of this view is on page 90:

> Because I am trying to describe the logical structure of organized society, it may be well to pause at this point to explain what is involved and to make explicit at least part of what is at stake. How can "organized society" have a "*logical* structure"? After all, society is not a set of propositions or a theory, so what is this talk about logical structure? On my account, social and institutional reality contain representations, not only mental representations, but even linguistic representations, as constitutive elements. These do have logical structures. I am attempting to lay bare the most fundamental of those logical structures.

If one of the aims of the *Construction* is to answer the question, "What is the logical structure of institutional reality?," Searle needs to say something about how institutional reality can have such a structure in the first place. Only propositions—not objects—have a logical structure. Searle answers by suggesting that institutional objects, like dollar bills and the presidency, are special kinds of objects that are in part constituted by mental and linguistic representations.

Why does Searle insist on embedding linguistic and mental representations into the analysis? The driving question of the *Construction* is, "How can there be an objective world of money, property, marriage, governments, elections, football games, cocktail parties and law courts in a world that consists entirely of physical particles in fields of force, and in which some of these particles are organized into systems that are conscious biological beasts such as ourselves?" (*CSR*, pp. xi–xii, 29). An appeal to explicit intentionality is required to answer Searle's own question. Searle exploits

the causal connection between institutional facts and intentionality in order to analyze institutions in terms of two primitives: brute facts (X) and collective intentionality.

However, we can imagine an account of institutional reality that looks very similar to Searle's, but that does not include collective intentionality or any claims as to how Y is *caused.* Hubert Dreyfus, in "Phenomenological Description Versus Rational Reconstruction," takes Searle to have offered a causal, phenomenological account of the creation of institutional facts, wherein institutional facts are "imposed on" or "assigned to" brute facts by conscious observers.

Dreyfus worries that the appeal to mental and linguistic representations makes a hash of the phenomenology. "If this is meant to be the causal claim that we 'conscious' observers and users actually 'assign' functions to brute stuff from 'outside' every time we observe or use some of the stuff as equipment, it contradicts the phenomenology".[8] Given the causal role Searle gives to collective intentionality it is not clear how we can read him as offering something less than a full-blown explanatory account of how institutional facts are created. Dreyfus contends that an overextension of Searle's otherwise plausible logical claim moves him to make irresponsible causal and phenomenological claims.

Dreyfus recommends that Searle drives a wedge between the "unquestionable" logical account of social reality and the problematic, empirically false phenomenological account, discarding the latter.[9] If the logical account is so tied to the causal account, a cacophonous inner life is falsely posited; Searle is doing "bad phenomenology".

Dreyfus' recommendation would result in an austere account of institutionality, wherein institutions are analyzed in terms of two primitives—brute facts (X) and institutional facts (Y), rather than collective intentionality. But the appeal to institutional facts is unsatisfactory, presupposing the very phenomena that Searle sets out to explain in the first place. He will have analyzed

institutions in terms of institutional facts, which is unhelpful. Searle needs the appeal to mental and linguistic representations or collective intentionality if he is not to beg the very question that the book purports to answer.

Difficulties with *CSR*, p. 90

Dreyfus and "bad phenomenology"

The answer to Searle's question seems to require the appeal to the causal efficacy of intentionality. For that reason, Dreyfus' austere analysis is unacceptable to Searle. But, then Dreyfus' own criticism remains unanswered.

In response, Searle criticizes Dreyfus for misunderstanding the aims of the *Construction*. "I make no claims about what *causes* make social reality possible.... I would assume the answer would list a whole lot of causes, most of them having to do with biological evolution. I am, to repeat, trying to describe the *constitutive* elements of social and institutional reality, not what causes us to be able to constitute them the way we do".[10] Indeed, it is often clear Searle is deeply sympathetic with Dreyfus' worry. Searle speaks of the "illusion that [a] person who is able to deal with money, to cope with society, and speak a language must be unconsciously following rules" (*CSR*, p. 142). Elsewhere he writes, "I have sometimes spoken as if the collective imposition of functions were always a matter of a deliberate act or set of actions. But except for special cases where legislation is passed or the authorities change the rules of the game, the creation of institutional facts is typically a matter of natural evolution, and there need be no explicit conscious imposition of function—whether status or other type of function—on lower-level phenomena" (*CSR*, p. 125–6).

But Searle's response to Dreyfus is unsatisfactory. If Searle is denying that intentionality has any place in his attempt to describe

the constitutive elements of institutional reality, then it appears that Searle has adopted Dreyfus' austere account and so begged his own question. Moreover, in the same response to Dreyfus, Searle seems cognizant of the importance of these representations to his program, as underscored in *CSR*, p. 90: "Contrary to [Dreyfus'] account, there is no wedge whatsoever between the ontology of institutional facts and the 'causation' by way of collective intentionality, because the imposition and maintenance of status-functions by collective intentionality is not something which just *causes* institutional reality, it is *constitutive of that reality precisely because it is constitutive of the ontology according to the constitutive rule*".[11] Here it is clear that Searle not only holds that collective intentionality *causes* institutional facts, but is also *constitutive* of that reality.

Searle seems to want to deny the charge that he is positing an untenable phenomenology, while still including collective intentionality among institutional reality's constituents. Is there any way to make sense of Searle's position?

In the *Construction* Searle helpfully describes the appeal to linguistic representations as largely propaedeutic: "To start with, in [the first] chapter and the next, I will use a first person intentionalistic vocabulary to try to lay bare certain elementary features of social ontology. Later, in Chapter 6, I will show how some, though not all, of the intentionalistic apparatus can be explained in terms of, and ultimately eliminated in favor of, what I have elsewhere called the 'Background' of capacities, abilities, tendencies, and dispositions" (*CSR*, p. 5, see also pp. 13, 126, 129, 142). Elsewhere, Searle writes that "I want to propose that in many cases it is just wrong to assume, and certainly unsupported by the evidence that has been presented in the course of these discussions, that our behavior matches the structure of the rules because we are unconsciously following the rules" (*CSR*, p. 145).

Contrary to the *CSR*, p. 90 quotation, these passages indicate that some constitutive rules need not require conscious, or even unconscious, representations. Where it is untenable to posit

explicit intentional imposition, Searle has recourse to a non-intentional Background of dispositions and attitudes. Given this, we should reject the *CSR*, p. 90 quote as an aberration, not reflecting Searle's overall view. If Searle can appeal to the Background instead of intentionality, then Searle may be able to provide a non-trivial analysis of institutional reality while avoiding a dubious phenomenology.

The Background as a proxy for intentionality

If Searle agrees that our institutions cannot be articulated in terms of brute facts and intentional imposition, might we appeal to Searle's concept of the Background as a proxy for mental and linguistic representations (when they are not present)? There is, after all, "a parallelism between the functional structure of the Background and the intentional structure of the social phenomena to which the Background capacities relate" (*CSR*, p. 142). Unfortunately, there are difficulties with the concept of the Background. I will argue Searle cannot use the Background to evade Dreyfus' criticism.

In Barry Stroud's essay, "Background of Thought," he points out that Searle recognizes that the mental cannot always adequately explain our engaging in rule-governed behavior. Searle agrees. In *Intentionality* he writes that it is *not* the case that "all Intentionalistic mental life and all cognitive capacities could be entirely reduced to representations".[12] He writes that intentional states or representations are rather "underlain" or "enabled" by a Background of "skills, abilities, preintentional assumptions and propositions, stances, and nonrepresentational attitudes".[13] Likewise, insofar as institutional reality is dependent on the intentional imposition of a status-function (Y), it is not the case that all such impositions must involve explicit representations. In many cases, Searle argues, the explicit intentional impositions can be "ultimately eliminated in favor of" the Background (*CSR*, p. 5).

But if it is *false* to contend that every participant follows rules in the form of conscious or unconscious representations, Stroud wonders whether the Background can seal this explanatory gap. Stroud points out that the more Searle tries to give flesh to the notion of the Background, the more difficult the Background is to understand. Searle vacillates between two characterizations of the Background, one emphasizing the sense in which it is "explicitly mental," the other emphasizing the sense in which it is not.

On the one hand, especially in *Intentionality*, Searle sometimes describes the Background as a collection of nonconscious propositions. However, given the fact that the Background is meant to *underlie* intentional states, it is clear that the Background cannot consist in representations, which are themselves intentional. This requirement, then, bars overly representational descriptions of the Background such as "assumptions," "presuppositions," or "attitudes"—Stroud writes that "to have any of those, it seems, is to be in an intentional state".[14]

Searle cannot employ the very terms that mandated the positing of the Background in the first place. Either Searle fails to avoid the "bad phenomenology" criticism, or the Background just is a kind of mental state *sans* the problematic phenomenological element, which he cannot exclude from mental states because of other aspects of his view. The Background is a placeholder for *whatever* imposes and maintains a status-function on some brute fact in a way that avoids Dreyfus' criticism; in which case it is mysterious.[15]

On the other hand, Searle also describes the Background nonpropositionally, in terms of "abilities," "capacities," or "practices". This appears to be the dominant characterization in *The Construction of Social Reality*, where Searle casts the Background as "a certain category of neurophysiological causation" (*CSR*, p. 129). Searle reiterates this characterization in his more recent *Rationality in Action*, where he implores the reader to think of the Background "ontologically speaking as a set of brain structures".[16]

However, if the Background is nonrepresentational, a category of neurophysiological causation, it is not clear that Searle has answered his own question. Here, an institutional fact would consist of some brute fact (X′) and some other brute neurophysiological fact (X″). It is less than clear how the coupling of these two brute facts can produce the normative component indicative of the status-function. That is, if the Y terms are characterized in nonintentional X terms, then lost are the deontology that gives those concepts significance in the first place (and with it, the loss of the recognition that it is in virtue of intersubjective agreement that X was made to count as Y).

In the *Construction* there are places where Searle appears to grant that the Background cannot serve as a proxy for intentionality, if intentionality is what imbues some brute fact with the required normativity. Rather than appealing to rules in the form of conscious or unconscious representations, Searle sometimes contends that "we evolve a set of dispositions that are sensitive to the rule structure" (*CSR*, p. 145).

But what is this "rule structure"? If the Background is a set of "motivational dispositions," as Searle claims (*CSR*, pp. 5, 129, 135), then the "rule structure" cannot even be identified with the Background, for otherwise Searle would be telling us that the rule structure is causally sensitive to itself. Sometimes, Searle characterizes the rule structure as "intentional" (*CSR*, pp. 142, 145) but then Dreyfus' criticism still looms large. In that case the Background fails to serve as a suitable substitute for explicit intentional imposition, as promised by Searle.

Like the Background, this rule structure is either mysterious or unhelpful. And, in places, it seems to carry much of the load originally shouldered by the appeal to mental and linguistic representations, or even the Background. The rule structure is unhelpful in answering Searle's question as to how there can be institutions in a world that consists entirely of physical particles and conscious biological beasts (and rule structures) (*CSR*,

pp. xi–xii). Here, institutional facts are characterized in terms of some brute fact X and, if not explicit intentional imposition, a Background of dispositions on which a rule structure has impressed itself. Putting the possibility of Platonism aside, it is not that Searle has said something false, but only that it is difficult to understand how this is an answer to the very question that purports to drive the *Construction*.

I have argued that if Searle characterizes the Background as "explicitly mental," he either fails to avoid Dreyfus' criticism or else the characterization of the Background descends into unintelligibility. If the Background is characterized nonpropositionally, as a category of neurophysiological causation, it is not clear how *this* (unlike, perhaps, explicit intentional representations) could imbue the status-function with the required normative aspect. Moreover, appeals to a rule structure do not appear to help.

As evidenced by the *CSR*, p. 90 quotation, Searle thinks he is justified in claiming that institutional reality has the structure "X counts as Y in C" because the intentional representations constituting that reality appear to have that structure. But if it turns out, as Searle suggests, that huge blocks of our institutional reality do not appear to be fortified by such representations, then what can explain Searle's confidence that the formula outlines the structure of this reality? Indeed, what confidence can Searle have in suggesting that institutional reality has a logical structure at all?

Social reality does not have a logical structure: Searle's conceptual relativity

There is good textual evidence to suggest that Searle does not think that every constitutive rule contains mental or linguistic representations as constitutive elements. So why, in the heart of the *Construction*, is Searle motivated to declare that mental and linguistic representations are constitutive of our institutions? Is it

possible that Searle, here, is attempting to survey where we have been, without feeling it necessary to rehearse the details?

Rather than an ellipsis, I think the *CSR*, p. 90 formulation is a critical component of the argument at this point; so crucial, in fact, that Searle finds it necessary to paper over some of the earlier qualifications regarding the role of explicit intentionality. Even if Searle could meet the criticisms we have been outlining, the appeal to the Background undercuts his ability to answer one of the *Construction*'s principal questions: what is the logical structure of institutional reality? (*CSR*, pp. 2, 90)

References to a "logical structure" are not limited to the *CSR*, p. 90 quotation. "I describe the elementary construction of social facts and the logical structure of the development of institutional facts from simpler forms of social facts" (*CSR*, p. 31; see also pp. 22, 90, 112). Moreover, "the structure of institutional facts is the structure of hierarchies of the formula 'X counts as Y in context C'" (*CSR*, p. 55; see also p. 56). Searle also writes that "It is no exaggeration to say that these iterations [of the form 'X counts as Y in C'] provide the logical structure of complex societies" (*CSR*, p. 80).

But, unlike the rest of these citations, the *CSR*, p. 90 quotation functions as a kind of keystone for Searle's program: "How can an 'organized society' have a '*logical* structure'? After all, society is not a set of propositions or a theory, so what is this talk of a logical structure?" (*CSR*, p. 90)

Searle's answer to his own question, however, is unsatisfactory. Since the "Background" is defined as "the set of nonintentional or preintentional capacities that enable intentional states" (*CSR*, p. 129), or a category of neurophysiological causation, and only propositions, theories, and mental representations have a logical structure (*CSR*, p. 90), the Background *cannot* have a logical structure. Searle, in agreement with Dreyfus, is very clear that mental and especially *linguistic* representations do not underlie all institutional facts. But then even if our mental and linguistic

representations have the form "X counts as Y in C," this only accounts for the structure of those institutions which can be traced back to such representations. Even Searle, as we have seen, grants that institutions are not typically underlain by representations.

So what is the logical structure of these nonintentional institutional facts? Better, does it even make sense to talk about their logical structure if, as Searle contends, we can only predicate a logical structure of propositions or theories? The worry is not just over whether or not Searle has got the *right* account of social and institutional reality. The worry is that Searle, even by his own lights, may be quite out of bounds in suggesting that organized society has a canonical structure, much less in suggesting that the logical structure is "X counts as Y in C".

At this point, we are in a position to turn Searle's defense of conceptual relativism against the *Construction* itself. Following Putnam and Goodman, Searle agrees that conceptual relativism "show[s] that different conceptual systems will generate different and apparently inconsistent descriptions of the same 'reality'" (*CSR*, p. 163). I had argued that Searle's remarks, here, only applied to descriptions of *brute reality*. Institutional reality, however, *can* be said to have a logical structure because it is constituted in part by linguistic representations. However, in this chapter we have seen that, for Searle, it is not the case that institutional reality is always constituted by linguistic representations. This undermines the claim that institutional reality has a logical structure (only propositions, theories, and representations have these), and so dispels our confidence in a strong correspondence between institutional descriptions and their objects. If I am correct to compare Searle's project with that of the atomists, then Searle falls victim to a classic criticism of atomism: without the appeal to linguistic representations, Searle does not seem justified in suggesting that institutional reality has the structure "X counts as Y in C," or that it has any canonical structure whatsoever. These considerations move me to apply Searle's conceptual relativism to the case of institutional

reality itself, so that different conceptual systems will generate different and apparently inconsistent descriptions of the same *institutional reality*. The constitutive formula outlines one such conceptual system.

This presents another option for interpreting Searle. If Searle, with the constitutive formula, has articulated one such conceptual system, then it is unclear that he has answered the driving question about the logical structure of institutional reality. There is no single such *logical structure*, but only a motley of "different conceptual schemes," each of which simultaneously highlight and obscure various features of institutional reality. This suggestion echoes that of Max Weber, who observes that "the infinity and absolute irrationality of the multiplicity of which everything concrete consists provides an epistemologically really cogent *demonstration* that it is an absolutely senseless thought to attempt a 'copy' of reality through any kind of science".[17] Weber contends that because social reality consists of an "infinite multiplicity," we should not expect to apprehend it under any one model or ideal type, no matter how general. We will have occasion to go into much more detail about Weber's view in Chapter 4. Weber's characterization of reality, institutional or otherwise (recall that Weber was a social scientist), as an infinite multiplicity prefigures what Searle, Putnam, Goodman, and others later observed: while we should not deny that there is a reality which can be described, we should not confuse the structure of any one of those descriptions with the structure of that reality. With this in mind, recall again Austin. The atomist reading of the significance of the constitutive formula cannot be correct, because that would be to fall once again into the error of reading back into the institutional world the features of one of the ways we might talk about it.

The following chapters will attempt to flesh out a more palatable version of what Searle calls "conceptual relativism," and its significance regarding the constitutive formula. Searle's own articulation—"any system of representation at all

is conventional, and to that extent arbitrary"—is misleading. Calling models "systems of representation" understates the role they play in our activities; thus, the suggestion that these schemes are "arbitrary" or "purely and absolutely conventional"[18] will strike the reader as, at best, imprudent and, at worst, injurious. There are quite obviously better ways and models by which to understand the world, and it so happens that the constitutive formula is extremely useful to this end. Nevertheless, I will argue that it is not the *only* way to depict institutional reality.

Notes

1. See Sam Page, "Searle's Realism Deconstructed," *The Philosophical Forum* 35, no. 3 (2004).
2. George Pitcher, *Truth* (Englewood Cliffs, NJ: Prentice-Hall, 1964).
3. "Defend" does capture the sense in which Strawsonian disquotation is the more widely accepted conception of truth.
4. Pitcher, *Truth*, p. 24.
5. Ibid., p. 32.
6. Incidentally, it is precisely this point which is affirmed in Searle's attempt to bridge the is-ought gap. See John R. Searle, "How to Derive 'Ought' from 'Is'," in *Concepts in Social & Political Philosophy*, ed. Richard E. Flathman (New York: Macmillan, 1973).
7. Searle, *Consciousness and Language*, p. 137.
8. Hubert L. Dreyfus, "Phenomenological Description Versus Rational Reconstruction," *Revue Internationale de Philosophie* 55, no. 1 (2001): p. 187.
9. Ibid., p. 188.
10. John R. Searle, "Neither Phenomenological Description nor Rational Reconstruction: Reply to Dreyfus," *Revue Internationale de Philosophie* 55, no. 1 (2001): p. 279.
11. Ibid., italics his.
12. John R. Searle, *Intentionality* (Cambridge: Cambridge University Press, 1983), p. 152.
13. Ibid., p. 151.
14. Barry Stroud, "The Background of Thought," in *John Searle and His Critics*, eds E. LePore and R. van Gulick (Mass: Basil Blackwell, 1991), p. 250.

15. If the Background is meant to explain intentional states, as is argued in *Intentionality*, Stroud makes a parallel argument when he complains that the Background just is whatever explains intentionality in the required way. See ibid., p. 253.
16. John R. Searle, *Rationality in Action* (Massachusetts: The MIT Press, 2001), p. 58.
17. Weber in Thomas Burger, *Max Weber's Theory of Concept Formation* (New York: Duke University Press, 1987), p. 66.
18. Austin in Pitcher, *Truth*.

4

Kuhn, Weber, and Instruments of Inquiry

The criticisms of Chapters 2 and 3 against Searle's institutional atomism motivate a different reading of the constitutive formula. While we shall continue to take the constitutive formula as a means by which we can answer questions about institutional reality, I think we need to scale back our expectations of what we should expect from such answers. Moreover, we need to say something about how articulation is possible, and how we might select among competing systems of representation, in the face of Weber's "infinite multiplicity". This will be the task of the remaining three chapters.

In Chapter 4 I will compare Thomas Kuhn's depiction of research in the physical sciences with Max Weber's discussion of research in the social sciences; it is not clear to me that the congruence between these two authors has been fully appreciated. I will argue that the role that a paradigm plays in Kuhn's depiction of scientific activity approximates the role that an ideal type plays in Weber's earlier account of social scientific activity. As a pedagogical tool, Kuhn's notion of a paradigm provides a good first approximation of Weber's central methodological concept, the ideal type. Both of these authors locate scientific and social scientific statements in the context of an ongoing, dialogical process. Assertions are answers to specific questions and get their significance only against that process. The extent that an answer allows an inquirer to overcome a pressing difficulty (anomalies, counterinstances, etc.) is the extent to which the answer is considered successful. Although the analogy between a paradigm and an

ideal type eventually breaks down, there are important similarities between the two modes of inquiry.

This chapter provides the groundwork for determining what we can and cannot expect of our answer to questions about institutional reality. In Chapters 5 and 6 I will apply Kuhn's and Weber's insights to the particular case of institutional analysis. This will be done by looking at what we can and cannot expect from the constitutive formula when applied to the example of money. Following this discussion I then endeavor to discuss the significance of, not just applications of the constitutive formula (i.e. money), but the formula itself.

The generation of puzzles in the natural and social sciences

In this section I argue that both Kuhn and Weber construe the natural and social sciences as puzzle-solving activities. Activities proceed according to certain recognizable patterns. I show that Kuhn and Weber have identified similar such patterns in these two modes of inquiry. In particular I will argue that Kuhn's paradigm and Weber's ideal type are instruments or tools in that they help the inquirer identify a potential puzzle.

Kuhn: paradigms are instruments that generate puzzles in the natural sciences

The textbook account of science is the view "that the content of science is uniquely exemplified by the observations, laws, and theories described in [the textbook's] pages".[1] Because the textbook account moves us to misunderstand the actual processes of discovery, Kuhn suggests that we rather take a motley of problem-solving activities, techniques, and procedures to be

indicative of science. According to the textbook, science is distinguished according to its subject matter, as opposed to a particular kind of activity—"patterns of education and communication".[2] The textbook account moves us to (mis-)read the nature of scientific activity back through an idealized portrait of our observations, laws, and theories. Kuhn, on the other hand, begins with an account of the activity and suggests that this might have some impact on how we might understand the significance of those observations, laws, and theories.

The textbook account is a useful picture for some purposes. For example, it serves pedagogic and persuasive functions.[3] It also rightly emphasizes the evolution our scientific theories have undergone.[4] But in other ways the textbook account of science is misleading. If science is understood as a body of truth, a certain picture of the nature of scientific activity is implied—"the piece meal process by which these items have been added, singly and in combination, to the ever growing stockpile that constitutes scientific technique and knowledge".[5] A history of science that takes for granted this "process of accretion," would then chronicle when and by whom each item in that body of truth was discovered. To explain failures of discovery, such an account would appeal to "error, myth, and superstition".[6]

However, this conjectural history of science, found in the textbook's pages, is undermined by investigations into the actual process of scientific development. Questions such as when and who discovered oxygen are elusive. Historians have difficulty distinguishing between the scientific and occult components of a scientist's worldview; idiosyncratic elements seem to be a dominant and even necessary feature of a scientist's view. The emphasis placed on science's ability to banish ignorance obscures the fact that past scientific activities are importantly similar to contemporary scientific activity. Within normal science, what contemporary scientists share with their ancestors is the ubiquitous adherence to a given paradigm or governing exemplar. The problem with the textbook

account of science, that emphasizes the *objects* of discovery over the *activity* of discovery, is that when the process or activity is read back through this picture it becomes almost impossible to account for these similarities.[7]

What does it mean to say that science is a set of activities? Minimally, science—in all of its incarnations—is an activity that aims toward understanding or explanation.[8] Kuhn sometimes paraphrases "understanding" as "the ability to recognize a given situation as like some and unlike others that one has seen before".[9] As we shall see, what guides this activity, for Kuhn, is a set of values and commitments embodied in received paradigms. Science might be called, following Weber[10] and Foucault,[11] a "discursive activity". This is to distinguish it from other kinds of activities that aim primarily, not at understanding, but at some other purpose. For instance, the activity of Wittgenstein's *Philosophical Investigations* §2 is not discursive because its primary goal is to construct buildings, not to develop a systematic and sufficiently general depiction of the world.[12] In this activity words like "block" and "slab" are tools or instruments that serve the purpose of building.[13] Just as a chisel or measuring tape might have aided in the shaping of a slab, "slab" aids in the movement and placement of that slab. To push the analogy, I suggest that paradigms are likewise tools or instruments that facilitate the aims of the scientist.[14] That is, paradigms help scientific inquirers increase their understanding of the natural world; they do this through a succession of puzzle-solving exercises. An example of a discursive activity might be the distinguishing and labeling of waterfowl, in Kuhn's "Second thoughts on Paradigms".

Why characterize such activities as discursive? Understanding is achieved by means of puzzle-solving: a puzzle is proposed and then solved. This increases our understanding, and makes possible new kinds of puzzles. In this way the process resembles a kind of conversation. The dialogical quality becomes even more obvious if we exchange the locution of "puzzle-solving" with that of

"question-answering": with a light enough touch, this discursive process might even be described as Socratic.

In the *Structure of Scientific Revolutions*, Kuhn writes that a paradigm "identifies challenging puzzles, supplies clues to their solution, and guarantees that the truly clever practitioner will succeed".[15] Following Kuhn, I will argue that paradigms are tools in *two* senses: first, they enable the scientist to see a puzzle in the first place. Second, they provide part of the means for the puzzle's resolution.

I should begin by saying a few words about what Kuhn has in mind by "puzzle". A puzzle, unlike other kinds of problems, provides the means to its own solution. Crossword puzzles provide hints in the form of written clues, usually a simple definition, supplemented by constraints placed on word length. A series of other conventions further constrain the possibility field: clues formulated in the past tense yield answers in the same tense. Clues that end in a question mark flag a pun. Tangrams are puzzles which are resolved when seven predefined polygons are arranged to form a given shape. These and other puzzles such as word riddles, anagrams, Sudoku, and the Rubik's Cube likewise contrast with other sorts of problems—such as how to reduce greenhouse emissions from internal combustion engines—in which the formulation does not provide clues to its resolution. The important thing to see is that part of what makes puzzles tractable, unlike other problems, is that they contain part of the means to their own solution—they are internally soluble. Naturally, there is no crisp dividing line between puzzles and other kinds of problems.

Science proceeds through solving puzzles, not problems. Paradigms are instruments in that they provide the scientist with the means, first, to see something as puzzling and, second, to help resolve that puzzle. I will begin by saying a few words about how a paradigm helps the inquirer see a phenomenon as puzzling.

It should be noted that Kuhn only rarely explicitly characterizes a paradigm as a kind of tool or instrument. First, he writes

that "So long as the tools a paradigm supplies continue to prove capable of solving the problems it defines, science moves fastest and penetrates most deeply through confident employment of these tools".[16] Second, "The law-sketch, say $f = ma$, has functioned as a tool, informing the student what similarities to look for, signaling the gestalt in which the situation is to be seen. The resultant ability [is] to see a variety of situations as like each other ...".[17,18] A paradigm is a tool in the sense that it helps the scientist "define," "state," or "identify" the problem to be solved. I will talk below about the sense in which a paradigm is a tool that *solves* a problem.

A paradigm is normative in that it helps the researcher see what counts as a (legitimate) problem in the natural sciences: "One of the things a scientific community acquires with a paradigm is a criterion for choosing problems that, while the paradigm is taken for granted, can be assumed to have solutions".[19] If the problem "cannot be stated in terms of the conceptual and instrumental tools the paradigm supplies" then the problem is "rejected as metaphysical, outside the scope of the discipline, or just too problematic".[20]

What sort of thing might come to puzzle a scientist? Kuhn gives us an important clue when he writes that an "anomaly appears only against the background provided by the paradigm".[21] If anomalies and counterinstances are deviations from a set of expectations, those expectations are embodied in paradigms; it is only against the history of a paradigm's success in accounting for a class of phenomena that the notions of anomaly, counterinstance, and crisis can gain currency.

To see this it may be helpful to review what has become a kind of scientific urban legend, wherein science has been said to have demonstrated that bumblebees cannot fly.

A bird or airplane wing, for example, is a model of the airfoil paradigm and Bernoulli's principle, whereby an increase in the speed of a fluid is inversely proportional to the pressure within

that fluid. The airfoils on birds and airplane wings force the air to move faster above the wing, creating a lower pressure and producing lift. It need not always have been the case that the airfoil was the dominant paradigm—at one point bird flight was unaccounted for by any model or, in the case of pre-paradigm science, perhaps this phenomenon was covered by too many models.

At this stage the airfoil paradigm is more or less effortlessly extended to cover a variety of seemingly similar phenomena. So that the airfoil is used to account for, not just avian flight, but the chiropteran and pterosaurian flight; the paradigm is easily extended to cover fish locomotion. While the Wright brothers used the paradigm to develop the *Flyer*, the cambered airfoil was originally developed to explain the behavior of kites. The paradigm has also been employed to cover the behavior of curve balls, tectonic uplift, and the tendency for passing cars to be drawn into large trucks.

But insect flight looks anomalous against objects such as birds and airplanes, which exemplify the airfoil paradigm. The claim that bumblebees should not be able to fly has its origins in a 1934 book by entomologist Antoine Magnan, who explained that if bee wings functioned like airplane wings, they could not produce nearly enough lift to suspend the bee in flight. Moreover, the shallow angle of attack a bee wing exhibits would appear to induce a *stall*, which, according to Bernoulli's principle, happens when a greater pressure is created above the wing.[22] If Bernoulli's principle, along with a shape that induces the required differences in air pressure, is the paradigm, exemplified in things like bird and aircraft wings, it is also the contrast class against which insect flight appears puzzling.

The bee wing myth illustrates the first sense in which a paradigm is a tool: it creates the possibility of a puzzle. The fact that insects fly is normally not surprising or interesting except against the background supplied by the airfoil paradigm. More importantly, the anomaly must bear some resemblance to the paradigm if the

difference is to be surprising, puzzling, or even comprehensible. The paradigm-as-contrast-class here could probably never be orbital mechanics, the set of all prime numbers, or Afghanis. Insect flight may be different from these things, but it is so different that it does not even rise to the level of counting as an anomaly.

Even if a paradigm fails to anticipate a feature of a given phenomenon (and so produce crisis), a paradigm *must* first illuminate a number of other features of that same phenomenon; a paradigm must have helped the research scientist solve much of a given puzzle before crisis is even possible. In this way, "In science ... novelty emerges only with difficulty, manifest by resistance, against a background provided by expectation".[23]

Weber: ideal types are instruments that generate puzzles in the social sciences

I will argue that, according to Weber, the ideal type plays something of the same role in the social sciences that a paradigm plays in the physical sciences. In particular, I suggest that both authors articulate their respective discursive activities in terms of puzzle-solving. In this section, I will focus on the sense in which paradigms and ideal types are tools that give rise to puzzles. In this way, I use Kuhn to help clarify Weber's ideal type as spelled out in Weber's classic and difficult text, " 'Objectivity' in Social Science and Social Policy".

Weber, like Kuhn, is less concerned with the *Sache* of social science, than with the process or activity that brings it about. And this process is one of problem-solving: "It is not the 'actual' interconnections of 'things' but the *conceptual* interconnections of *problems* which define the scope of the various sciences".[24] To see this, I will review Weber's critique of Karl Marx, who adopts what might be called a textbook account of social science.

Weber's critique of Marx's periodizing dialectic illustrates a stance that bears a close affinity to that of Kuhn's. Weber

does not challenge Marx on empirical grounds, but on a metatheoretical level. Weber does not indict him for being *wrong* or *false,* but for overstating the significance of his historical account. Weber argues that Marx has given us, not a hypothesis about economic "laws," but an ideal type.[25] Characterizing Marxian concepts in terms of ideal types, rather than laws, implicates their *heuristic* significance—the role they play in a discursive activity.

Weber's criticism of Marx is subtle, displaying sincere admiration. In recharacterizing Marx's historical materialism as an ideal type, Weber simultaneously underscores its importance, while avoiding the excesses of Marx and his adherents.

> We will only point out here that naturally all specifically Marxian "laws" and developmental constructs—insofar as they are theoretically sound—are ideal types. The eminent, indeed unique, *heuristic* significance of these ideal types when they are used for the *assessment* of reality is known to everyone who has ever employed Marxian concepts and hypotheses. Similarly, their perniciousness, as soon as they are thought of as empirically valid or real (*i.e.*, truly metaphysical) "effective forces," "tendencies," etc. is likewise known to those who have used them.[26]

Weber claims that the "so-called 'materialistic conception of history' as a *Weltanschauung* or as a formula for the causal explanation of historical reality is to be rejected most emphatically".[27] But the fact that Marx overstated the significance of his own claims should not move us to deny the "eminent, indeed unique, *heuristic* significance of these ideal types".

Weber, like Kuhn, responds by drawing our attention to the role these concepts play in the *activity* of social science itself. Marx, Weber argues, has rather formulated an ideal type—one that is bound to run against anomalies and counterinstances. As a puzzle is raised, crisis ensues. In this way, ideal types are useful; they are tools in the sense that they provoke crisis. Like paradigms,

ideal types, then, are a tool or an instrument in the sense that it gives rise to puzzles. I have suggested, and will argue in detail below, that this is the first step in a discursive process, where the resolution of these puzzles makes possible new sorts of puzzles, thereby enhancing our understanding. Consider Weber's articulation of the same process:

> The history of the social sciences is and remains a continuous process passing from the attempt to order reality analytically through the construction of concepts, the dissolution of the analytical constructs so constructed through the expansion and shift of the scientific horizon, and the reformulation anew of concepts on the foundations thus transformed.... [T]his process shows that in the cultural sciences, concept-construction depends on the setting of the problem, and the latter varies with the content of culture itself. The relationship between concept and reality in the cultural sciences involves the transitoriness of all such syntheses.[28]

By "concept," here, Weber means "ideal type". Ideal types are constructed and face dissolution in the face of "the expansion and shift of the scientific horizon". This empirical expansion is one of the sources of anomalies and counterinstances that engender "the setting of the problem," or crisis. As we will see when I discuss various responses to crisis, below, this in turn eventuates in the "reformulation" of new concepts and discursive foundations. "The history of the social sciences is and remains a continuous process," which I have characterized as discursive.

In particular, Weber was interested in demonstrating the "transitoriness" of Marx's materialistic conception of history. Marx loses sight of the role his concepts play in puzzle-generation, and social scientific activity in general, when he characterizes historical materialism as a body of truths, hypotheses about "truly metaphysical" laws or tendencies. This evokes the textbook's

characterization of scientific phenomena. Marx's mistake, according to Weber, was not that he said something wrong or even unhelpful, but that he did not understand the significance of his own insight. In this way, Weber treats Marx with a remarkably light touch.

Looking beyond historical materialism, less spectacular social scientific ideal types can help inquirers formulate puzzles. To this end, it will be helpful to look closely at the ideally typical economic agent, *homo economicus*. Such an agent is instrumentally rational; an action is rational just in case it maximizes expected utility, or brings about the object of desire, usually wealth. Minimally, agents are assumed to display a strong preference for their own interests over the desires of others. This picture might be called the rational choice, neoclassical, or Bayesian model of economic behavior.

Economic behavior that traces overt supply and demand curves is easily accounted for in this theory. Buyers and sellers who have access to a broad amount of information regarding the asking prices of other merchants, and where transactions are largely anonymous, are most easily brought under the neoclassicalist model of economic behavior. The behavior of those in a medieval marketplace or an Internet auction-house is probably exemplary in this respect.

While the neoclassical ideal type has been primarily employed to explain economic behavior, the model is easily extended to cover a broader array of cases. The variety of such patterns explained by an explicit appeal to the economic rational choice model is astounding: researchers have used the model to explain (and hopefully predict) target choice by residential burglars by taking into account major arterial routes, house attributes such as visibility and affluence, occupancy indicators (a car in the driveway, a large number of children), and security precautions in use.[29] A staggering amount of research has been dedicated to "competitive marriage market models," which use the neoclassicalist ideal type as a basis for explaining marriage formation and

dissolution.[30] Finally, in the most extreme deployment of the model, some authors have endeavored to explain the swarming behavior of honeybees by appealing to the Bayesian ideal type: "We assume that a honeybee is rational in choosing to stay in the old hive or to leave to build a new one. Rationality here refers to a bee's behavior to maximize the food (honey) that it can share or contribute".[31]

In spite of its explanatory scope, some behavior appears mysterious against the Bayesian model. Thus, the ideal type gives rise to a puzzle. For example, an economist might be puzzled at the persistence of "inshopping" in rural communities, where members continue to shop at more expensive locally owned retail stores, instead of discount chain stores. The fact that it is not uncommon for sellers to place maximum-bidding prices on Internet auctions is puzzling to the neoclassicalist. Such researchers have even struggled to explain a practice so ubiquitous as restaurant tipping.

Non-economic behavior can also be puzzling against what the neoclassical ideal type would have us expect. Authors have pointed out that addiction, suicide, and even religious belief seem wholly unmotivated by Bayesian lights.

Again, just as in the natural sciences, paradigms embody a set of expectations that are occasionally dashed by anomalies and counterexamples; in the social sciences, ideal types also give rise to the possibility of a puzzle. The fact that people tip in restaurants or maintain religious beliefs is not normally surprising, except against the background provided by the neoclassical ideal type. But for adherents of the neoclassical model, tipping, inshopping, or suicide may be sources of puzzlement and crisis. The Bayesian model, or ideal type, is useful precisely *because* it is a condition for the possibility of puzzle-solving in the social sciences. They are, in this sense, tools or instruments of inquiry. These puzzles, which demand resolution, are indicative of a discursive activity that ultimately increases our understanding of these phenomena.

I have argued that both paradigms and ideal types help generate puzzles in the natural and social sciences, respectively. They embody pictures against which applications could come to count as anomalous or puzzling. Confronted with puzzles, both Kuhn and Weber recognize that inquirers have a variety of strategies at their disposal. If both paradigms and ideal types help generate puzzles, we might likewise also compare how they might be used to solve puzzles in these discursive activities. Before doing so, I should say something about the etiology of a paradigm or ideal type.

The etiology of a paradigm or ideal type

If both Kuhn and Weber agree that paradigms and ideal types are conditions for the possibility of a puzzle arising, I think there is also substantial agreement as to the etiology of a paradigm or ideal type—the conditions which lead to and justify their formation.

Starting with Weber, I would now like to draw the reader's attention to those passages where Weber talks about the genesis of a given ideal type. In the human sciences "the guiding 'point of view' [or evaluative idea] is of great importance for the *construction* of the conceptual scheme [or ideal type] which will be used in the investigation".[32] The researcher's points of view or evaluative ideas play a formative role in the *creation* of an ideal type. In this way, Weber locates the formation of a given ideal type in the antecedent commitments and values of the researcher.

Evaluative ideas, for Weber, are antecedent commitments or beliefs that help shape how a researcher interprets the data. That is to say, ideal types build from, and are grounded in, a community's understanding. "Genuine artistry" in the human sciences, Weber says, "manifests itself though its ability to produce new knowledge by interpreting already *known* facts according to

known viewpoints".[33] Weber's notion of an "evaluative idea" points to conspicuous aspects of this general understanding—aspects which pay a role in differentiating one resulting ideal type in the social sciences from another. As we have seen, an evaluative idea which underlies Marxism is that modes of production seek to overcome their own limitations.

Larry Wright, in *Teleological Explanations*, writes that "Expressions, or sets of expressions which have originated and developed their logical nuance in one context and by reference to a limited subject matter, will sometimes be metaphorically extended to other contexts and referents. If done with care and logical sensitivity, this process is unobjectionable and frequently fertile of insight".[34] An ideal type might be compared to a metaphor in that it packages antecedent understanding and patterns (evaluative ideas) in such a way that it can be harnessed to illuminate novel phenomena.

While evaluative ideas may be distinguished from ideal types (ideal types are "formed" by them), the difference is a matter of degree. Ideal types are just evaluative ideas or antecedent commitments which have been made *more* public and articulate (since evaluative ideas are already given to our understanding, there is a sense in which they must be already public). An ideal type's publicity is the basis for the possibility of a normal or quasi-normal discursive activity. Weber emphasizes an ideal type's potential to be more articulate than the evaluative idea which underlies its formulation. A researcher who employs ideal types, Weber believes, is less subject to build hypotheses on the basis of hidden ambiguities: one consequence of not formulating explicit ideal types is that the researcher "consciously or unconsciously uses other similar concepts without formulating them verbally and elaborating them logically or that he remains stuck in the realm of the vaguely 'felt' ".[35]

It is misleading to talk of "the" evaluative idea which led to a formulation of an ideal type, as if it were a discrete entity. While

ideal types emphasize features of our general understanding (evaluative ideas), it is a mistake to think that those features can be isolated and separated from the entirety.

Weber grounds an ideal type in our antecedent understanding, or more particularly, in conspicuous evaluative ideas about how the world works. Similarly Kuhn grounds the formation or construction of a paradigm in our antecedent understanding, so that, in addition to satisfying Kuhn's five criteria of theory selection, a theory must be intuitively compelling. Of course, especially in the natural sciences, the right kind of training[36] is often required in order to be a competent judge of what is to count as intuitively compelling.

In talking about the formation or construction of a paradigm, I have in mind the transition from pre-paradigm science or revolutionary science to normal science.[37,38] In both cases, even as a new paradigm becomes formulated, it seeks to make reference to the understanding acquired by previous generations. In the transition from pre-paradigm science to normal science, Kuhn writes that "no natural history can be interpreted in the absence of at least some implicit body of intertwined theoretical and methodological belief that permits selection, evaluation, and criticism. If that body of belief is not already implicit in the collection of facts—in which case more than 'mere facts' are at hand—it must be externally supplied, perhaps by a current metaphysic, by another science, or by personal and historical accident".[39] In the transition from revolutionary science to normal science, Kuhn in "Logic of Discovery" notes that "a man trained as puzzle solver will wish to preserve as many as possible of the prior puzzle solutions obtained by his group".

This is seen in Kuhn's remarks about the formulation of a paradigm: in the movement from pre-paradigm science to normal science, electricity might be characterized as a "fluid" which ran through conductive material by one school, and "effluvium" which emanated from non-conductors by another school.[40] These

are both intuitive pictures, which ground an understanding of electricity in a pre-existing understanding of how the world works. Fluidity became the dominant paradigm because of its ability to account for data more accurately, with greater scope and simplicity. Here, the prospective paradigm is able to illuminate a new phenomenon (electricity) by evoking antecedent understanding (the behavior of fluids). This intuitive appeal does not just in fact lead to discovery of a paradigm, but is part of the justificatory force of that paradigm. Naturally this antecedent understanding need not be "general" or "common" understanding, for otherwise it would be difficult to account for the formation of modern physics or even the heliocentric view, but it must exist if only among an educated elite.

First possible response to a puzzle

Both the physical sciences and the social sciences are discursive activities in the sense that both aim to increase our understanding. Resemblances between disparate phenomena are sought by means of a set of paradigms or ideal types. As the paradigm or ideal type is applied to increasingly disparate phenomena, puzzles emerge in the form of anomalies and counterinstances. Paradigms are tools that eventuate in an increase of understanding by means of puzzles. I have argued that Weber shares a similar conception of social scientific activity.

How can researchers respond to such puzzles? Both Kuhn and Weber outline a series of overlapping but non-identical responses to a puzzle. The difference in Kuhn's and Weber's views, here, points to significant differences in the natural and social sciences.

Kuhn outlines three responses to a crisis.[41] First, the dominant paradigm may be able to absorb, by means of further articulation or extension, the anomaly. This response is part of the "cumulative process" by which normal science proceeds. Second, the

puzzle can be bracketed for a future generation of problem solvers. Finally, there is the possibility of revolution, wherein the dominant paradigm can be supplanted by a new paradigm.

Weber proposes two typical responses to crisis.[42] First, a researcher may stretch or modify the paradigm or ideal type so as to absorb the anomaly. The researcher assumes that the paradigm or ideal type provides a solution-sketch to the puzzle, even if the sketch must be appropriately modified. Understanding, here, is achieved by means of a fairly linear trajectory or, for Kuhn, a "cumulative process": the solving of one puzzle allows for new puzzles to arise that are in turn solved, and so on. Ideal types, like paradigms, are tools or instruments that drive this process. The process proceeds as puzzles that arise naturally over the course of inquiry are resolved.

Weber's first response, then, exactly corresponds to Kuhn's first response to crisis, wherein the dominant paradigm is extended to cover the anomalous case.

In Weber's second response to crisis, a researcher may take the anomaly to be an indication of the *limits* of the scope of an ideal type. Weber argues that, while ideal types can be stretched or modified, they *must* break down eventually. This may also be true of paradigms but I will not argue as much here. Here, crisis is an indication that a *different* paradigm or ideal type might help us better explain the offending phenomenon.

Weber's second response to a puzzle offers an alternative to the claim that understanding is achieved by means of a linear process, wherein a single ideal type is continuously modified or stretched in the face of anomalies. Rather, a number of different ideal types might be used *simultaneously* to understand a given phenomenon. The possibility of this response softens the perhaps caustic suggestion that discursive activities always proceed along a linear plot curve; paradigm-modification is only one possible reaction to anomaly.[43] Sometimes a new ideal type is used *in addition* to the old ideal type; the old ideal type is not abandoned.

Weber's second response is related to, but not identical to, Kuhn's second response to crisis. I will also discuss why Kuhn's third response to crisis does not appear to be a genuine possibility in the social sciences.

Inquiry may turn up applications that resist being subsumed under the dominant paradigm or, according to my comparison, an ideal type. When this happens, the inquirer's expectations are dashed and the offending phenomenon becomes puzzling.

Kuhn focuses on one response to the discovery of new and defiant phenomena. "Novelties of fact" are routinely instrumental in refining the dominant paradigm:

> Discovery commences with the awareness of anomaly, i.e., with the recognition that nature has somehow violated the paradigm-induced expectations that govern normal science. It then continues with a more or less extended exploration of the area of anomaly. And it closes only when the paradigm theory has been adjusted so that the anomalous has become expected.[44]

In attempting to resolve a question or puzzle, the ideal type gets adjusted, refined, expanded, qualified, or rearticulated. Researchers "will devise numerous articulations and *ad hoc* modifications of their theory in order to eliminate any apparent conflict".[45] Crisis provides "an occasion for retooling" an entrenched paradigm.[46]

Weber writes about ideal types, mirroring Kuhn's discussion of the paradigm-modification:

> Wherever the strictly economic explanation encounters difficulties, various devices are available for maintaining its general validity as the decisive causal factor. Sometimes every historical event which is *not* explicable by the invocation of economic motives is regarded *for that very reason* as a scientifically insignificant "accident." At other times, the definition of "economic" is stretched beyond recognition so that all human interests which

> are related in any way whatsoever to the use of material interests are included in the definition.[47]

The puzzle-solving aspect of the discursive activity involves reconciling offending data. Just as Kuhn describes scientific activity in terms of puzzle-solving, Weber draws the reader's attention to the social scientist's task of overcoming "difficulties" or "problems".[48] The two problem-solving tactics Weber mentions here include treating the data as "accidental" or isolated, and "stretching" the ideal type ("economic," "rational," "distribution"). Weber sometimes characterizes this "stretching" in terms of the "reconstruction" or "reformulation" of an ideal type.[49]

Compare Weber's remarks to the kind of maneuvers made by researchers employing the phlogiston paradigm in the face of anomaly. While it had been observed that burned metal *gains* weight, providing prima facie evidence against the paradigm that moved the researcher to expect that burned metal would *release* something into the atmosphere, there was not a call to revise the paradigm:

> In the seventeenth century [the call to repeal the phlogiston paradigm] seemed unnecessary to most chemists. If chemical reactions could alter the volume, color, and texture of the ingredients, why should they not alter weight as well? Weight was not always taken to be the measure of quantity of matter. Besides, weight-gain on roasting remained an isolated phenomenon. Most natural bodies (e.g., wood) lose weight on roasting as the phlogiston theory was later to say they should.[50]

That metal gained rather than lost weight seems to challenge the phlogiston paradigm. However the researcher in question has recourse to some of the same puzzle-solving strategies mentioned by Weber: the event could be treated as isolated, or the significance of the term "weight" can get recharacterized, modified or,

in Weber's words, "stretched". It could be treated as importantly similar to volume, color, and texture or the possibility of negative weight could be postulated.[51] We will talk about the isolation response in the next section.

Examples of paradigm modification are relatively easy to come by in the natural sciences. Recall the case where researchers tried to understand insect flight by means of the airfoil model. Researchers could not explain insect flight because, given the speed at which an insect flies, the wing was not large enough to produce lift. Moreover, the paradigm suggested that the low angle of attack would result in a stall.

The fact that insect wings move (unlike an airplane) does not, in fact, render the airfoil paradigm inapplicable. Bird wings move and are effectively treated under the airfoil paradigm. Rather, the shallow angle of attack becomes an important hint in reconciling the paradigm with the facts. Stall is not instantaneous—just prior to a wing's losing lift by means of an improper angle of attack, the velocity of the air above the wing actually accelerates, producing a momentary and dramatic increase in lift. Because airplane and bird wings are relatively static, they cannot take advantage of this phenomenon. However, the high rate at which insect wings flap creates a kind of loop that exploits the phenomenon of delayed stall. It is important to see that before an insect wing can stall, it must first behave much like a standard airfoil. As the angle of attack is changed, a stall is induced which, following Bernoulli's principle, momentarily increases lift.[52] In this way, the very paradigm that gave rise to the anomaly (the failure of Bernoulli's principle to account for the lift of bee wings; the stall) was instrumental in solving the problem. What the investigation has shown is that, indeed, there is a certain sense in which insect wings *do* behave like airplane wings after all; the paradigm is not rejected, but refined and expanded to cover the puzzling case.

Note that the clarification of the airfoil paradigm to account for insect flight is just the first step in a series of clarifications, each of

which represent an advance in our understanding of the phenomenon. The expanded or stretched airfoil paradigm might be applied to explain or account for other insect-wing-like phenomena: hummingbirds and helicopters, perhaps. But new problems arise: while the delayed stall phenomenon can account for an insect's ability to fly, it cannot account for the fact that they can lift objects which are twice their weight. In response to this puzzle a researcher recently modified the paradigm to include the phenomena of wake capture and rotational circulation.[53] Still more questions remain: what of insects with four wings (lacewings are two pairs of wings that flap out of sync; tiger beetles have two pairs of wings, one of which remains stationary). How do insects turn?

Similarly, as inquirers are confronted with puzzles in the social sciences, one possible response, according to Weber, is to stretch or refine the very ideal type that gave rise to the puzzle in the first place. I have outlined several potential anomalies that have confronted various theorists deploying the Bayesian ideal type. These include restaurant tipping, community inshopping, and the fact that some sellers place bid caps on auctioned items. Insofar as the Bayesian model is extended to cover non-economic behavior, researchers have pointed out that addiction, suicide, and perhaps even religious belief appear puzzling. Inshopping, tipping, addiction, and suicide are *puzzles* for the economist and socio-economist. They represent prima facie challenges to the prevailing Bayesian ideal type. That is, in an interrogative mode, this behavior seems *puzzling* or *surprising* against the background of the ideal type. Why would a person tip if in so doing utility is unlikely to be maximized?

In all of these cases defenders of the neoclassical ideal type have offered solutions to these puzzles by stretching or modifying the ideal type. Indeed a great deal of research both in economics and socio-economics has been dedicated to the identification and resolution of precisely these sorts of anomalies or counterinstances.

Generally, refined and expanded analyses are required to explain seemingly anomalous economic behavior under the Bayesian

model. N.J. Miller, for example, endeavors to explain inshopping in rural communities, where residents continue to shop at more expensive, locally owned retail stores instead of at discount chain stores. He appeals to gains in "social capital" and "reciprocity" that outweigh the overt economic advantages offered by the larger stores.[54]

The practice of restaurant tipping might strike the neoclassicalist as irrational: "From a rational-choice perspective, tipping makes sense only if desired outcomes are contingent on how much is tipped. Three possibilities are that people tip in order to buy social approval, equitable relationships and/or future service".[55] Given the ideal type of the ideally rational agent, tipping might seem puzzling. Lynn and Grassman attempt to solve that puzzle by arguing that, indeed, it is *rational* to tip; they do this by, to use Weber's term, "stretching" the *ways* in which an agent can maximize "expected utility". *Wealth* is *one* of any number of aims that a rationally self-interested agent might pursue.

Even the case of the online auction site—generally taken to be an exemplary free market system where the Bayesian model best explains and even predicts behavior—displays anomalies or puzzles. Consider the abstract of "Buy Prices in Online Auctions: Irrationality on the Internet?":

> Buy prices are puzzling: it makes sense for a seller at auction to set a minimum bidding level (i.e., a reserve price), surely, but a maximum? We explore the use of maximum bidding levels (buy prices) in online auctions and provide a rational explanation for this seemingly irrational auction mechanism. We show that augmenting an English auction with a buy price can improve the seller's profits by partially insuring (and therefore increasing the expected payment from) some risk-averse bidders.[56]

These authors expand "expected utility" to include social capital, values, reciprocity, and security—perhaps in the form of future service.

Ideal types are always "in constant tension with the new knowledge which we can and *desire* to wrest from reality. The progress of cultural science occurs through this conflict".[57] I have argued that Weber's first response to a puzzle or conflict is importantly analogous to Kuhn's first response to crisis. By stretching the ideal type the "result is the perpetual reconstruction of those concepts through which we seek to comprehend reality".[58] Perpetual reconstruction traces a linear trajectory by which our understanding increases. Ideal types are proposed and applied in order to point out similarities between increasingly diverse sets of phenomena. At one point, an application may resist assimilation under a given ideal type, rebuffing the inquirer's expectations. As I have argued while the ideal type (contrast-class) must differ from the anomaly, in order for this difference to be significant, it must be set against a host of similarities. According to Weber's first response to crisis, the ideal type is stretched or modified so as to assimilate the errant phenomenon. With the revised and clarified model in hand, it is applied to additional cases until a new puzzle arises. One puzzle-solution is often a condition for the possibility of new types of puzzle.

As the paradigm is refined, this creates the possibility of new sorts of puzzles. This virtuous circle has the effect of increasing our understanding of science's subject matter. I have argued that Weber, likewise, recognizes a similar such circle—or spiral—in social scientific inquiry. In this way, by means of Weber's and Kuhn's first response to anomaly, our understanding increases by means of a fairly linear, dynamic process. In Weber's second response to anomaly, we will consider an alternative process by which we can increase our understanding of a given phenomenon.

Second possible response to a puzzle

Inquirers in crisis have recourse to another response besides that of model-modification. In particular, I will argue that sometimes an

anomaly is an indication of the scope or applicability of a given ideal type. While Kuhn has thus far proved helpful in my exposition of the ideal type, at this point I am going to appeal directly to Weber.

I will begin by looking at an articulation of the ideal type, which will strike many readers as baffling. Weber characterizes the ideal type as a kind of "*utopia*".[59] He goes on to say, only somewhat more helpfully, that "in its conceptual purity this thought-picture (*Gedankenbild*) cannot be found empirically anywhere in reality".[60] Ideal types provide a thought-picture or "utopia" against which a particular phenomenon may be compared. This is to say that it is not necessary that the particular phenomenon under investigation resembles the ideal type in every way. Indeed, it is Weber's contention that social reality is *always* more complex than any ideal type that might be used to investigate it, and so it *never* totally resembles the object of investigation: "All knowledge of cultural reality ... is always knowledge from *particular points of view*".[61] In general, an ideal type does not "copy," but "accentuates" and suppresses, or approximates various features of a given social phenomenon.[62]

Why is Weber committed to saying this? Weber follows the neo-Kantian Heinrich Rickert when he suggests that "the infinity and absolute irrationality of the multiplicity of which everything concrete consists provides an epistemologically really cogent *demonstration* that it is an absolutely senseless thought to attempt a 'copy' of reality through any kind of science".[63,64]

There are less caustically ontological ways to paraphrase "the infinity and absolute irrationality of the multiplicity of which everything concrete consists". In an intellectual climate more like our own, Weber might have put the insight this way: we only have access to this reality in virtue of a background understanding which itself is vast and, on the whole, inarticulate. While particular features of this background understanding can be articulated, there are limits to what we can expect from any one such

characterization. Given the inarticulateness of this background understanding, and given that this understanding necessarily structures our experience of empirical reality, Weber's temptation to characterize empirical reality itself as an "infinite multiplicity" is understandable.

Puzzles as an indication of the scope of a paradigm or ideal type

One response to conflict or crisis, when an ideal type challenges our expectations, is to "stretch," "reformulate," or "reconstruct" the very ideal type against which the phenomenon appeared anomalous. But, depending on the context, another response is available to the inquirer. In particular, sometimes puzzles can be so intractable that the researcher may take the anomaly as an indication of the limits of that ideal type's applicability. Weber writes: "The great attempts at theory-construction in our science were always useful for revealing the limits of the significance of those points of view which provided their foundations".[65]

Given this infinite multiplicity, we should expect that any one of the ideal types that aid in understanding should give a partial, finite rendering of this reality. Here, Weber draws our attention to the fact that we can always imagine asking additional questions about a phenomenon. No puzzle and its solution is canonical in the sense that we can not imagine additional puzzles we might have asked about the phenomena. Indeed, in many cases, not only can we *imagine* having additional puzzles, but these concerns might turn out to be, not just imaginable, but highly motivated.

Weber explains an ideal type's limitation by appeal to the infinitude of the object it is used to illuminate:

> ... the synthetic concepts used by historians are either imperfectly defined or, as soon as the elimination of ambiguity is sought for, the concept becomes an abstract ideal type and reveals itself as a theoretical and hence "one-sided" viewpoint

> which illuminates the aspect of reality with which it can be related. But these concepts are shown to be obviously inappropriate as schema into which reality could be completely *integrated.* For none of those system of ideas, which are absolutely indispensable in the understanding of those segments of reality which are meaningful at a particular moment, can exhaust its infinite richness.[66]

While ideal types are "absolutely indispensable," we should also be aware that even the most general and abstract models cannot "exhaust [reality's] infinite richness".

Because the ideal type is underdetermined by the object it is supposed to help explicate, we might compare the ideal type to a metaphor. Similar to a metaphor, because the means of comparison necessarily differs from the object of comparison, we have principled reasons to expect that any given ideal type must eventually break down in the face of this infinitude.

Sometimes Weber uses this language of comparison: ideal types "are of great value for research and of high systematic value for expository purposes when they are used as conceptual instruments for *comparison* with and the *measurement* of reality. They are indispensable for this purpose".[67] But because the object of comparison—social reality—is not only different, but infinitely more complex than the means of comparison, any given ideal type *must* eventually *break down.* That is, social reality is, on Weber's lights, ultimately inarticulable. This is not to say that we cannot talk about it: we should not, however, expect canonical articulations of this reality.

The second of Kuhn's three responses to crisis is that the puzzle can be isolated and left to the ingenuity of future generations of researchers: "scientists may conclude that no solution will be forthcoming in the present state of their field. The problem is labeled and set aside for a future generation with more developed tools".[68]

For Weber, an inquirer may respond to an anomaly by taking it to be an indication of a given ideal type's scope. While recognizing a model's scope of applicability is not obviously the same thing as bracketing it for future generations, the difference may be just a question of emphasis. In drawing attention to an ideal type's limitations Weber underscores its sphere of applicability. Kuhn, however, draws our attention to a paradigm's potential, and so downplays its historical successes. But these two formulations are compatible: when Weber suggests that an anomaly might be an indication of an ideal type's scope, he is not precluding the possibility of reevaluation at some later date. Moreover, in spite of the anomaly, the paradigm in question must be useful for some purposes if it is worth even bracketing.

Nevertheless, I will argue below that the difference in emphasis, here, anticipates substantive differences in how Kuhn and Weber characterize certain kinds of developments in the natural and social sciences.

On the use of multiple paradigms or ideal types

I want to look at one more passage where Weber underscores the partiality or incompleteness of any given ideal type: "Insofar as [an ideal type] traces a specific element of cultural life through the most diverse cultural contexts, it is making an historical interpretation from a specific point of view, and offering a partial picture, a *preliminary* contribution to a more complete historical knowledge of culture".[69] Weber thinks that all ideal types are necessarily "partial" and "preliminary"; in the face of an anomaly the social scientific inquirer is not always able to stretch an ideal type successfully. If so, in what sense is it possible to have "a more complete historical knowledge of culture"?

When a metaphor breaks down or else remains silent, one is always free to appeal to additional metaphors in order to draw out the sought characteristic. Consider Richard's attempt to

describe his dog to Rachel, the protagonist of Virginia Woolf's *The Voyage Out*:

> "We had a dog who was a bore and knew it," he said, addressing her in cool, easy tones. "He was a Skye terrier, one of those long chaps, with little feet poking out from their hair like—like caterpillars—no, like sofas, I should say".[70]

Richard is not denying that his Skye terrier is similar to a caterpillar, but is appealing to the sofa to draw out perspicuous features of the dog that the initial metaphor does not capture. Similarly, any number of ideal types may be used to increase our understanding of social phenomena. Even as the researcher is confronted with the task of saying something about that which resists articulation, be it a Skye terrier or social reality.

The triangulation of these various ideal types expands or rounds out a given portrait of this phenomenon. Along these lines, Weber discusses the diversity of ideal types or utopias that might be used to represent, for instance, the capitalist economic system:

> It is possible, or rather, it must be accepted as certain that numerous, indeed a very great many, utopias of this sort can be worked out, of which *none* is like another, and *none* of which can be observed in empirical reality as an actually existing economic system, but *each* of which however claims that it is a representation of the "idea" of capitalistic culture. *Each* of these can claim to be a representation of the "idea" of capitalistic culture to the extent that it has really taken certain traits, meaningful in their essential features, from the empirical reality of our culture and brought them together into a unified ideal-construct.... Inasmuch as the "points of view" from which they can become significant for us are very diverse, the most varied criteria can be applied to the selection of the traits which are to enter into the construction of an ideal-typical view of a particular culture.[71]

Any number of incommensurable ideal types, ideal-constructs, or utopias may be brought to bear on a phenomenon. Each ideal type can draw out interesting and significant features of this phenomenon, but Weber warns us that we should not expect any more from them.

Of course, Weber is primarily interested in reframing Marx's historical materialism. Above I suggested that part of the force of reconstituting Marx's concepts under the ideal type is that it draws our attention back to the activity of social science. But in doing so, a new face of Weber's critique of Marx comes to light: historical materialism is, perhaps, one of many utopias by which we might understand economic activity. To otherwise construe Marx's view in terms of "truly metaphysical" laws or forces is "pernicious" because it illicitly obscures this possibility. In this way, Weber criticizes Marx, not only for drawing our attention away from the role concepts play in a discursive process, but for not understanding the limitations of his own insight (or even for not understanding that insights can have limitations). Ideal types, for Weber, accentuate and suppress various features of a given phenomenon so that a number of different ideal types may be used simultaneously to both pose and answer questions about the object under investigation.

In looking for a more compelling illustration of Weber's view, I will return to the ideally typical rational agent in economics. For several years there has been a debate about how an economic agent is most usefully conceived. We have already discussed the neoclassical ideally rational agent. But according to the "behavioralist" view ("prospect theory" is a version of the behavioralist view, advanced by Daniel Kahneman), a variety of factors inform an agent's decisions. Behavioralists use findings in psychology to demonstrate that people's actual choices are less rational than the neoclassical model would suggest (given what the neoclassicalist means by "rational"). People, for example, tend to evaluate their own overall well-being against the perceived well-being of others.

People are more motivated by the fear of loss than the possibility of gain. People routinely overestimate their own competence in unfamiliar situations. According to prospect theory, an economic agent is conceived as an anthology of various dispositions.

The debate between the neoclassicalist and the behavioralist often gets cashed out as an empirical matter: which framework more accurately or precisely accounts for human behavior? An article in *The Economist*, for example, notes that behavioralist economics "suggests that people, contrary to the basic assumptions of the standard approach, do not always behave rationally". The article goes on to worry about the consequences if behavioralism turns out to be correct: "What if people cannot even calculate the amount they are willing to pay for a pound of butter or a hair cut, or have any idea what prices will be in the future? In such a world, such basic constructs as demand and supply curves—which show the quantities that people and firms would be willing to demand or supply at given prices—lose their meaning".[72]

But under Weber's ideal type conception, the neoclassicalist need not make any such claim: the neoclassicalist has merely formulated an ideal type, and has given us the means to depict how an agent, so conceived, might behave. This implies only that the neoclassical model is significant, useful or even necessary for economics—not the stronger claim that people "always behave rationally". Behavioralism does not disqualify demand and supply curves, as *The Economist* hyperbolically suggests: I can yet feel some confidence in the projections I have made about the price of butter. For Weber "an ideal type is formed by the one-sided *accentuation* of one or more points of view",[73] and the neoclassicalist view represents one such kind of accentuation.

Likewise, the behavioralist view groups together a variety of different formulae to form a different portrait of human behavior than that exhibited by *homo economicus*. More empirical research will not necessarily be decisive in adjudicating the differences between the two warring camps[74]—the features accentuated by

either the neoclassicalist or the behavioralist are *there*. The question is less about the facts, most of which are accounted for by either ideal type. The question is, rather, normative: how ought research in economics to proceed? Which ideal type is more fecund, less misleading? Moreover, if both ideal types look promising there is no reason not to pursue more than one line of inquiry.

This brings us back to the sense in which ideal types are tools or instruments. Since the 1940s the ideal type or model of the rational economic man has been a dominate element of economic research; to use such a model was, in part, what constituted legitimate research in the field. Eventually this ideal type started to show cracks as economists struggled, for example, to explain the stock market crash of October 1987 in terms of the efficient market hypothesis.[75] It was at this point when some economists—"especially the younger, more ambitious ones"—began to probe psychology and biology for a richer portrait of human motivation. Perhaps the debate between the neoclassicalist and the behavioralist is more usefully understood as a dispute over the kinds of ideal types that ought to provide the framework for explanation in economics.

Fortunately a number of attempts are underway to reconcile the neoclassical ideal type with the insights of the behavioralists. Along these lines, *The Economist* strikes a conciliatory tone:

> In fact, the battle between rationalists and behaviouralists may be largely in the past. Those who believe in *homo economicus* no longer routinely ignore his emotional and spiritual dimensions. Nor do behaviouralists any longer assume people are wholly irrational. Instead, most now view them as "quasi-rational": trying as hard as they can to be rational but making the same mistakes over and over.[76]

One way this reconciliation can take place is that the ideal type of the rational agent can be stretched, redefined, or delimited to carve a space for "irrational" elements.

However, a reconciliation can take place that does not necessarily involve the creation of a new super-ideal-type, a variation on Weber's first response which brings together insights of both the neoclassicalists and behavioralists under one roof. There is no reason the economist could not deploy a variety of tools, or else various camps of economists might each champion different ideal types.[77] Several ideal types can be brought to bear on a phenomenon to produce a more nuanced understanding of human behavior without our having to say how those ideal types relate to each other.

Kuhn, Weber, and the structure of revolution

Kuhn's first and second responses to crisis are, roughly, innovation and the bracketing of the paradigm in question. I have argued that these correspond with Weber's first and, to some extent, second responses to a discursive puzzle. In Kuhn's third response to crisis, the dominate paradigm is abandoned in favor of a new paradigm: "a crisis may end with the emergence of a new candidate for paradigm and with the ensuing battle over its acceptance".[78] In this way, social scientific inquiry can proceed linearly, when there is a dominant paradigm or ideal type, or conjunctively, when inquirers deploy multiple paradigms or ideal types to help account for the same phenomenon.

If there is a difference in Kuhn's and Weber's view, then, it must be found in the third response and Weber's advocacy of the deployment of multiple ideal types. Weber points out that in the face of such limitations, the researcher is always free to bring the phenomenon in question under multiple, alternative ideal types to increase our understanding. Comparing a Skye terrier to both a caterpillar and then a couch gives us a better understanding of the nature of the dog. Weber, however, does not articulate anything like revolution as a genuine or even possible response to crisis in the social sciences. Ideal types may be stretched or

accepted as limited, at which point understanding may continue to increase by means of different ideal types. A cursory reading of Kuhn, on the other hand, downplays the likelihood of this simultaneity while stressing the possibility of scientific revolution. To the extent that multiple paradigms coexist, it is in the case of *abnormal* science: preparadigmatic schools and brief periods of revolution exhibit this plurality.[79] "During the transition period there will be a large but never complete overlap between the problems that can be solved by the old and new paradigm".[80] *Normal* science is almost defined, then, in terms of the dominance of a single paradigm. Science, for Kuhn, is ideally monoparadigmatic. This also explains why, in the second response, Kuhn emphasizes the future potential of a failed paradigm; the luster of that potential is found in its *promise* to unify, to draw diverse phenomena under a single parasol.

This stands in stark contrast to Weber, who not only resists the idea that inquiry in the social sciences proceeds by means of a single, dominant ideal type, but gives principled reasons to think that this cannot be the case: every ideal type must eventually break down in the face of the infinite multitude that is social reality. In the second response, Weber celebrates an anomaly's ability to draw the researcher's attention to the limitations of a given ideal type; the inquirer has learned something about the tools he or she is using. Weber emphasizes an ideal type's past successes in accounting for a phenomenon, so as to emphasize that it is still in play despite new-found limitations in its scope.

Given the remarkable isomorphism that exists between Kuhn's and Weber's depiction of inquiry in the natural and social sciences, how should we understand this divergence? Why would Kuhn shun multiparadigmatic activity even as Weber appears to celebrate its possibility?

Kuhn provides the easiest answer when he characterizes the social sciences in terms of a collection of preparadigmatic schools: "it remains an open question what parts of the social sciences have

yet acquired such paradigms at all".[81] While both the natural and the social sciences are discursive activities, there is something different about their respective objects of discourse. Natural science seems unique in its ability to generate dominant, unifying paradigms; this is a condition for the possibility of revolutionary science in which a new paradigm *supplants* an old one. While inquiry in the social sciences resembles inquiry in the natural sciences in many respects, for various reasons, the similitude breaks down because the natural sciences appear able to produce dominant paradigms.

Against a background of similarity, we should expect differences between the discursive activities, and Kuhn's own characterization on this point is not unhelpful. Indeed, if at the end of our comparison, we did not find any differences between inquiry in the natural and the social sciences, *that* would have been surprising.

I will end this chapter with a brief, speculative note, if only to motivate the aforementioned comparison. Perhaps there is a sense in which Kuhn is overly committed to science's monoparadigmatic stance, thus overstating the differences between the natural and social sciences. Kuhn is perhaps correct to observe that scientists often, in fact, struggle to formulate even one paradigm, rendering Weber mute.[82] But Kuhn's characterization of a multiparadigmatic stance as "prescientific" or "abnormal science" might suggest a wider normative agenda. If so, perhaps Weber's treatment of the ideal type might help us temper Kuhn's depiction of crisis, and acceptable responses to it.

Notes

1. Thomas S. Kuhn, *Structure of Scientific Revolutions*, 3rd ed. (Chicago: University of Chicago Press, 1962), p. 1.
2. Thomas S. Kuhn, *The Essential Tension: Selected Studies in Scientific Tradition and Change* (Chicago: University of Chicago Press, 1977), p. xvi.
3. Kuhn, *Structure of Scientific Revolutions*, p. 1.

4. The suggestion that science is primarily an activity might lead one to caricature science as hopelessly contingent—as one of Kuhn's critics put it, "a matter of mob psychology". Kuhn argues, in "Objectivity, Value Judgment, and Theory Choice," that even though theory choice is often markedly idiosyncratic, it is not unobjective and baseless.
5. Kuhn, *Structure of Scientific Revolutions*, pp. 1–2.
6. Ibid., p. 2.
7. Ibid., pp. 2–3.
8. Kuhn, *The Essential Tension: Selected Studies in Scientific Tradition and Change*, 233, Bas C. van Fraassen, *The Scientific Image* (Oxford, New York: Clarendon Press; Oxford University Press, 1980), p. 154.
9. Kuhn, *Structure of Scientific Revolutions*, p. 192.
10. Max Weber, " 'Objectivity' in Social Science and Social Policy," in *The Methodology of the Social Sciences* (New York: The Free Press, 1949), p. 94.
11. Michel Foucault, *The Archaeology of Knowledge*, [1st American] ed. (New York: Pantheon Books, 1972), p. 46.
12. This is not to deny that there is a discursive element—it is necessary to systematically distinguish blocks from slabs. But part of Wittgenstein's purpose in choosing this example is to demonstrate that it would be absurd to suggest that the distinguishing between these objects is the primary purpose of the language-game.
13. Ludwig Wittgenstein, *Philosophical Investigations*, trans. G.E.M. Anscombe, 3rd ed. (New Jersey: Prentice-Hall, 1958).
14. The characterization of a paradigm as a tool was originally suggested to me by Mary Morgan's work on models as mediating instruments. See Mary Morgan and Margaret Morrison, "Models as Mediating Instruments," in *Models as Mediators*, eds Mary Morgan and Margaret Morrison (Cambridge: Cambridge University Press, 1999).
15. Kuhn, *Structure of Scientific Revolutions*, p. 179.
16. Kuhn, *Structure of Scientific Revolutions*, p. 76.
17. Ibid., p. 189.
18. Elsewhere Kuhn compares the scientist who, in the face of crisis, abandons a paradigm without finding a replacement to "the carpenter who blames his tools". (Ibid., pp. 79, 80, 84.) In "The Essential Tension" Kuhn writes that "in mature science there are always far too many . . . areas in which no existing paradigms seem obviously to apply and for whose exploration few tools and standards are available." To the extent that paradigms can be identified with theories, Kuhn writes that "theories are, even more than laboratory instruments, the essential tools

of the scientist's trade". See Kuhn, *The Essential Tension: Selected Studies in Scientific Tradition and Change*, p. 234.

19. Kuhn, *Structure of Scientific Revolutions*, p. 37.
20. Ibid.
21. Ibid., p. 65.
22. Michael Dickinson, "Solving the Mystery of Insect Flight," *Scientific American* (2001).
23. Kuhn, *Structure of Scientific Revolutions*, p. 64.
24. Weber, " 'Objectivity' in Social Science and Social Policy," p. 68.
25. Ibid., p. 90.
26. Ibid., p. 103.
27. Ibid., p. 69.
28. Ibid., p. 105.
29. Yochanan Shachmurove, Gideon Fishman, and Simon Hakim, "The Burglar as a Rational Economic Agent," *CARESS Working Paper* (1997): p. 1.
30. Amyra Grossbard, "Towards a Marriage between Economics and Anthropology and a General Theory of Marriage," *American Economic Review* 68, no. 2 (1978).
31. Chung cheng Lin, Tze wei Chen, and Ching chong Lai, "The Economics of Honeybee Swarming," *Regional Science and Urban Economics* 33, no. 5 (2003): p. 581.
32. Weber, " 'Objectivity' in Social Science and Social Policy," p. 84.
33. Ibid., p. 113.
34. Larry Wright, *Teleological Explanations: An Etiological Analysis of Goals and Functions* (Berkeley: University of California Press, 1976), p. 13.
35. Weber, " 'Objectivity' in Social Science and Social Policy," p. 94.
36. Ibid., p. 79.
37. Following Kuhn's use of the word, when it comes to the transition from the pre-paradigm sciences, "paradigm" is a success term: "to be accepted as a paradigm, a theory must seem better than its competitors." But when it comes to the transition from revolutionary sciences to normal science, "paradigm" is not used as a success term, so that Kuhn speaks of "competing paradigms" and "paradigm destruction". Given Kuhn's usage, "paradigm destruction" would be impossible in the transition from pre-paradigm science to normal science. While there is nothing that in principle is wrong with Kuhn's use, it does mask over some important similarities that exist between pre-paradigm science (where there are competing theories or schools) and revolutionary science (where

there are competing paradigms). See Kuhn, *Structure of Scientific Revolutions*, pp. 17, 67.

38. Note that while I have confined my attention to the transition from pre-paradigm and revolutionary science to normal science, I believe that similar remarks almost certainly apply to the transition from pre-science discursive practices (religion, etc.) to pre-paradigm science.
39. Kuhn, *Structure of Scientific Revolutions*, pp. 16–17.
40. Ibid., p. 14.
41. Ibid., pp. 84, 97.
42. Weber, " 'Objectivity' in Social Science and Social Policy," p. 70.
43. I do not want to suggest that these are the only possible responses to crisis. For example, the counterinstances in questions could be found to be false, the product of, say, inaccurate instrumentation. For example scientists have been unable to explain the prolific amount of lava produced by Mount Fuji until they recently discovered a crack in the tectonic plate just below the mountain; solving this anomaly did not stress any existing paradigms about magma production.
44. Kuhn, *Structure of Scientific Revolutions*, pp. 52–3.
45. Ibid., p. 78.
46. Ibid., p. 76.
47. Weber, " 'Objectivity' in Social Science and Social Policy,' p. 70.
48. Ibid., p. 81.
49. Ibid., p. 105.
50. Kuhn, *Structure of Scientific Revolutions*, p. 71.
51. Ibid.
52. Dickinson, "Solving the Mystery of Insect Flight," p. 55.
53. Ibid., p. 57.
54. N.J. Miller, "Contributions of Social Capital Theory in Predicting Rural Community Inshopping Behavior," *Journal of Socio-Economics* 30 (2001): p. 477.
55. Michael Lynn and Andrea Grassman, "Restaurant Tipping: An Examination of Three 'Rational' Explanations," *Journal of Economic Psychology* 11, no. 2 (1990): p. 169.
56. Eric B. Budish and Lisa N. Takeyama, "Buy Prices in Online Auctions: Irrationality on the Internet?" *Economics Letters* 72, no. 3 (2001): p. 325.
57. Weber, " 'Objectivity' in Social Science and Social Policy," p. 105.
58. Ibid.
59. Ibid., pp. 90, 91.

60. Ibid., p. 90.
61. Ibid., p. 81.
62. Ibid., pp. 90, 91.
63. Weber in Burger, *Max Weber's Theory of Concept Formation*, p. 66.
64. Compare Weber's view to John Dupré's promiscuous realism: it is not the case that "science can be expected to provide . . . a goal-independent set of classifications. The reason I reject this idea is not, as is generally supposed by constructivists, that nature provides no objective divisions between kinds but rather that it provides far too many. This is the promiscuity in the thesis of promiscuous realism." See John Dupré, "Are Whales Fish?" in *Humans and Other Animals* (Oxford: New York Oxford University Press; Clarendon Press, 2002), p. 54.
65. Weber, " 'Objectivity' in Social Science and Social Policy," pp. 105–6.
66. Ibid., p. 105.
67. Ibid., p. 97.
68. Kuhn, *Structure of Scientific Revolutions*, p. 84.
69. Weber, " 'Objectivity' in Social Science and Social Policy," p. 66.
70. Virginia Woolf, *The Voyage Out* (London: Duckworth, 1915), p. 43.
71. Weber, " 'Objectivity' in Social Science and Social Policy," p. 91.
72. The Economist, "Behaviourists at the Gates," *The Economist* (2003).
73. Weber, " 'Objectivity' in Social Science and Social Policy," p. 90.
74. No question further research might help: Gary Becker, at the University of Chicago, aims to demonstrate that behavior which appears to be irrational is indeed rational. But even these moves stretch the notion of rationality, moving economists to anticipate behavior in new ways. If this is the case, there is only a sense in which the old neoclassical model is vindicated by such authors.
75. *The Economist*, "Rethinking Thinking," (1999).
76. Ibid.
77. Helen Longino explicitly advocates the latter approach. Advocates of competing ideal types, paradigms, or epistemologies are brought together in an "interactive dialogic community". Longino recommends this stance as an effective way of highlighting and eliminating assumptions, prejudices, and socio-political commitments which help shape theory. See Helen E. Longino, *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry* (Princeton, NJ: Princeton University Press, 1990).
78. Kuhn, *Structure of Scientific Revolutions*, p. 84.

79. See note 38, above.
81. Kuhn, *Structure of Scientific Revolutions*, p. 85.
82. Ibid., p. 15.
83. Larry Wright, *Critical Thinking* (Oxford: Oxford University Press, 2001), p. 232.

5

Searle and the Ideal Type—Applications of the Constitutive Formula

I am concerned with the significance of the constitutive formula as a means by which we solve puzzles in discursive activities.

The criticisms outlined in Chapters 2 and 3, combined with my Chapter 4 discussion of Kuhn's notion of a paradigm and Weber's notion of an ideal type, will allow me to argue that the constitutive formula is an ideal type. This formulation allows me to avoid the two criticisms directed at the atomist reading of the constitutive formula while simultaneously connecting Searle's remarks to an established methodological stance within the social sciences.

We can distinguish the more general claim that social reality can be described according to the "X counts as Y in C" formula from the more specific articulations of our institutions—green pieces of paper count as money. Building off my previous discussion of money, I will be confining my attention to particular claims about our institutions made with the constitutive formula. I claim that if a particular application of the constitutive formula tells us something about our institutions, it does so in something of the same way that a paradigm tells the natural scientist something about the physical world. In brief, I will argue the orthodox and chartalist accounts of money are best understood as ideal types.

I will conclude by arguing that the ideally typical formulation of these competing ways of describing our financial institutions helps Searle avoid Wisdom's objection. My discussion of money aims to contest the suggestion that institutional reality has a particular logical structure. In this chapter, then, I hope to underscore

the difficulties of the atomist reading of the constitutive formula while showing how an appeal to the ideal type helps us avoid these problems.

This chapter also sets the stage for the more important discussion that takes place in the next chapter. There, I will step back and try to say something about the constitutive formula itself. Locating Searle's project within territory already traversed by Weber, I will claim that the formula "X counts as Y in C" (and not *just* any particular institutional depiction) is helpfully understood as a kind of ideal type.

In Chapter 2, I argued that money can be articulated in at least two ways. The orthodoxy defines money as a medium of exchange whereas the chartalist contends that this view is incompatible with the historical evidence. Money is rather an indication of debt-owed, and whatever object is used to tabulate debt counts as money. Thus money might be a name on a clay tablet, although a gold token might also serve to indicate that a central authority is somehow indebted to the coin's owner.

This discussion aims to show that the orthodox and chartalist accounts of money are ideal types that we might use to better understand economic phenomena. Both labels gesture at clusters of interrelated ideally typical pictures. Let me briefly try to disentangle *some* of those underlying pictures.

First, money as a medium of exchange implicates the marketplace (village fair) as the ideal context; rational transactors are its ideal participants; these participants are motivated by the desire to acquire something they do not have but need or want (accretion); improvements in efficiency suffice as an ideal explanation of change and innovation.

Second, money as an indication of debt implicates the asymmetrical relationship between ruler and subject as the ideal context; participants are rationally responsive, not just to market value, but to power (coercion, duress, etc.); these participants are motivated, not so much to gain what they do not have, but to

compensate for what they have taken. Innovation (bank currency, etc.) is explained as a response, not just to gains in efficiency, but primarily to the demand for financing.

These are two competing accounts of money. Both of these accounts are quasi-historical narratives that help explain how the use of money evolved from a fundamental application (medium of exchange or debt-owed). What is the significance of those narratives?

One function of the narrative is obvious: it organizes and illustrates the relationships between the various pictures (the context, participants' psychologies, their motivations, explanation for change) that constitute, say, the orthodox view. The narrative may also accomplish other functions. For example, Wray and Ingham explicitly treat their origins-account as *empirical evidence* for the correctness of their account of money or the incorrectness of the orthodox view: "There is no evidence of barter-based markets (outside of trivial prisoner-of-war cases), and all the evidence about the origins of money points to state involvement".[1] "In short, monetary practice has its logical origins in money of account and its historical foundation in the chartal money of early bureaucratic empires. It was not, *pace* Menger, the spontaneous product of the market".[2] Both authors use anthropological evidence to support the view that coins, like tallies, were first used, not to lubricate barter, but to mark the state's debt to administrative and military bodies. These debt-markers could be then exchanged for goods and services with the taxed, those who *owed* a debt to the state. "Money, then, originated not as a cost minimizing medium of exchange, but as the unit of account in which debts to the palace (tax liabilities) were measured".[3]

Kevin Dowd, who champions the orthodoxy, concedes that there is little anthropological evidence in support of the village fair conception of the origin of money. But he worries whether those who focus on historical accuracy have missed the point of such narratives: "A conjectural history provides a benchmark to

help assess the world we live in, but it is important to appreciate that it is *not* meant to provide an accurate historical description of how the world actually evolved".[4] He goes on to describe the conjectural history as a "useful myth".

Dowd is largely right, so long as we are clear that not any old conjectural history will do. Separating legitimate conjectural histories from illegitimate ones, we might be able to get clearer on the sense in which anthropological evidence can *also* be helpful in assessing what money might be.

The founding of these ideal types in our background understanding

There are a number of functions that quasi-historical narratives perform. These genealogies, for example, harness our existing understanding. Obviously, citing historically accurate cases is an effective way of doing this but, *à la* Hobbes or Rawls, this function does not require that the narrative be veridical. In the last chapter I suggested that for both Kuhn and Weber, part of what makes a paradigm or ideal type compelling is the extent to which it appeals to and builds from our existing antecedent understanding. Electricity was characterized as a "fluid" by one camp and "effluvium" by another, bringing to bear a host of extant perceptions and expectations already developed in other areas of experience and inquiry. Old skill sets are often applicable to new activities.

Part of what makes the orthodox view or "real analysis"[5] so compelling are the intuitive pictures evoked. Geoffrey Ingham is clear about this when he writes:

> Real analysis and, ultimately, the equations of general equilibrium models are not, as is generally supposed, purely the results of the axiomatic-deductive method. The "real economy" abstraction actually derives from an inaccurate historical

> concept of a small scale, pre-capitalist "natural economy" or the "village fair". In this model, economic activity is seen to involve routine spot trades in which media of exchange can be readily taken to be the direct representation of real commodities—that is, as their "vehicles"—by continuously transacting economic agents.[6]

Fundamentally, Ingham states, the modern economic marketplace is modeled on a certain conception of a "village fair," in which buyers and sellers freely come together to exchange goods and services. The means by which these goods and services are exchanged has evolved, as have the number of goods and services which might be traded. But throughout the genealogy the implicit circumstances, purposes, and motivations of the fair-goers remain the same, untouched by marked improvements in efficiency.[7]

The narratives that underlie the competing accounts of money, whether historical or conjectural, are pedagogical. There is much in our existing understanding that we can appeal to in order to give us a tighter grip on the nature of money. For example, reviewing the orthodox account, we might think back to our experiences at a swap-meet or even purchasing items in an online auction. In these contexts we understand something about the motivations of the various participants: the livelihood of the seller, perhaps, depends on his getting the highest price for an item while remaining competitive. In the online auction format, we see prospective buyers negotiate with others over the perceived worth of a given item. These experiences, or perhaps just the canonical articulation of these experiences, instill a certain picture of what kinds of activities take place in a village fair. The conjectural narrative draws off this antecedent understanding, and gives significance to the pieces of metal and bits of paper in our pockets.

Likewise, our understanding of taxation or debt is a large part of what makes the chartalist account not only intelligible, but plausible. Dowd has correctly argued that if the conjectural history

mobilizes our extant understanding in the right way—even if the conjectural history is not in fact supported by anthropological evidence—it has served one of its primary functions.

The ideal types of money and the first response to crisis

However, Wray is also right to be concerned when he noticed disparities between the historical record and expectations laid down by the orthodox ideal type. These disparities give rise to a puzzle, and it is up to the defenders of the orthodox view to somehow respond to this puzzle.

I argued that ideal types are tools or instruments in two senses. First, they allow us to see a phenomenon as puzzling. A puzzle or crisis arises when there is a disparity between the data and the expectations embodied in the ideal type. Second, ideal types are tools in that they help the inquirer resolve the puzzle. For Weber there are at least two possible responses to such a puzzle, the first of which I will discuss in this section. According to the first response, while the ideal type provides a rough solution-sketch to the puzzle's resolution, it must also be stretched or modified to cover the aberrant phenomenon. This response initiates a discursive trajectory by which our understanding is gradually increased: this refined and expanded ideal type, in turn, creates the possibility of new kinds of puzzles.

With respect to the orthodoxy, historical evidence is a principal source of puzzlement. If, for example, the use of precious metals for exchanges brings about such massive gains in efficiency, why was the first coin not issued prior to the seventh century BC? Why were denominations of these coins so high that they could not be used for "ordinary" marketplace-type exchanges? Why did the clay tablets that tallied debts appear at least 2,000 years before these first coins were issued?[8]

Bracketing Wray's questions for a moment, it is worth emphasizing how successful the orthodox view has been in answering other puzzles: for example, what would motivate someone to exchange precious metal for *bank currency*? What are we exchanging upon our departure from the gold standard? Is not money also a unit of account or a store of value? The evolutionary account ties up these loose ends in a process resembling normal science—the same constellation of ideal types is tailored or modified to provide solutions to these puzzles. The solution sketch is there—the introduction of every novel phenomenon is explained by an increase in efficiency. The task is to fill in the details so as to explain how, for example, exchanging one's readily tradable gold coins for seemingly less tradable bank currency or receipts could possibly constitute a gain in efficiency:

> The use of coins still involves considerable costs, particularly those of storing, protecting and moving coins around. To save on some of these costs, some people come to be prepared to pay others to store their gold for them. Goldsmiths and some merchants already have facilities to keep large amounts of gold, and can therefore keep additional quantities of it at relatively low marginal cost. These people find it profitable to accept gold for safekeeping for a fee that many current holders of gold are willing to pay, and depositors are issued with receipts that gives them the right to demand their gold back.[9]

Thus, a case that prima facie looks importantly different from the county fair is in reality just a special instance of it. This is the second sense in which I have pointed to the ideal type or paradigm as an instrument or tool: to help resolve a puzzle raised by the paradigm itself. More fundamentally, the narrative reminds the reader that these *are* differences that need to be reconciled. It is only against the background of a certain conception of the

function of money that the idea of bank currency as a kind of money might even strike one as remotely problematic or untypical in the first place. It is precisely in this sense that an ideal type is a tool which allows us to see a phenomenon as potentially puzzling.

But despite the orthodox ideal type's success in addressing many questions (why would someone exchange gold for bank currency?), there are other questions (why would the first coins be issued in such large denominations?) that remain intractable. Not incidentally, these are the same questions that are effortlessly handled by the chartalist ideal type. This suggestion anticipates Weber's second response to crisis, wherein additional ideal types are evoked to help account for the puzzle.

Still, it is not absurd to think that a researcher might be able to account for the fact that gold coins came so late onto the scene in such large denominations by appeal to something like gains in efficiency. In this way, a linear trajectory—puzzle, followed by a solution via a refinement of the ideal type, leading to new puzzles, and so on—is maintained. Details of the basic orthodox conjectural history might shift (it turns out that gold coins, as a medium of exchange, have unanticipated consequences), but the overall integrity of the ideal type is largely preserved.

The ideal types of money and the second response to crisis

While the linear progression implied by serial modifications of the same ideal type is a helpful way to advance understanding, it is not the only way. Both Kuhn and, in particular, Weber argue that a variety of ideal types or paradigms can be simultaneously brought to bear on a given phenomenon.

Weber argues that, because of the infinite multiplicity of reality, the most we can expect of a given ideal type is that it accentuates and suppresses various features of this reality: if a given ideal type cannot purport to provide the means to "copy"

the infinite multiplicity, then the door is open to using several different ideal types, each one of which draws out various features of this multiplicity.

Refining an ideal type in the face of counterinstances is not the only way by which we can advance our understanding. A particularly intractable counterinstance can also be seen as an indication that the researcher has stumbled upon the limits of an ideal type's scope. Indeed if Weber is right about the infinite multiplicity, we must assume that every ideal type breaks down eventually. In this case, it might be appropriate to employ an additional ideal type in an attempt to cover the recalcitrant phenomena.

This does not imply that the original ideal type is entirely abandoned—it may have well established credentials for certain kinds of puzzle. Nevertheless, the inquirer comes to see that the ideal type may not be helpful in solving the particular puzzle at hand. Of course, there is no definitive way to judge whether an ideal type is being overextended; the ideal type can always be pushed and modified in different degrees, although one has to assume that, as far as the ideal type's usefulness is concerned, at some point something like a law of diminishing returns sets in. Stretching may incur costs, both in terms of the perspicuity and even intelligibility of the ideal type.

Looking at money, I want to return to Dowd's characterization of the orthodox history: "A conjectural history provides a benchmark to help assess the world we live in, but it is important to appreciate that it is *not* meant to provide an accurate historical description of how the world actually evolved".[10] Dowd offers a sophisticated recharacterization of the orthodoxy's significance. In particular, he contrasts "the world we live in" with the historical case,[11] which delimits the scope of applicability of the neoclassicalist ideal type. The point of the model was to tell us something about how money normally functions *today*. The conjectural history is a tidy way of bringing a number of seeming disparate phenomena under the roof of the orthodox ideal type. Money is a

medium of exchange in today's world because, the story goes, its use leads to gains in efficiency. This may or may not account for why, say, coins were introduced in the past, but helps account for their current use.

In the face of all the questions or puzzles we might have about a phenomenon, there must be a limit to what we can expect of a particular ideal type. Moreover, we should be able to say something about where these limitations lie; Dowd does this temporally, distinguishing the world we live in from the historical case. The inability of the orthodox model to resolve historical counterinstances does not *falsify* it (as Wray seems to suggest); it is rather that we seem, at least for the time being, to have stumbled upon the kinds of phenomena it is less adept at illuminating.

With this in mind, let us return to Wray's shrewd attempt to articulate the significance of his chartalist ideal type. The neoclassicalist model cannot be reconciled with the anthropological facts:

> The problem is that the Never-Never Land imagined by the Paul Samuelsons and other textbook writers simply never, ever, existed. There is no evidence of barter-based markets (outside trivial prisoner-of-war cases), and all the evidence about the origins of money points to state involvement.... [O]n closer inspection, it becomes obvious that [the historical facts] do not support the Samuelsonian hypothesis about the origins of markets and money.[12]

Wray takes it for granted that the point of the conjectural history was to map the outlines of an actual history; the ideal type was intended to be a "hypothesis about the origins of markets and money". Dowd, as we have seen, does not think this but surely other neoclassicalists have. In a parenthetical remark Wray manages to dismiss as "trivial" the kinds of prisoner-of-war situations where the orthodox ideal type does seem to work because this situation is significantly different from the highly centralized

economies of the ancient Phoenicians, and so gets the history wrong. Wray is right to distinguish the prisoner-of-war case from these temple and palace communities, but he is wrong to dismiss the former as "trivial" because of that difference. Dowd is right to claim that Wray has missed the point of the orthodox conjectural history. The fact that it does not cover all the historical instances does not falsify the orthodox ideal, but rather helps us get a better sense of the limitations of its reach.

The chartalist view is also an ideal type with its own limitations. Wray's model nicely accounts for the accoutrements and behavior exhibited by those in these centralized debtor bureaucracies. But this alone does not imply that Wray's ideal type is applicable to modern economic activity; after all, perhaps modern economies resemble the prisoner-of-war case (or the village fair or Internet auction site) more than they resemble the economies of temple and palace communities. However, to be fair to Wray, it seems unlikely that the way we use money has changed so radically over time that his ideal type is no longer applicable; that the chartalist covers the historical cases is prima facie evidence for its capacity to illuminate at least aspects of modern economic activity. But even if the chartalist model is applicable to the modern world, it does not preclude the possibility of other explanatory models, as he seems to maintain.

But, if Wray is right, the respective domains of the two ideal types is probably not best described as reflecting differences between the world we live in and the historical world, as Dowd suggests. Keeping our sight fixed on the present, it seems as though Wray's model might be more useful in characterizing highly centralized, regulated economies whereas the orthodox ideal type might be more useful in characterizing free-market situations. The two differing ideal types reflect the not-at-all absurd suggestion that money functions in a different way for governments and its debtors, than it does when it is in the hands of private consumers. The historian can likewise probably find

evidence for both of these patterns in history. Money is a complicated sort of object that functions in different ways in a variety of overlapping contexts. Again, if social reality consists of an infinite multiplicity, as Weber suggests, the fact that it cannot be articulated under a single, neat, canonical model should not surprise us.

Conclusion: returning to John Wisdom's objection

According to Searle, iterations of "X counts as Y in C" provide the logical structure of complex societies. I characterized this claim by means of an analogy with the atomists, who explicated the structure of brute reality in terms of analytic propositions. I also claimed that given the extent to which the analogy between Searle and the atomists obtains, objections originally brought against the atomists might also apply to Searle's thesis. I would like to briefly return to John Wisdom's objection as I related it to institutional analysis. In analyzing what it is for, say, England to declare war, Wisdom worried that it is always possible to imagine alternative ways England might have declared war. Any analysis, then, will involve an infinite disjunction of conjunctions of statements, or else be cut short prematurely with an etc-clause.[13] According to this objection, if institutional analysis is understood on the model of philosophical analysis then it seems that analysis is impossible: we can always imagine alternative ways to extend the analysis of a given phenomenon. We have seen this possibility play out in the case of money, which seems to defy simple analysis under the constitutive formula. When Wisdom's objection is applied to the particular case of the constitutive formula, the infinitude extends outward from *two* points—from both the X and the Y terms.

First, we can always imagine alternative brute facts (X) which might satisfy a given Y term. For example, any number of objects—paper with markings printed by a government, paper with markings printed by private banks, wooden tallies, gold

coins, electronic data, Marlboro cigarette packs, cowrie shells, clay tablets, woodpecker scalps, feathers, pig tusks, etc.—are understood to count as media of exchange (Y). We might speak of an infinitude of possible brute facts (X), rendering analysis impossible.

We can see how preconceiving the constitutive formula as an ideal type caps the proliferation of X terms by considering the actual process by which puzzles arise. I claimed that a puzzle arises when some observation conflicts with our expectations. Confining our attention to the orthodox ideal type, bank currency might be seen to produce such a conflict: why would someone exchange a readily tradable commodity (gold coins) for potentially less tradable bank currency? A similar conflict of expectations arises when we look at our departure from, for example, the gold standard. The ideal type is often refined in one way or another to account for these anomalies.

Notice the kinds of puzzles that did not arise over the course of investigation: if X is the bank currency, or currency not tied to precious metals, we have a puzzle. If X is a gold coin, woodpecker scalp, feather, electronic data, or even a Marlboro cigarette pack, we do *not* necessarily have a puzzle (although we could). The sense in which coins might make a good medium of exchange is obvious to anyone who already understands the orthodox ideal type. If pressed, we might be able to say something about why a coin (X) might count as a medium of exchange (Y): while they are not easily divisible, we effortlessly recognize them to have a number of other features that make them good candidates for a medium of exchange: they are easily identifiable, relatively unusual but not rare, storable, durable, compact, and easily transportable.[14] On the other hand, because gold coins have all of these features, it is difficult to understand why self-interested agents would be motivated to use bank currency—we may, then, have a genuine puzzle.[15]

Of course, in defense of the shift to bank currency, the orthodox economist will respond by noting that large quantities of gold

coins are *not* compact and easily transportable, unlike bank currency. "The use of coins still involves considerable costs, particularly those of storing, protecting and moving coins around".[16] The errant application is thereby made congruous with the ideal type. If institutional analysis is successful, it is to the extent that it is able to anticipate all the X's that might count as Y. This is impossible. Articulation is necessary, not in the case of coins, but only when the inquirer is confronted with an actual puzzle. Why would bank notes (X1) as opposed to gold coins (X2) come to count as a medium of exchange (Y)? Puzzles arise piecemeal, as we happen across aberrant data. Likewise, when the orthodoxy is conceived as an ideal type, success comes when the inquirer can reconcile a puzzling X with the expectations embodied in the Y term.

Second, it seems possible that there are an infinite number of ways to recharacterize the Y term. We have seen how money could be conceived as a medium of exchange or an indication of debt-owed. Additional characterizations, perhaps in terms of unit of account or repository of value, are available. I have suggested that ideal types (as well as paradigms) explicitly consolidate an array of largely inarticulate expectations. I traced the intuitive appeal of the orthodox ideal type back to, among other things, our experiences and expectations in village-fair-type economic situations. Likewise, the chartalist ideal type can be traced back to our expectations that arise from our experiences of paying taxes, etc. While our financial experience is complex and could be used to generate a great variety of ideal types, this same experience puts intuitive constraints on which of these models are worth pursuing in detail.

This point can be put as follows: just as a given ideal type (Y) constrains the kinds of puzzles we might actually have about a given X (why is bank currency a more suitable medium of exchange than gold coins?), our general understanding of how the world actually functions places constraints on the puzzles we might actually have about the nature of money. Money might be

plausibly construed as an indication of debt-owed or a medium of exchange. Perhaps it might also be conceived as a unit of exchange. In a world a little bit different from our own, other ideal types might seem to hold more promise for inquiry. But the possibility of being able to imagine such ideal types should not, contra Wisdom, in any way bear on whether or not inquiry is successful. Just as our understanding of what it is for England to declare war should not depend on our ability to articulate all the alternative ways in which England might have declared war, we need not articulate all the possible ways in which something might come to count as money. A discursive activity only requires one or two conspicuous articulations to advance our understanding. Thus, by conceiving applications of the formula "X counts as Y in C" as ideally typical in Weber's sense, it is possible to sidestep the force of Wisdom's objection. This is done, not by showing Wisdom to be wrong (indeed, Weber's contention about the nature of reality seems to corroborate Wisdom's point), but by recasting the atomist's criterion for success.

For the atomist, success is possible only when a given phenomenon has been exhaustively analyzed. Wisdom argues that this is impossible. Weber asks us to rethink our criterion for success by reflecting on the actual process of a discursive practice. Speaking of the constitutive formula in particular, it is not necessary to subsume every imaginable application under the formula "X counts as Y in C". It may be the case that there are infinite ways to characterize money under the constitutive formula, but we do not need to specify those *in advance* as required by the atomist reading. Success comes as actual puzzles are actually resolved, not when all real and imaginable puzzles are resolved. Inquiry is thus reconceived as a kind of cavalcade. For this reason, institutional analysis is no longer impossible; we are no longer threatened by Wisdom's infinitude.

But even if it is the case that competing articulations of the institution of money are best understood as ideal types, this does not yet

imply that society does not have a logical structure. That is, even if we have shown that there are significant difficulties in articulating the logical structure of a given institution, Searle may still contend that society *does* have such a logical structure, if only on a very general level. Our discussion thus far implies, at best, that the logical structure of our institutions cannot be articulated except in the most general terms—"X counts as Y in C". In the next chapter I will argue that the constitutive formula *itself* is best understood as a kind of ideal type—a claim that takes aim at the heart of the atomist reading.

Notes

1. Wray, "The Property Theory of Interest and Money," p. 42.
2. Ingham, "Modern Money," p. 26.
3. Wray, "The Property Theory of Interest and Money," p. 43.
4. Dowd, "Aristotle on Money," p. 139.
5. "Real analysis" is the label that J.A. Schumpter, in his classic *History of Economic Analysis*, gave to the view that all financial transactions, and the function of money in general, can be ultimately understood as a version of barter. See Joseph Alois Schumpeter, *History of Economic Analysis* (New York: Oxford University Press, 1954).
6. Ingham, "Modern Money," p. 17.
7. In this passage Ingham only suggests that the village fair model of economic transaction helps explain the endurance of the model of the ideally rational economic agent.
8. Wray, "The Property Theory of Interest and Money," pp. 44–5.
9. Dowd, "Aristotle on Money," p. 144.
10. Ibid., p. 139.
11. Dowd might also be implying a distinction between a "benchmark" and "an accurate ... description," suggesting that he and, say, Wray are working with fundamentally different methodological pictures. The notion of a "benchmark" seems to readily dovetail with the notion of an ideal type or paradigm, as explicated above. It evokes Wittgenstein's discussion of "standards" in *Philosophical Investigations*, and suggests, as per my discussion in Chapter 4, that these models constitute the contrast-class against which questions arise. "An accurate ... description,"

on the other hand, calls to mind the Popperian portrait of falsification, where intractable anomalies function as counter-examples and provide the basis for the rejection of competing theories. Wray seems to implicitly endorse the latter methodological stance when he says that anthropological findings "do not support the [orthodox] hypothesis about the origins of markets and money". This suggestion is merely speculative, as it banks on enunciating "an accurate description" and "hypothesis" in peculiar ways. Wray, "The Property Theory of Interest and Money," p. 43.

12. Ibid., pp. 42–3.
13. Wisdom, "Metaphysics and Verification," p. 478. See also Urmson, *Philosophical Analysis; Its Development between the Two World Wars*, pp. 151–2.
14. Baumol and Blinder, *Economics, Principles, and Policy*, p. 201.
15. I am not suggesting that we cannot imagine genuine puzzles when it comes to the use of woodpecker skulls or other such objects. Consider, for example, the case of the Yap community of the Western Caroline islands of Micronesia. Carved stone disks were in wide circulation until the mid-nineteenth century, some of which had a diameter greater than four meters. An economist or social scientist might be puzzled to find that these disks are still used today "for various social exchanges, their worth dependent on the size, shape, quality of stone and history behind each particular piece" given the availability of other media of exchange. See Fitzpatrick, "A Massive Undertaking: Examining Stone Money in Its Archaeological Context," p. 332.
16. Dowd, "Aristotle on Money," p. 144.

6

Searle and the Ideal Type—the Constitutive Formula and the Status-function

In this chapter I will argue that both Searle's notion of the constitutive formula and the status-function are ideal types.

According to Weber, social reality consists of an "infinite multiplicity" that resists articulation under any one schema. Ideal types, then, are tools of inquiry that unavoidably both accentuate and suppress various features of this reality. One way of getting at the ideally typical nature of Searle's proposal is to show that there are other models that highlight aspects of our institutions otherwise obscured by the constitutive formula or status-function. I will propose an alternative model, appropriating Alasdair MacIntyre's discussion of virtue.

Before arguing that Searle has formulated an ideal type, I will begin by revisiting Searle's various characterizations of the "X counts as Y in C" formula.

Is the constitutive formula an ideal type?

Would Searle agree with my characterization of the constitutive formula as an ideal type? In surmising an answer to this question, Searle's own remarks display a striking ambivalence. I will first consider his statements that would seem to shun the suggestion that the constitutive formula is an ideal type. I will then consider other of his claims that appear to endorse it. If this endorsement holds fast, this has the happy consequence of aligning my reading of the constitutive formula with that of its author.

The constitutive formula is not an ideal type

Is the constitutive formula an ideal type? If we confine our attention strictly to the *Construction*, I think the answer would have to be "no". As we have seen in Chapter 1, some variation on the atomist reading of the formula is closer to what Searle seems to have had in mind. Searle claims that "The structure of institutional facts is the structure of hierarchies of the form 'X counts as Y in context C' " (*CSR*, p. 55) and, further, "It is no exaggeration to say that these iterations provide the logical structure of complex societies" (*CSR*, pp. 22, 56, 80, 90, 191).[1] In *Speech Acts*, Searle asserts that "every institutional fact is underlain by a rule of the form 'X counts as Y in C' ",[2] and again, in the *Construction*, "all institutional forms of human culture ... must always have the structure X counts as Y in C ..." (*CSR*, p. 40). When Searle describes the *Construction* as an attempt to "investigate the logical structure of institutional reality" (*CSR*, p. 94), it does not seem as if he is (just) giving us an instrument that may aid some research program. Elsewhere Searle writes that "I am simply describing the structure whereby institutional reality actually works in real human societies" (*CSR*, p. 45). It is difficult to understand the sense in which Searle has offered us an ideal type (which implies the possibility of other ideal types) and at the same time is doing foundational ontology (*CSR*, p. xii), describing the basic, logical structure of institutional reality. If, in the *Construction*, Searle has provided us with an ideal type, he has done so in exceptionally misleading terms. Indeed, Searle's own framing of the project, so akin to that of the atomist's, appears almost irreconcilable with the ideally typical reading.

Nor is the atomist reading unusual among Searle's readers, as evidenced by Dreyfus' and Hacking's understanding of the aims of the *Construction*. Dreyfus, following an extended clarificatory dialogue between Searle and himself, feels as though he has finally understood Searle's position: "I now understand that, when Searle analyzes the role of propositional representations in

constituting actions and institutional facts, he is doing logical analysis ...".[3] Hacking's characterization of the *Construction* is even more explicit in tying Searle's project to that of the atomists: "Searle uses the word 'construction' in a rather literal way. Chapter 1 is called 'The Building Blocks of Social Reality'. He has the bricks-and-mortar connotation of 'construction' in mind. He writes in a philosophical tradition that includes Bertrand Russell ('logical construction') and Rudolf Carnap (Aufbau)".[4]

The constitutive formula is an ideal type

There is also very good reason to think that Searle *would* endorse my characterization of the constitutive formula as an ideal type. Recall my discussion of Searle's reply to Barry Smith's critique of the constitutive formula (Chapter 1). Because there are free-standing Y terms (promises remain applicable even after the words/sounds are uttered), Smith argues, it cannot be the case that X counts as Y in C represents the logical structure of institutional reality.

Searle responds by denying that he was interested in articulating conditions for the possibility of institutional reality in the first place; he contends that it is possible to articulate the logical structure of reality while not doing logical analysis. Searle also offers a positive characterization of what he is doing in the *Construction*: first, the constitutive formula is just "useful mnemonic" and, second, he is anyway principally interested in the status-function.

Searle writes that "the formula *X counts as Y in C* is intended as a useful mnemonic".[5] Following my characterization of the ideal type, a fair paraphrase of a "useful mnemonic" is a tool or instrument that supports and increases our understanding. The constitutive formula, I will argue, is useful in getting us to see aspects of social reality that do not have brute concomitants, flagged by the status-function. So, "This useful mnemonic is not intended as a

definition of *social objects* or even of institutional facts . . .".[6] Searle explicitly and unambiguously distances himself from the atomist reading of the constitutive formula.

Depending on what Searle means by a "useful mnemonic," he might be sympathetic to my characterization of the constitutive formula as an ideal type. X counts as Y in C is a mnemonic, a reminder, an ideal type, or a tool that helps us advance our understanding of a particular feature of institutional reality. One sense in which ideal types are tools is that they help the inquirer solve puzzles that come up over the course of a discursive activity—they provide solution sketches.

Not just Smith, but other readers of the *Construction* such as Dreyfus and Hacking, would find Searle's response surprising. Searle has at once distanced himself from the atomist reading of the *Construction*, endorsed the ideally typical reading the same text, and displaced the centrality of the constitutive formula in favor of the status-function. Indeed, Smith argues this "fateful admission" marks a significant break from the overall tone of the *Construction* itself.[7]

After so characterizing the constitutive formula, Searle goes on to highlight the status-function as the principal concern of the *Construction*. Recall that the Y term denotes the status-function or institutional fact. It contains two aspects: it is an agentive function, but one that confers rights and obligations to its bearer.

For Searle, it makes no difference whether Smith has discovered institutional facts without corresponding X facts, thereby falsifying the formula, because he is principally interested in just the Y term anyway. Searle affirms that "My concern, in short, is with institutional reality, which is a special case of social reality. It is a matter of status-functions, it is about the deontic powers accruing to status-functions, and it is utterly naturalistic".[8] In this way Smith's counterexamples do not threaten the overall thrust of the *Construction*: an institutional fact "need have no physical realization; it may be just a set of status-functions".[9]

In spite of Searle's response to Smith, I am going to argue that the atomist, rather than Weber, more perfectly represents Searle's metaphilosophical impulses. If, as suggested in Chapter 1, the X term is limited to persons, it is not clear that there is any significant difference between the constitutive formula and the status-function. Rights and obligations must be embodied. Nevertheless, if we follow Searle in not limiting the X term in this way, then it is clear that the constitutive formula is just a special case of the status-function. If we can characterize the status-function as an ideal type, *a fortiori* the constitutive formula, which Searle has already described as a "useful mnemonic," can also be rendered as such. For that reason, I will inquire into the status of the status-function.

Is the status-function an ideal type?

Searle would seem to agree that the constitutive formula is an ideal type or mnemonic that may be useful in helping the sociological or philosophical inquirer overcome certain puzzles that arise from framing social phenomena in such a way that the normative aspect is lost or abridged. If this is a fair characterization of the role of the constitutive formula, then on one level Searle appears to be in agreement with Weber's general depiction of discursive activities: ideal types are tools that help resolve puzzles. Nevertheless, I do not think that Searle's treatment of the status-function is entirely analogous to his treatment of the constitutive formula, limiting the applicability of the ideal type to Searle's project. Is the status-function, like the constitutive formula, *another* useful mnemonic or ideal type? And if not, how are we to understand *its* significance?

There are passages in the *Construction* where Searle claims that the logical structure of institutional reality is constituted by iterations of X counts as Y in C (*CSR*, pp. 55, 80); these sort of

statements provide part of the basis for the atomist reading of the constitutive formula. However, after looking at deontic status-functions in particular, he describes *these* as constituting the logical structure of institutional reality (*CSR*, pp. 109, 110, 112). The constitutive formula is a useful mnemonic to get us to see what *must* lie at the core of institutional reality—the status-function and, in particular, the rights and obligations that it denotes. In this way, it seems that Searle is still interested in providing necessary conditions in line with atomist ambitions, in articulating logical structures, in doing *ontology*.

If the constitutive formula is a means to get us to see the status-function—what really matters—this helps reconcile the terminology of the *Construction*, which prompted the atomist reading of the constitutive formula in the first place, with Searle's "useful mnemonic" characterization in his discussion with Smith. Despite some slippage here and there, it is the status-function, and not the constitutive formula, that constitutes institutional reality's foundation. The constitutive formula serves as an ideal type to get the reader to see *this*, but the status-function itself is not an ideal type—it is a *description* of institutional reality, and it is unavoidable. Searle is, after all, trying to say something about *what* constitutes the "foundations of the social sciences": the status-function whose content is articulated in terms of rights and obligations.

There are problems with my characterization of X counts as Y in C as an ideal type. The constitutive formula is an ideal type or mnemonic, but not as a part of an *ongoing* or *indefinite* discursive activity. Its propaedeutic value is localized within a more traditional analytic project: the development of a foundational ontology, within which institutional reality can be located. *Any* account of institutional reality must implicate status-functions, in Searle's view. As we have seen, once that ontology has been properly understood, the constitutive formula can be kicked away.[10] So, in spite of what seemed like a promising dialogue with Smith,

Searle does not totally break free from the perniciousness of the atomist ideology.

In this way, we still have, therefore, a fundamental disagreement between Searle and Weber about what we can expect from our social scientific concepts. In spite of Searle's promising characterization of the constitutive formula, there remains an important difference in the way that Weber and Searle construe, not the process by which a discursive activity proceeds, but the ultimate aims and ends of such an activity. As far as Weber is concerned, this dynamic process of asking and answering questions, only to have a new set of questions arise, proceeds indefinitely.

> The stream of immeasurable events flows unendingly towards eternity. The cultural problems which move men form themselves ever new and in different colors, and the boundaries of that area in the infinite stream of concrete events which acquires meaning and significance for us ... are constantly subject to change. The intellectual contexts from which it is viewed and scientifically analyzed shift. The points of departure of the cultural sciences remain changeable throughout the limitless future as long as the Chinese ossification of intellectual life does not render mankind incapable of setting new questions to the eternally inexhaustible flow of life. A systematic science of culture, even only in the sense of a definitive, objectively valid, systematic fixation of the problems which it should treat, would be senseless in itself. Such an attempt could only produce a collection of numerous, specifically particularized, heterogeneous and disparate viewpoints....[11]

As old puzzles or "cultural problems" in the social sciences are solved, novel ones emerge in "ever new and in different colors". Moreover, Weber rejects the possibility of a Hempelian "ossification" that would result in "a definitive, objectively valid, systematic fixation of the problems which [the social sciences] should treat".

In another passage, Weber explicitly locates his characterization of the ideal type in a philosophical tradition that finds its roots in Kant:

> If one perceives the implications of the fundamental ideas of modern epistemology which ultimately derives from Kant; namely, that concepts are primarily analytical *instruments* for the intellectual mastery of empirical data *and can be only that*, the fact that precise genetic concepts are necessarily ideal types will not cause him to desist from constructing them.[12]

Our concepts are "*instruments*" or tools that draw our attention to various features of social reality, which is, *in itself*, infinitely complex. We can no doubt increase our understanding of this reality by clarifying and cross-referencing various ideal types, what Weber calls a "*process of synthesis*".[13] We should not, however, expect to provide canonical or unavoidable interpretations of it, even in the form of necessary conditions. It is in this way that Weber can assert that "there are sciences to which eternal youth is granted," within which we can expect the "transciency of *all* ideal types *but* also at the same time the inevitability of *new* ones".[14]

The variety of ways we were able to characterize money in the last chapter, and even the case of the constitutive formula, which Searle characterizes as ultimately superfluous, illustrates this transiency. But the status-function, as characterized by Searle, flouts Weber's conviction in suggesting that there are specifiable limits to what we can expect from our social scientific inquiries into institutional reality. There may be a variety of ways by which to characterize the significance of money. But Searle cautions us that those characterizations must ultimately be rendered in terms of the distribution of enablements and requirements; rights and obligations, then, are the bed rock of institutional analysis. The comparison of Searle's project with that of the atomist's is not unjustifiable.

What remains is to argue, against Searle, that the status-function itself, like the constitutive formula, is best understood as an ideal type. I will do this by showing that there are features of institutional reality that remain inaccessible to the status-function, but are clearly perceivable to the structural functionalist, and vice versa. I will conclude by suggesting that there are additional ideal types that illuminate our institutions in ways that are unavailable to either Searle or the structural functionalist.

Searle's functionalism

For Searle institutional facts or status-functions are agentive functions that implicate certain rights and obligations suggesting a kind of affinity between Searle and the structural functionalist. I will now explore the extent to which Searle does and does not endorse a variety of functionalism, laying groundwork for the formulation of a different model of institutional reality.

Recall Searle's taxonomy of functions. He distinguishes between nonagentive and agentive functions. Agentive functions, unlike nonagentive functions, exist in virtue of an agent's purposes and activities. A rock functions as a paperweight only in virtue of certain purposes that we may have; this is not the case for a heart, which pumps blood irrespective of our purposes. Agentive functions can be further subdivided: there are noninstitutional and institutional agentive functions, so that rocks that function as paperweights are importantly different from rocks that function as media of exchange (money). For now, I will identify institutional-agentive-functions with Searle's notion of the status-function.

Building on certain suggestive passages found in the *Construction*, I would like to recommend that institutional-agentive-functions can be further subdivided. Ian Hacking distinguishes "indifferent kinds" from "interactive kinds".[15] Indifferent kinds are brute objects that are not conscious or self-aware like quarks or green

pieces of paper. Interactive kinds, by contrast, are conscious and self-aware; people—agents—are generally interactive kinds, because they experience themselves in a certain way. Hacking argues that descriptions of interactive kinds are subject to a "looping effect" in a way that is not the case when we try to describe indifferent kinds. When a person is described as having Multiple Personality Disorder, that attribution may fundamentally *change* the object of description.[16] Interactive kinds are, as Hacking puts it, *moving targets*: they respond to their being characterized, changing the significance of both the object and the characterization itself.

I would now like to appropriate Hacking's distinction to discuss, not how different kinds of objects are or are not affected by the assertions we make about them, but how these different kinds of objects can refine our taxonomy of functions. Status-functions or institutional facts may be imposed on either indifferent or interactive kinds. Searle's principal example of an institutional fact is that of money, in which a status-function is imposed on an indifferent X, a green piece of paper. But Searle also talks about presidents, or we even might use Hacking's example of Multiple Personality Disorder, in which case a status-function is imposed on an interactive X, a person. Searle mostly seems to treat the cases of money and the presidency as analogous.

But we might distinguish between two kinds of status-function: *indifferent*-institutional-agentive-functions and *interactive*-institutional-agentive-functions.[17] While Searle does not explicitly subdivide the status-function, it is clear that he would not be unsympathetic to the distinction. Indeed, Searle asserts that interactive status-functions are "fundamental": "it is not the five dollar bill as an *object* that matters, but rather that the *possessor* of the five dollar bill now has a certain power that he or she did not otherwise have" (*CSR*, p. 97). By "power," Searle is gesturing at the normative component indicative of all (nonhonorific) status-functions—enablements and requirements, or rights and

obligations.[18] Interactive X's, then, are the sort of things that can be in accord with extant rights and obligations: this includes people, but more typically a person's behavior. Money, however, is not an interactive X because money cannot be remiss in the same way that a person, or a person's behavior, may be remiss.

Searle's remarks, here, suggest an alternative means by which to characterize this hierarchy of functions. Every function, agentive or otherwise, implies the possibility of *failure*. But indifferent functions fail in importantly different ways than agentive functions.

On Larry Wright's analysis of functions, X has function Y if Y is a consequence of X's being there and X is there because it does Y.[19] The second clause is necessary because not all consequences of X are functions of X. Pens *can function as* (or be good as) drumsticks, projectiles, and pointers but these are not normally their function; these are "accidental" in the sense that they do not explain X's being there. Wright's analysis concerns indifferent phenomena such as pens, rocks, and hearts.

I would now like to extend this discussion to the more complicated case of functions attributed to interactive X's. When we characterize interactive phenomena from an internal point of view, X's bringing about some goal is not sufficient for X's being good (or for X to count as Y). The *means* by which Y is brought about is important in a way that it is not when functions are attributed to indifferent X's.

Consider the case of President Bush's 2005 nomination of John Bolton as US envoy to the United Nations. Detractors were concerned that Bolton, an outspoken and even caustic critic of the UN, would have the effect of further isolating the US from its allies. Former ambassador to the UN, Jeane Kirkpatrick, observed of Bolton: "He loves to tussle. He may do diplomatic jobs for the US government, but John is not a diplomat".[20] Kirkpatrick's assessment of Bolton reveals how a function may be attributed to interactive phenomena differently than indifferent phenomena.

The possibility of failure for interactive agents cannot be characterized strictly in terms of dysfunction. The problem with someone who spouts ungrammatical sentences is not (only) that they might fail to be understood. Even if they are understood, the speaker is still *remiss*—they have, in some sense, violated an obligation to, say, make sure that subjects and verbs agree. Interactive agents are subject to standards, possibly in the form of rights or obligations, that govern *the way* in which some end is to be accomplished.

Returning to the Kirkpatrick quotation, there are two senses in which Bolton could be taken not to be a diplomat. First, if Bolton is being treated as an indifferent X, then Bolton is not a diplomat in the sense that he does not bring about Y, or the goals indicative of diplomacy. I think this is a misunderstanding of Kirkpatrick's sense, but nevertheless pedagogically interesting.

Second, if Bolton is being treated as an interactive agent, then if Bolton fails at being a diplomat, it is not (necessarily) because he fails to bring about some Y. Indeed, if the aims of diplomacy are in part to resolve political disputes nonmilitarily, Bolton may have a strong record of accomplishment on this front; in fact, he *does* tend to bring about Y, which is one of the reasons why he keeps getting "diplomatic jobs". Bolton's problem has rather to do with the means by which Y is brought about; he somehow fails to go about engaging in diplomacy in the right way. Somehow, Bolton's hawkish tenacity (X) contravenes the *means* by which Y should be accomplished.

With interactive agents, the *means* by which Y is brought about is an important part of what it is for Y to be brought about. Thus, Y has two functional components: the goal itself (Y_1) and the means (Y_2). If Bolton is to be a "good diplomat" it is not enough that Bolton succeeds in bringing about nonmilitary solutions to crises, but that he does this in ways that are collectively recognized as *appropriate*. Adapting Wright's definition of function to the special case of Y_2, here, we are looking at the means, Y_2, that tend to

bring about Y_1 (in our society), and contrasting it from *other means* that might bring about Y_1. That is, for interactive X's, X is not good or appropriate *just because* it brings about Y_1—it must also bring it about in the right way. In the first paragraph of *The Da Vinci Code*, author Dan Brown writes, "A voice spoke, chillingly close. 'Do not move.' " Linguist Geoffrey Pullum responds, noting that "A voice doesn't speak—a person speaks; a voice is what a person speaks with".[21] Pullum raises a question as to whether or not Brown's phrasing (Y_2) is the most appropriate means by which to convey the narrative (Y_1). The problem here is not that Brown failed to bring about Y_1. The question rather lies in whether or not Y_2 is an appropriate way for English speakers to bring about Y_1.

The portrait of what it is for an interactive X to be a Y is distinctly Aristotelian. On Wright's account of indifferent X's, "it is clear that the demonstration of the goal-directedness of something's behavior does not involve us at all in the discussion of the internal structure of that something".[22] Not so for the functions of interactive X's, where the means (Y_2) is constitutive of Y. Alasdair MacIntyre puts this point as follows: For indifferent X's,

> the means and the end can each be adequately characterized without reference to the other; and a number of quite different means may be employed to achieve one and the same end. But the exercise of the virtues is not in this sense *a* means to the end of the good of man. For what constitutes the good for man is a complete human life lived at its best, and the exercise of the virtues is a necessary and central part of such a life, and not a mere preparatory exercise to secure such a life.[23]

Aristotle might agree that, for persons, an interactive function Y is minimally constituted by Y_1 and Y_2; the means is constitutive of the X's end. MacIntyre also makes a substantive claim about the Y_2 by which Y_1 is brought about—it consists in virtue

or excellence. This is to anticipate a discussion I will take up in detail below. Suffice it to say that Searle and the Aristotelian disagree as to the makeup of Y_2, where, for Searle, Y_2 consists in the satisfaction of rights and obligations.

Searle affirms that it is the means (Y_2) by which an end is brought about that is constitutional to Y, and further, that this means principally consists of rights and obligations: "Because the Y content is imposed on the X element by collective acceptance, there must be some content to these collective acceptances (recognitions, beliefs, etc.); and I am suggesting that for a large class of cases[24] the content involves some conventional power mode in which the subject is related to some type of action or course of actions" (*CSR*, p. 104). The power modes are enablements and requirements, or rights and obligations. The disagreement between Searle and the Aristotelian, then, comes down to a disagreement between what behavior (X) should count as Y_2, where Y_2 is the appropriate means by which Y_1 is brought about. Searle thinks that constitutive rules, described in terms of the distribution of rights and obligations, are sufficient to specify the contours of an institution, whereas the Aristotelian looks towards excellence or virtue.

It is important to see why Searle does not include Y_1 as part of the Y's "content," contrary to the suggestion that interactive-institutional-agentive-functions contain *both* an end (Y_1) and means (Y_2). Searle does not want to say this because an X's being a diplomat should not be contingent on X's actually succeeding in bringing about Y_1; on one hand this is correct, but on the other hand this worry is misguided. It may be the case a given X may fail to bring about its goal, a fact which Searle's exclusion of Y_1 from the content of Y captures, but the idea that such X's might *always* fail is nonsensical. X can only be dysfunctional in virtue of the fact that typically it does tend to bring about Y_1. Not all hearts succeed in pumping blood, whereas no pinecones succeed in pumping blood. When it comes to Y_1, only the former rises to the level of being *dysfunctional*, or bad at pumping blood; a heart has

the function it does, unlike pinecones, because it already tends to do such things. In that way, so long as we are clear that dysfunction must be a genuine (but not systematic) possibility, there is no problem with including Y_1 in Y; indeed, it is highly misleading not to do so.

Extending Wright's analysis of functions, I have argued that for interactive X's the means by which Y is brought about is an important part of what Y is. This point can be rendered in a Heideggerian mode. John Haugeland points out that,

> not all norms are of the sort that entail responsibility—commitment and entitlement—on the part of that to which they apply. Thus, the parts of a complex organism or system can often be understood as having functional roles, defined by norms of performance. For instance, a heart is "supposed to" pump blood, much as a carburetor is "supposed to" mix fuel; these are their normal roles, and we may even call them "faulty" if they fail. But they are in no sense committed or obliged (still less entitled) to fill these roles; hence they cannot be found culpable, or blamed as irresponsible if they fail. The structure of the who—of accountable agents—is, by contrast, precisely that of entities who *are* responsible for what they do: they *can* be committed and entitled, and hence held to account for how they perform.[25]

A heart has a nonagentive function, whereas a carburetor has an agentive function, and yet both can be contrasted with interactive status-functions in the sense that "culpability" and "responsibility" can only be predicated of the latter. Carburetors cannot be remiss. That is to say, that while an interactive X may be treated as an indifferent X, failure for Dasein cannot be articulated exclusively in terms of equipmental breakdown. Furthermore, even when we confine our attention to the status-function (institutional-agentive-functions), interactive status-functions can be distinguished from indifferent status-functions in the same way. If an

indifferent X fails to function as a medium of exchange, it does so differently than an interactive X that fails to function as a president. Only the latter can be remiss, or be found culpable or responsible. This is because the *means* by which Y is brought about is important for interactive X's. Searle's propensity to treat money and presidents as similar is problematic, not because there are not similarities between the two kinds of object, but because the conflation often obscures the important differences between these kinds of status-function. Fortunately, Searle is nevertheless clear that interactive functions—objects to which we can attribute enablements and requirements—are the primary concern in the *Construction*. It is these functions that Searle most properly identifies with the status-function.

Having elaborated on Searle's functional taxonomy, compare Searle's mode of inquiry in the *Construction* to that of the structural functionalist. Functionalists recommend that social systems are importantly analogous to biological systems. Talcott Parsons argues that "the same principles [drawn from biology] are equally relevant to social systems".[26] The tradition finds its roots in Auguste Comte, Herbert Spencer, Emile Durkheim, and was subsequently developed by sociologist Robert Merton and anthropologists Bronislaw Malinowski and A.R. Radcliffe-Brown. Both biological and social entities are systems that regulate or maintain themselves relative to changes in their environment; biological and social systems are homeostatic in that they have mechanisms that allow them to sustain themselves against external forces that threaten to disintegrate their integrity. The functionalist assumes that a system has a kind of unity, and the parts of the system tend toward integration, contributing to its equilibrium and fitness.

The social scientist assumes that the actions of participants contribute to the maintenance of the social system, just as the evolutionary biologist assumes that most phenotypes contribute to the fitness of a given species. The biology analogy gives an important indication of the sort of functions with which the structural

functionalist is especially concerned. Biological systems, including cells and their constituent features, organs, and organisms within a given environment, are *indifferent* entities. As we have seen, interactive entities can be treated as indifferent entities, but in doing so, the way in which we describe cases of failure must likewise shift. The structural functionalist, I want to suggest, is principally concerned with interactive entities but then explicates their behavior by treating them as indifferent entities. The functionalist is concerned with *all* the agentive functions listed in Searle's hierarchy except for interactive-status-functions.

Many societies make use of conflict-solving rituals to resolve disputes between two or more people. The Bedouin Arabs employ the practice of *Bisha*, or ordeal by fire.[27] In case of an otherwise irresolvable dispute between two parties (perhaps a husband suspects his wife of infidelity), they meet before the *Mobasha*, a disinterested third party who is to administer the Bisha. Both parties agree to accept the results of the Bisha. Before witnesses, the Mobasha heats a metal tool until it glows hot. After the iron is inspected by the witness, the accused is then instructed to lick the metal tool. Afterward the Mobasha inspects the tongue. If it is found to be harmed the accused is declared guilty; if the tongue is otherwise unscathed, the accused is deemed innocent.

The above analysis of the Bisha ritual treats the Bedouin Arab community as a social system: the behavior (X) of the ritual's participants tends to bring about conflict-resolution, and more generally, social integration (Y). The ritual can be dysfunctional if it fails to resolve conflicts and so fails to bring about social integration. So, the functionalist treats the participants of the Bisha as indifferent X's. The functionalist adopts what Searle calls an external point of view. Searle underscores this distinction in the following passage:

> Another formal feature to note is that the usual distinction between internal and the external points of view applies to

> institutional facts. In this book we are interested primarily in the internal point of view, because it is only from the internal point of view of the participants that the institution can exist at all (*CSR*, p. 98).

Searle is not interested in the external perspective of the structural functionalist, but the standpoints of participants acting inside the world; in particular, he is interested in the standpoint of semi-self-conscious participants, who often, but do not always, understand and meet their obligations. In other words, if Searle is interested in X's function, he is interested both in its end (Y_1) and the *means* (Y_2) by which Y_1 is brought about. Since the means is articulated in terms of rights and obligations, iterations of the status-function would depict the nested structure of participants' rights and obligations. When agents are treated as indifferent entities, we are adopting the external point of view; this is the point of view adopted by the functionalist. When agents are treated as interactive entities or participants, we are adopting the internal point of view. This implies the possibility, not just of *dysfunction* (the failure to bring about Y_1), but *remission* (the failure to bring it about in the right way, Y_2). Moreover, Searle claims that while our institutions can be rendered externally, the internal point of view represented by the interactive-status-function is a precondition for the existence of institutionality. Thus, a functionalist depiction of institutional reality is derived from the participant's point of view.

This is easy to see if we consider functionalist attempts to construe the rights and obligations indicative of interactive-status-functions. A functionalist might contend that structural-functionalism is a broader, more inclusive account of social reality than that of Searle's because the former can account for rights and obligations in addition to features of society that make no reference to such norms. Since Searle is strictly concerned with rights and obligations, it seems the functionalist accounts for at least those features covered by Searle's model.

The functionalist does not deny that norms are an important feature of society. One of the functionalist components of our society are norms and values, perhaps in the form of rights and obligations, that mediate our interactions. These *can* be treated under the functionalist framework, so that these norms are brought under the form of teleological explanation—"A tends to bring about G". The behavior of participants of the Bisha ceremony (A) tends towards conflict-resolution and even social stability (G) *by means of* the obligations imposed on participants (i.e., the accuser may be obligated to abandon the allegation). In that way, the functionalist might argue, structural-functionalist analysis includes the very features of institutional reality with which the status-function is principally concerned. The structural-functionalist analysis, thus, seems more general than that of Searle's.

While it is true that the functionalist need not abandon talk of rights and obligations in depicting society, I now want to argue that the functionalist characterizes these terms in such a way that mars our normal understanding of them. When placed within the form of teleological explanation, rights and obligations are treated here as a means, mechanism, or instrument by which G is brought about. And as a means, there is an important sense in which the evocation of such norms is largely *incidental*, or even *irrelevant*, to the functionalist's question.

To see this I would like to survey a feature of Wright's analysis of teleological explanation. In explaining the behavior of goal-directed objects, such as guided missiles, Wright argues that drawing our attention to the underlying mechanisms or means by which the feedback loop is obtained is often a non sequitur:

> In the case of consciously designed mechanisms, the teleological characterization of their behavior ... will be the design criterion. The explanation for there being any machine there at all, as well as much of its detail will be in terms [of its tendency to bring about G]. The mechanism is explanatorily incidental.

> What we wanted was something (it did not really matter what) that would behave in a way that would tend to produce G, at least in certain circumstances. That the mechanisms turned out to have certain *mechanical* details is interesting and explanatory only in certain narrowly circumscribed contexts: for example, where we are incredulous at its possibility, or where mechanical defects affect its performance.[28]

If the question that the functionalist asks is, "Why does this tribe perform Bisha rituals (A)?," the answer will involve stating some G that explains A's being there. For the functionalist, the particular mechanism (i.e., a rule structure of rights and obligations) by which A brings about G *may* be interesting and explanatory only in certain cases—for example, if it is unclear how G is the result of A. But if, as in the usual case, the means is not puzzling, a discussion of the mechanism by which A occurs is, at best, incidental, and at worst, highly distracting.

But showing that the evocation of norms is incidental, given the question the functionalist is asking, does not yet show that the sense in which norms are evoked—to the extent that they are evoked—robs the notion of a critical aspect. To see this, we only have to point out that the functionalist is none the worse for evoking alternative mechanisms, if a question about these mechanisms does happen to arise: A could tend to produce G, not by means of rights and obligations, but by some alternative mechanistic instrumentality—overt compulsion through fear, force, or intimidation, covert coercion, neural implants controlled by mad scientists, or whatnot.[29] Given the sort of questions the functionalist is answering, the two explanations are intersubstitutable. Just as a natural or artificial heart are functionally equivalent, even though their underlying mechanisms are quite different, either norms or neural implants could answer any question the functionalist might have about the relationship between A and G. But, to insist these mechanisms are importantly different is to ask a *different*

question than the one asked by the functionalist. For those of us, including Searle, who are uncomfortable with the consequences of intersubstitutability, this is an important indication that the functionalist has radically reconstrued the normal significance of rights and obligations. The functionalist has not denied the possibility of rights and obligations, but has treated them as mechanisms by which a given G is brought about. If this is unsatisfactory, it suggests that functionalism has only highlighted certain aspects of social reality, even as it obscures others.

Building upon Searle's own distinctions, there are aspects of institutional reality that cannot be satisfactorily articulated under a functionalist regime. My discussion of different kinds of failure (remission vs. dysfunction) as well as the intersubstitutability argument were intended to show that the functionalist's conception of normativity is too thin to adequately account for many institutional phenomena.

Institutional phenomena that resist articulation under the status-function

There are manifest institutional phenomena that the status-function strains to characterize. To see this I would like to look at two instances of normativity.

First case: For Searle, an archetypical case of an institutional fact might be something like a Japanese tea ceremony (*Sado*). The ceremony *consists* of a number of prescribed rituals and movements, both on the part of the host and the guest. For example, participants are generally expected to keep conversation to a minimum during the one to five hour ceremony. The guest of honor is the first to receive the bowl of tea from the host. Upon receiving the tea, the guest is obliged not to drink from the front of the bowl. Following the first sip, a prescribed phrase is uttered, the rim is wiped with a cloth, and the bowl is passed to the next guest. If a critical number of these constitutive rules are violated, the

activity no longer counts as a *Sado*. But, if this threshold has not been reached, then participants who violate these rules are considered remiss.

Second case: normativity of a different kind can be found in the animal kingdom. Consider the case of cooperation observed in rats.[30] Here, a Plexiglas partition separates the two rats. On each side of the partition are two buttons and a food tray. On one side, the buttons would light up indicating the order in which the buttons needed to be pressed to eventuate the opening of the food tray. But the tray would only open if the "follower" rat, whose buttons were not lighted, mirrored the behavior of the "leader" rat.

The rats must succeed in cooperating in order to gain access to the food. It would be inappropriate to characterize the follower rat's behavior in terms of an *obligation* to mirror the leader rat's cue, or to characterize the activity itself as an *institution*. If the rats fail to open the tray it is not because one of the rats is remiss. However, the normativity involved is easily brought under a functionalist analysis, so that we might speak of functional and dysfunctional behavior.

I now want to suggest that there is human activity that straddles these two portraits of normativity in telling ways. I have in mind cases such as the language game of *Philosophical Investigations* §2. The activity of the builder and the assistant bears a striking resemblance to that of the leader and follower rat. We would effortlessly characterize this activity as "institutional," but I will argue that the ascription of rights and obligations to the participants is inapt. Certainly, we can *say* that the assistant has an obligation to bring the block upon hearing the builder say the word "block". But it is not clear what is barring the further ascription of obligations or status-functions to the rats. Moreover, the ascriptions of rights and responsibilities to the builder and assistant overarticulates the normativity involved. In other words, I will argue, contrary to what Searle's status-function would have us expect, that

the building activity *is* institutional, but does *not* involve the ascription of rights and responsibilities to the participants.

To see this, consider Searle's attempt to distinguish mere cooperative behavior from those activities that can be brought under the status-function:

> Animals running in a pack can have all the consciousness and collective intentionality they need. They can even have hierarchies and a dominant male; they can cooperate in the hunt, share their food, and even have pair bonding. But they cannot have marriages, property, or money. Why not? Because all these create institutional forms of powers, rights, obligations, duties, etc., and it is characteristic of such phenomena that they create reasons for action that are independent of what you or I or anyone else is otherwise inclined to do. Suppose I train my dog to chase dollar bills and bring them back to me in return for food. He still is not buying the good and the bills are not money to him. Why not? Because he cannot represent to himself the relevant deontic phenomena. He might be able to think "If I give him this he will give me that food." But he cannot think, for example, now I have the *right to buy* things and when someone else has this, he will also have the right to buy things (*CSR*, p. 70).

Searle's case of the dog is helpfully analogous to the position of the follower rat and the building assistant. Searle is explicit here: he does not want to ascribe rights and obligations to the dog, and so would not characterize the dog's activity as institutional. He would likewise resist describing the rat's cooperative behavior under the status-function.

Here Searle seems to be putting forth a criterion for whether or not an activity is institutional: minimally, the participants must, at least in principle, be able to represent to themselves the relevant deontic phenomena. This includes the requirement to be able to represent or articulate both the function of the activities in which

they are engaged, as well as the normative component indicative of a self-reflexive interactive agent. This representation-requirement finds its roots in Searle's discussion of agentive functions, of which the status-function is a variety: "for all cases of agentive function, someone must be capable of understanding what the thing is for, or the function could never be assigned. At least some of the participants in the system of exchange must understand, consciously or unconsciously, that money is to buy things with, screwdrivers are for driving screws, and so forth" (*CSR*, p. 22). Searle's dog presumably has the capacity to represent certain agentive functions (pieces of paper are for getting food), but he is quite unable to represent the interactive-institutional-agentive-functions (status-functions) constitutive of institutional reality. This is because, while he might be able to represent the end (Y_1) of X (to obtain food), he cannot represent the appropriate means, Y_2 (by way of certain rights and obligations). In other words, dogs are indifferent X's, and so cannot be brought under the status-function.

But with respect to the ability to represent the deontology, it becomes difficult to see why the activity of the builder and the assistant is importantly different from that of the dog or the rat. Stipulated constraints on their linguistic and conceptual resources guarantee that these agents would not be able to articulate the rights and obligations we are tempted to attribute to them: "We could imagine that the language of §2 was the *whole* language of A and B; even the whole language of the tribe".[31] Accordingly, the building assistant simply would not be able to think, consciously or unconsciously, "now I have an *obligation* to bring forth this block and when someone else hears this sound, he will also have the obligation to do the same." According to Searle, Wittgenstein's language game is a cooperative activity, but lacks the deontology necessary for institutionality.

Perhaps if Searle wanted to secure this activity's place in institutional reality, he might claim that the builder and his assistant

could articulate or codify the unconscious rules in the sense that they are not barred from doing so on account of their neurophysiology alone (unlike rats or dogs). But this response seems ad hoc, an indication that we have lost track of the question we set out to answer: what distinguishes institutional phenomena from merely coordinated behavior? What is it about unrealized neurophysiological potential that moves us to a clearer understanding of *this* difference, if it does not manifest itself in the activities themselves? And if some difference does exhibit itself in the activity, wouldn't *that* be a better candidate for articulating the difference in question? I will return to this below.

Perhaps, then, Searle might petition to exclude the building activity from the sphere of institutionality. If it is the case that the activity of the builder and his assistant cannot be brought under the status-function, then Searle might concede the result and remind us that not all cooperative behavior can be brought under the status-function. After all, not all social facts are institutional facts. Wittgenstein's language game is anyway highly artificial, and so tells us little about the kinds of archetypical institutions (e.g., Japanese tea ceremonies, the presidency) Searle is most interested in explicating.

But this strategy seems mistaken. Certainly, Wittgenstein took this activity to be paradigmatic of our institutions, rather than the derivative case; the Japanese tea ceremony is an institution that has the required deontology, according to Searle, but there is good reason to think that *it* is in some ways exceptional. Even by Searle's own terms, I will argue, it would be a mistake to exclude the language game of the builder and the assistant from the sphere of institutionality, even if the behavior cannot be described in terms of rights and obligations.

Before proceeding, I would like to briefly discuss the consequences of such an argument. Obviously, the functionalist has little difficulty in characterizing the activity of the builder and his assistant in terms of its tendency to bring about Y. The

functionalist is able to characterize institutional phenomena that is inaccessible to the status-function. Of course, the functionalist cannot say exactly what distinguishes the building activity from that of the rats or dog. But the functionalist, unlike Searle, never sets out to give an account that distinguishes institutions from other kinds of cooperative behavior. Indeed, the functionalist's driving metaphor is that social and institutional systems are analogous to biological systems. Still, I will argue that if an attempt to assimilate the building case into a rubric of rights and obligations *overarticulates* Wittgenstein's institution, the functionalist *underarticulates* the same institution by bringing the building case under the biological metaphor; as we have seen above, the failure of self-reflective agents cannot be fully articulated in terms of dysfunction or breakdown. If there is no problem in characterizing the activity of the builder and the assistant in terms of the tendency to bring about Y, this should not imply that there is not more to say. The difficulty is that the status-function does not quite say it—the builder and assistant are not constrained by a potentially explicit set of rights and obligations. Consequentially, I will suggest that we need *another* model of normativity, in addition to the status-function and functionalism; there must be another way to characterize failure besides dysfunction and remission. I will return to this point in the next section. In the meantime, it is necessary to further buttress the claim that some institutions, like that of the builder and his assistant, resist articulation under the status-function.

A primary difficulty with Searle's dog example is that it makes the ascription of the status-function contingent on whether or not the actor would be able to review the rights and obligations in question. This would exclude the rats' and dog's activity in the right sort of way, but also problematically bars Wittgenstein's building activity from the sphere of institutionality.

Peter Winch, in *The Idea of a Social Science and its Relation to Philosophy*, disagrees with Searle's subjunctive as a criterion of

institutionality. Consider his response to Michael Oakeshott's distinction between habitual and meaningful behavior.

> Oakeshott appears to think that the dividing line between behavior which is habitual and that which is rule-governed depends on whether or not a rule is *consciously* applied. In opposition to this I want to say that the test of whether a man's actions are the application of a rule is not whether he can *formulate* it but whether it makes sense to distinguish between a right and a wrong way of doing things in connection with what he does. Where that makes sense, then it must also make sense to say that he is applying a criterion in what he does even though he does not, and perhaps cannot, formulate that criterion.[32]

Winch is right to draw our attention to the fact that rule-following is possible, not only in the case where the rule is consciously applied, but even in cases where the participants *cannot* formulate the rule. Analogously, I want to suggest that the status-function, which requires that an agent *could* articulate the otherwise unconscious rules, offers too narrow a constraint on what sort of activities are to count as institutional. My aim, here, is to show that there are institutional phenomena that are accessible to the functionalists but not to Searle, and further, that there is a gap that needs to be filled by a third account of normativity.

Winch's point can be put as follows: *one* way to manifest the requisite understanding of the institution is to be able to talk about it, perhaps in terms of rights and obligations. But an agent's inability to talk or think about the activity in question is not, alone, a reason to suggest that he or she is not following rules (constitutive or otherwise). Other manifestations of institutional understanding are sufficient. Winch cites Wittgenstein's case of the student who is instructed to count by twos and does so correctly until 1000, at which point he continues on abnormally—1004, 1008, 1012, etc. One way to demonstrate competence is to be

able to formulate a rule: "n + 2". But this is hardly the typical way of demonstrating mastery. Someone who could count by twos in a variety of situations (starting from 0, 1000, −10, 90,000, etc.), but who was unable to formulate the rule, would be judged no less competent at the activity. In the extreme case there is the idiot savant who can, when given the date, name the day of the week over a span of 8000 weeks, but who may be quite unable to say which rule was being followed.

Wittgenstein and Winch might agree that, when it comes to many of our activities and institutions, we are all idiot savants. Mastery of formal grammar is probably an indication of linguistic competence, but it is hardly necessary. If someone was quite unable to wrap their head around basic grammatical concepts, but could nevertheless deploy well-formed sentences, we would not doubt this person's linguistic mastery; more tellingly, we might agree that if this person has a problem, it is with quite a different skill-set. Likewise is the case with other of our activities and institutions, many of which resist rendering under the status-function. For these I would like to turn back to Searle's own text:

> A test for the presence of genuine institutional facts is whether or not we could codify the rules explicitly. In the case of many institutional facts, such as property, marriage, and money, these indeed have been codified into explicit laws. Others, such as friendship, dates, and cocktail parties, are not so codified, but they could be. . . . Such institutional patterns could be codified if it mattered tremendously whether or not something was really a cocktail party or only a tea party. If the rights and duties of friendship suddenly became a matter of some grave legal or moral question, then we might imagine these informal institutions becoming codified explicitly, though of course explicit codification has its price. It deprives us of the flexibility, spontaneity, and informality that the practice has in its uncodified form (*CSR*, p. 88).

While Searle appears to treat the concluding remark as a casual aside, it contains an important qualification concerning the very possibility of articulation and codification. Here, Searle asserts that all institutional facts "could be" codified into patterns of rights and obligations, and yet in the same paragraph he rightly observes that an explicitly codified institution *differs* from its uncodified antecedent: it lacks "the flexibility, spontaneity, and informality that the practice has in its uncodified form".

Along these lines, Norbert Elias, in *The Civilizing Process*, wonders about the status of Erasmus' widely circulated treatise on table manners, *De Civilitate Morum Puerilim* (On Civility in Boys). Its publication, Elias argues, marks a "considerable" shift in behavioral change. But this change "did not take place, of course, in such a way that one ideal of good behavior was suddenly opposed by another radically different from it".[33] It was widely understood that meat was only eaten by hand (forks were communal, and were for taking meat from the main serving dish), but it was only to be handled by three fingers according to Erasmus. It is not correct to lick greasy fingers or wipe them on your garments. Moreover, "To dip bread you have bitten into the sauce is to behave like a peasant, and it shows little elegance to remove chewed food from the mouth and put it back on the *quadra*. If you cannot swallow a piece of food, turn round discreetly and throw it somewhere".[34] The important thing to see here is in codifying an institution we rarely just describe that institution, but we often end up *changing* it.[35]

But this should be deeply troubling to Searle: he requires that otherwise unconscious rights and obligations underlying institutional facts *could be* articulated by the participants, which is why Searle's dog cannot be said to buy things in spite of its behavior. If codification changes the institution, then uncodified institutions such as friendships, dates, cocktail parties, and table etiquette prior to Erasmus' publication would appear to be, like

Wittgenstein's language game, protoinstitutional. They resemble, but fall short of, Searle's legalized exemplars.

If Winch is correct, however, the participant's inability to successfully codify or articulate a cocktail party should not bar its admission into the sphere of the institutional. Moreover, the flexibility, spontaneity, and informality manifest by the participants of the cocktail party is a much more telling indication of the presence of a genuine institutional fact than the possibility of their being able to *talk* or even think about the activity in terms of rights and obligations. And it is precisely *this* mastery that separates the activity of the builder and his assistant from that of the rats or Searle's dog.[36] That the builder and the assistant are following a rule in a way that the animals are not, is evidenced by their behavior when circumstances change: if there are no blocks available, upon hearing "block" it is probably not the case that the assistant will mechanically go through the motions of carrying an invisible block to the builder. Maybe the builder occasionally makes mistakes, requesting a block when he needs a slab, and the assistant is able to correct for the error. Perhaps, even, the assistant might one day anticipate what the builder needs before he says it. All of these spontaneous *deviations* from the paradigm case are indications of understanding that the follower rat or Searle's dog could not exhibit without further habituation.

If the participants of an activity, such as the host of a Japanese tea ceremony or newlyweds, cite or could cite associated rights and obligations, the activity is almost certainly an institution. But, I have argued, the ability to articulate—even in principle—cannot be a necessary condition of institutionality. Here, the net is cast too narrowly and we exclude not only the activities of the builder and assistant but presumably friendships, dates, and cocktail parties; we can delimit the scope of a given institution by looking at associated rights and obligations but doing so excludes phenomena that we would normally take to be institutional.

Moreover, I have argued that it is in the spontaneous ability to go on together in novel situations, or be puzzled together in the case of breakdown, that institutionality most vigorously demonstrates itself. The ability to talk about an activity, however, is an often unreliable manifestation of competence. Requiring that the participant be able to talk about the activity is only one response to such situations, and one that presents its own unique difficulties; talking about a skill is itself a skill. A better indication of self-reflexive mastery is the ability to go on together in novel situations. Searle is right to say that institutions imply a "normative component" but I will argue that something like Aristotelian expertise or competence provides, in many cases, a better way to articulate this normativity without being overly exclusionary. Competence of this kind will allow us to include the activity of the builder within institutional reality.

In spite of Searle's discussion of agentive functions and the dog example, sometimes he recognizes difficulties with the articulability-requirement: sometimes "people who are participating in the institution are not conscious of the rules and do not appear to be following them, either consciously or unconsciously" (*CSR*, p. 127, 137). As a proxy for conscious or unconscious intentionality, Searle falls back on his notion of the Background. But we have seen that the Background is a troubled concept: it is an artifact which is nonintentional (a category of neurophysiological causation) but nevertheless, as a proxy for intentionality, must account for the deontology indicative of constitutive rules. In other words, the Background appears to be *whatever* satisfies Searle's seemingly incompatible requirements. Perhaps in recognition of this problem Searle again shifts his characterization of the Background: the Background is "a set of dispositions that are sensitive to the rule structure" (*CSR*, p. 145). But it was precisely this rule structure that could not be explicated in terms of the conscious and unconscious beliefs of participants, that prompted Searle to evoke the Background in the first place: "I will show how some, though

not all, of the intentionalistic apparatus can be explained in terms of, and ultimately eliminated in favor of, what I have elsewhere called the 'Background' of capacities, abilities, tendencies, and dispositions" (*CSR*, p. 5) Searle's appeal to a rule structure, which is now accounted for neither by conscious or unconscious intentional attitudes, nor the Background, begs the very question the *Construction* purports to answer: how are institutional facts possible? How does institutional reality exist in virtue of a world made up of physical particles and conscious beings? To further include a rule structure in his fundamental ontology blunts the force of the question Searle takes himself to be answering.

On the other hand, if the status-function is to be construed as an ideal type such a breakdown is not an indication that Searle has got it *wrong*, but only that we reached the limits of applicability of this particular ideal type. As Searle has done with his appeal to the Background, the inquirer always has the option of stretching the ideal type when faced with an anomaly. But, such maneuvers can come at a price. Sometimes a counterexample (in this case, difficulties with the articulability-requirement) is better treated as an indication of the limitations of the scope of a particular ideal type. There may be a number of reasons that will lead us to resist modifying the ideal type any further. For instance, part of what may be valuable about a given paradigm is what Wittgenstein calls its surveyability or perspicuity.[37] One way of putting Wittgenstein's point is that increased understanding does not always follow from the unification of two competing ideal types, especially if the resulting ideal type is somehow difficult to grasp. Recall the neoclassicalist-behavioralist debate. Both the neoclassicalist and the behavioralist might be able to characterize tipping behavior; one looks to deliberate cost-benefit calculations, whereas the other looks to blind and seemingly irrational instinctive responses. We could unify the models by insisting that all such impulses *must* be, by some trick of evolution, ultimately rational. But, it is not clear that an ideal type which manages to combine

the behavioralist and neoclassical insights would necessarily be more helpful than if the two models were applied to this particular case separately; unification seems to blur the helpful distinction between rational and instinctive responses.

In Searle's case, attempts to repair difficulties with a model that begins with overly articulate codified exemplars strains our conception of institutional phenomena, which seem neither articulable nor codifiable. We can stretch the ideal type to deal with such cases, but in doing so we are burdened with huge networks of unconscious rules, or an entity as tortured as the Background, or a rejection of the very question which motivated the project in the first place. In other words, while ideal types can be indefinitely modified to account for aberrant phenomena, the result can be obfuscatory, even unintelligible. What is needed, in this case, is an *additional* ideal type to complement the useful mnemonic articulated in the *Construction*. What is needed is a thicker model of normativity than offered by the functionalist (we are still interested in the point of view of participants) but thinner than that implied by the status-function. Is there a middle ground between the poles of dysfunction and remission?

Aristotelian virtue: an alternative account of the participant's point of view

What is needed is an *additional* model of normativity—another way to describe the possibility of failure—besides that provided by Searle or the functionalist.

Searle's discussion of codification gives an explanation for when explicit codification is necessary—when "it mattered tremendously," such as in the case of some grave moral or legal question. This intriguing observation gives us some sense as to when Searle's status-function is most aptly applied to institutional phenomena, and conversely, when another model of normativity might be more appropriate. To help germinate the seed contained in

Searle's observation, I would like to turn to an unlikely source: Alasdair MacIntyre's explication of Aristotelian virtue in his celebrated book, *After Virtue.*

Searle articulates the participant's point of view exclusively in terms of rights and obligations, enablements and requirements. But we are not necessarily driven to Searle's rendering of the within. MacIntyre wonders whether talk of *rights* is the only way available to characterize the normative aspect:

> [T]he possession of rights ... presuppose ... the existence of a socially established set of rules. Such sets of rules only come into existence at particular historical periods under particular social circumstances. They are in no way universal features of the human condition. ... [T]hose forms of human behavior which presuppose notions of some ground to entitlement, such as the notion of a right, always have a highly local and socially specific character, and that the existence of particular types of social institution or practice is a necessary condition for the notion of a claim to the possession of a right being an intelligible type of human performance. (As a matter of historical fact such types of social institution or practice have not existed universally in human societies.) Lacking any such social form, the making of a claim to a right would be like presenting a check or payment in a social order that lacked the institution of money.[38]

In this passage, MacIntyre has *inverted* Searle's view that institutions *presuppose* the distribution of rights and obligations. Rather than *constituting* our institution of money, as Searle claims, MacIntyre maintains rights and obligations are *among* the "types of social institution or practice," such as money. The normative, internal aspect remains—MacIntyre is obviously not arguing that other pre-Enlightenment human societies somehow got along without rules and norms. He is only arguing that these

rules can be characterized in a way that need not refer to enablements and requirements, and that such terms impose a tacit and highly contingent portrait of how participants comport themselves to one another.

MacIntyre helpfully goes on to discuss two varieties of normativity, which may help us see to which aspects of institutional reality Searle's status-function most aptly applies. MacIntyre distinguishes between the Aristotelian virtues on the one hand and a "morality of law" on the other. As a first approximation, virtues are skills—qualities of mind or character—that bring about a shared end, most generally characterized as the good. The *phronimos* or expert has the skill to both discriminate and respond appropriately to a variety of unique situations. A morality of law, however, is primarily a set of prohibitions on injurious actions, ones that intolerably undermine the possibility of the good (murder, theft, etc.). Only the latter may be readily explicated in terms of the status-function—obligations or requirements. Of these two models of normativity, MacIntyre writes:

> The need for *both* these types of practice arises from the fact that an individual member of such a community could fail in his role as a member of that community in two quite different ways. He could on the one hand simply fail to be good enough.... Conversely to fail the community by committing an offence against the law is *not* simply to fail by not being good enough. It is to fail in a quite different way.[39]

Someone who fails to meet articulated or even unarticulated obligations fails in a different way than someone who fails to act virtuously or excellently. The activity of the builder and his assistant, as well as attendees of a cocktail party, is best characterized under the latter model. These are not cases where "grave moral and legal questions" threaten the integrity of the institution, and so the channeling that comes with explicit or inexplicit

prohibitions, requirements, and obligations are not necessary. Rather, participants exhibit the plasticity, spontaneity, and informality indicative of cocktail party attendees and, more generally, Aristotle's conception of the good (i.e., model of normativity).

Aristotle compares the *phronimos* to experts of craft—a medical practitioner, for example. Dreyfus, following Aristotle's lead, compares the *phronimos* to the skilled expert who spontaneously responds to changing situations appropriately and effortlessly. In one sense, the *phronimos* cannot be said to follow rules,[40] in that the behavior cannot be accounted for in terms of the instructions the novice may have received when introduced to the activity. And thus, Dreyfus writes, "according to Aristotle, since there are no rules that dictate that what the *phronimos* does is the correct thing to do in that *type* of situation the *phronimos*, like any expert, cannot explain why he did what he did".[41] MacIntyre puts the same point as follows: "the exercise of the virtues requires a capacity to judge and to do the right thing in the right place at the right time in the right way. The exercise of such judgment is not a routinizable application of rules".[42] Something like *competence*, rather than rule-following, is the dominant model by which Aristotle would render the normative component of institutional reality.

If we have difficulty articulating the institution of a cocktail party in anything like the detail reflective of the behavior of its most competent participants, it is probably because institutional mastery is importantly analogous to the mastery we exhibit in driving a car or speaking a language. Construing this mastery in terms of rule-following does capture the sense in which it is possible to get it wrong, but we should not conclude from this that a *phronimos* is actually following rules, much less able to articulate those rules. Here, goodness does not trace the dictates of rights and obligations, but of expertise: two different judges, for example, may both exemplify the institution, and yet respond to the same case in different ways. The same may be said for Wittgenstein's assistants or the attendees of cocktail parties. Expertise is

not always determinable by the attribution of particular rights and obligations, although third parties often easily recognize failure even in the case where the actors have not violated the status-function.

In discussing Searle's text, of course we are not talking about the good life as a whole, as Aristotle does, but what it is to be a diplomat, English speaker, or cocktail party guest. Nevertheless, Aristotle's *moral* theory is particularly well suited for a discussion about the kind of normativity involved in institutional reality, because it is not clear that the competence necessary to navigate our institutions can be sharply distinguished from moral or even prudential considerations.[43] While Aristotle may be talking about the good life in general, that life is necessarily embodied in our institutions.

To underscore the difference between Searle's model of normativity and that articulated by Aristotle, I would like to return to Searle's concluding recapitulation of the main body of the *Construction*. Searle writes,

> that where human institutions are concerned, we accept a socially created normative component. We accept that there is something wrong with the person who when the baseball is pitched at him simply eats it; something wrong with the person who doesn't recognize any reason to do something after he has made a promise to do it; something wrong with the person who goes around spouting ungrammatical sentences. And all these cases involve something wrong in a way that is different from the way there is something wrong with the man who stumbles when he walks; that is, there is a socially created normative component in the institutional structure, and this is accounted for only by the fact that the institutional structure is a structure of rules, and the actual rules that we specify in describing the institution will determine those aspects under which the system is normative. It is precisely because of the rule that making a promise counts as undertaking an obligation that we recognize

> that certain kinds of behavior within the institution of promising are acceptable and other kinds are remiss (*CSR*, pp. 146–7).

A man who stumbles when he walks behaves dysfunctionally, and so fails in a different way than someone who is remiss. Human institutions, for Searle, imply the possibility of remission, and so are left underdescribed by the functionalist. There is little doubt that some institutions are effortlessly characterized by the status-function, but this is only one type of case. Other institutions may be awkwardly crammed into the frame provided by Searle's model, but we are left with the sense that in doing so, Searle is illicitly conflating importantly different phenomena. Upon reviewing the three examples of human institutions listed in the above passage the reader is left with a sense of uneasiness. Searle is right to suggest that these cases of failure are not simply cases of dysfunction, but is eating the ball during the baseball game the same kind of mistake as reneging on a promise or spouting ungrammatical sentences? That is, is a violation of obligations or requirements the best way to describe what is wrong with the baseball player?

The problem with the baseball player is not that he has violated a constitutive rule of baseball. If a pitched ball hits him in the face, whether or not he proceeds to eat it, these rules only dictate that he walks to first base. As he is required to walk, only if he decides not to do so might he be called remiss. The player who eats the ball is not remiss, but rather, unskillful in an Aristotelian mode. Another kind of normativity, then, may assert itself on an institution's participants: they can, not just violate an obligation or prohibition, as MacIntyre puts it, but *be defective in being good.* Indeed, if I am the assistant in Wittgenstein's language-game, such a defect might manifest itself in precisely my bringing forth the block *every* time I hear the sound "block". My problem, here, is not that I have violated a requirement.

Aristotle's model of normativity seems to characterize the very activities that Searle's model of normativity struggles to capture:

where Searle has in mind codified institutions such as money, marriage, and the Japanese tea ceremony as exemplars, Aristotle, Wittgenstein, and Dreyfus look to friendships, dates, and cocktail parties as exemplary of institutional reality. And MacIntyre is correct—ultimately *both* models, along with a discussion of how they relate to one other, may help round out our portrait of the institutional. Weber's infinite multiplicity drives us to expect this much. Institutions often do carry explicit and inexplicit prohibitions, obligations, and requirements, and these are best articulated by way of Searle's status-function.

While Searle's status-function remains the principal focus of this manuscript, I would like briefly to survey Margaret Gilbert's account of walking or traveling together. In addition to illustrating what institutional reality looks like conceived in terms of the *phronimos*, I am also going to use these cases to further develop the Aristotelian ideal type and show how it cannot be collapsed into the status-function.

In Chapter 1, I discussed Margaret Gilbert's framing of her account of plural subjecthood as a way of expounding the atomist foil, under which I had initially brought Searle's constitutive formula. I would now like to turn my attention to Gilbert's account of plural subjecthood itself. She has drawn our attention to what she takes to be exemplary cases of social phenomena—traveling together or walking together. And yet it is these cases that implicate a variety of normativity that is only awkwardly brought under the dictates of the status-function (but also is left underarticulated by the functionalist ideal type).

Building on Georg Simmel's observations, Gilbert notes that understanding social groups or collectivities is essential to understanding social phenomena. Social groups or collectivities are "*special plural subjects concepts*".[44] Given the concept's importance, it is incumbent on us to provide an analysis of the plural subject.

To this end, Gilbert suggests that "one is willing to be the member of a plural subject if one is willing, at least in relation to

certain conditions, to put one's own will into a 'pool of wills' dedicated, as one, to a single goal (or whatever it is that the pool is dedicated to)".[45] Plural subjects are, then, constituted by joint commitments. Individuals who are part of a plural subject "volunteer" or give over their will to the project of the group. Gilbert describes this as a process of "joining forces" with others. Gilbert's evocation of goals, like Searle, thus locates what is distinctive about institutional reality in terms that are susceptible to a functionalist analysis. And yet, like Searle, Gilbert is actually interested in phenomena which cannot be satisfactorily brought under the functionalist's external point of view: she is interested in how interactive agents themselves understand the significance of these goals. Because she is interested in the point of view of participants, her book *On Social Facts* does not just concern "social facts" in Searle's sense, but *institutional* facts—plural subjecthood involves a normative component that cannot exhaustively be articulated in terms of the possibility of dysfunction.

Consider the case of traveling together. To travel together is different from two people who, say, happen to be traveling to the same destination. What is that difference? Gilbert draws our attention to three features of traveling together (or walking together):

First feature. Both people must make it a goal that the other person arrives at their destination. If I am traveling alone, my goals have not been compromised if another passenger who I was not traveling with happens to miss the plane. Her purposes have been undermined, though mine have not. However, if I am traveling together with someone and she missed the plane, then the aims of both have been undermined.

Second feature. Gilbert suggests that each of the two people must somehow understand that it is the other's goal to see to their safe arrival. She must know that it is my goal that she (the other person) arrives at her destination and vice versa. Two people who, perhaps, secretly harbor this goal do not count, on Gilbert's

view, as traveling together. This goal has to be manifested to each other in some way (of course there are a variety of ways of doing this, ranging from subtle to overt). It must be somehow clear that the goal is "our" goal. These first two features amount to Searle's requirement that institutional reality must be the result of a collective imposition of a status-function, where collective intentionality ("we") is distinguished from individual intentionality ("I"). A plural subject emerges when at least two interactive agents impose a function on themselves.

"Social," for Gilbert, denotes any phenomenon that implicates plural subjects. Obviously, for Gilbert, only *people* can be plural subjects (a "we") whereas for Searle, status-functions can be imposed on both interactive X's and indifferent X's. But we have seen that for Searle interactive X's are the fundamental case, since indifferent X's cannot be remiss. Perhaps the only other significant difference between the two accounts is that Gilbert thinks that collective intentionality can be further analyzed in terms of configurations of individual intentions (I intend and I believe that you believe ...). This is why Searle would collapse Gilbert's two features of collective action (shared individual goals plus a coordinating action) into one feature (simply, collective intentional imposition).[46] This difference, however, is localized within a broad agreement as to the overall contours of these kinds of fact.

Third feature. Even in the case that we have the goal that both the other person and I should arrive at the destination, this does not yet count as traveling together. "But still, at this point each one knows only that both individually have a certain goal. Were they to go no further in their assumptions about each other, each might still feel obligated to attempt to fulfill the goal in question independently".[47] Gilbert's point is that if I act to make sure the other arrives safely, I do not do so because of moral obligations (I feel obligated to ensure that the other meets her goals) or prudential considerations (I want to ensure that I accomplish my goals).[48]

To count as traveling together, Gilbert contends, it is not enough that each member has the goal that both he and the other arrive at their destination. They must, further, "join forces". The joint commitment or goal is pursued by each party as the constituent of a plural subject.[49] In short, the goal that both you and I arrive at the destination is "our" goal. These are not the exactly similar *personal* goals of two individuals, but the identical goal of a plural subject.

Gilbert, here, is presenting a version of the intersubstitutability argument that I formulated above. A goal may be brought about by a variety of mechanisms including prudential considerations, or moral considerations, nonmoral or nonprudential obligations, or even coercion or compulsion; these cases, for the structural functionalist, are functionally equivalent. And yet, to count as a social (institutional) fact, it is not enough that there is an X which tends to bring about Y (by whatever means). The means, Y_2, by which the Y_1 is accomplished is constitutive of goal Y. Like Searle, it is at this point where she breaks with the functionalist. The possibility of making a mistake cannot be collapsed into the possibility of dysfunction.

Finally, Gilbert positively specifies the mechanism by which this goal-directed behavior is to count as a social (institutional) fact: plural subjects are constituted by joint commitments, which imply certain *obligations*.[50] For example, if we are traveling together I have an obligation not to spontaneously decide to embark for a different destination. Or if I decide to do so, I am at least under some obligation to inform my partner. Here Gilbert, like Searle, identifies the means, Y_2, by which our Y_1 is brought about in terms of joint commitments or obligations. Gilbert is right to suggest that dysfunction will not adequately characterize the failure to travel together. But, following my criticism of Searle, at this point Gilbert misstates or overcharacterizes the kind of normativity normally indicative of traveling or walking together.

I recommend, contra Gilbert (and Searle), that the possibility of failure when traveling or walking together is more naturally articulated as a defect in being good, rather than as a violation of a commitment, obligation, or requirement. We can see this by considering the ways in which a plural subject might come to fail to walk together. There are mistakes that are only awkwardly brought under the deontology of a status-function: people who do not adjust or compromise their pacing according to the requirements can be defective in Aristotle's sense, but are not remiss in the sense of violating an obligation. Likewise is the case for a companion who walks too near or too far from the other. Depending on the circumstances (there is no rule), using headphones, not using headphones, refusing to converse, refusing not to converse, or straining the conversation with inappropriate responses can constitute a failure at being good, and so a failure in walking together. But this failure is not easily described in terms of remission, the violation of tacit obligations. Even the person who impulsively changes destination seems more akin to an unskillful craftsperson than a violator of the law. Where there is failure in these cases, it is not clear that one or the other participant is remiss; we are more likely to be left feeling puzzled than wronged. We can sometimes characterize this in terms of a violation of rules, but we risk overstating the socially created normative component that underlies such phenomena.

Thus far, I have sought the limits of Searle's ideal type by surveying institutions that largely resist articulation under a rubric of rights and obligations. Searle's deontology does not in every case seem necessary to institutionality, and those are exactly the cases where something like Aristotelian excellence seems a better way to capture the normative component indicative of institutional reality. I will now briefly consider cases where a deontology of rights and obligations does not seem sufficient to something's being an institution, and so must be *supplemented* by Aristotelian virtue.

As we have seen, for interactive agents, not only must X tend to bring about some goal (Y_1) in accordance with Wright's analysis of functional phenomena, but X must also aim to bring about Y_1 in the appropriate way (Y_2). Which X's are appropriate is determined by the collective acceptances of a given society.

Given these two conditions placed on an interactive X, Y_2 is often most helpfully construed in terms of excellence, rather than in terms of rights and obligations. Even in institutions that have the requisite rights and obligations to be satisfied, it is sometimes the case that behavior, which is in fact in accordance with the deontology, is *not* sufficient to bring about Y_1. That is, while Searle's deontology allows to us distinguish appropriate from inappropriate means (Y_2), it often does not satisfy the more basic requirement of being sufficient to actually bring about Y_1.

Rights and obligations typically mark boundary conditions, barring behavior that would promptly undermine the institution itself. A batter, for example, has an obligation to stop batting if struck out three times. But it is not at all clear that a batter has an obligation to *sprint* to first base upon successfully hitting a ground ball. If he strolls to the base and so is tagged out, his problem is not one of remission. Likewise, to use a version of Searle's example, consider the shortstop who, upon catching the ground ball, begins to eat the ball instead of throwing the stroller out. The shortstop is not remiss, so much as unskillful. We can imagine innumerable examples where participants engage in behavior that is not remiss, but nevertheless, do not—as required by Y_2—tend to bring about Y_1. If so, the institution of baseball cannot be constituted by functions of the form Y_2 (and Y_1), whose content is delineated strictly in terms of obligations or requirements. Except, perhaps in extreme cases, such as a Japanese tea ceremony or chess, such requirements are almost always such that they alone could not bring about Y_1.

Again, it is possible to appropriate MacIntyre's discussion of the good life as a whole, to say something about the normativity

required for institutionality: MacIntyre discusses two ways to fail—we can fail by violating certain obligations and we can fail by not being good enough. "An offense against the laws destroys those relationships which makes common pursuit of the good possible; defective character, while it may also render someone more liable to commit offences, makes one unable to contribute to the achievement of that good without which the community's common life has no point".[51] An offense against the law, or a violation of rights and obligations, are particularly egregious actions that destroy the capacity for participants to achieve the aims of a given institution (Y_1). While such rules circumscribe an institution's most vulnerable flanks, such trespasses are nevertheless not the *only* way to undermine Y_1—"For both [kinds of failure] injure the community to some degree and make its shared project less likely to be successful".[52]

Searle distinguishes between the constitutive rules and regulative rules of an institution; the former are conditions for the possibility of the institution, whereas the latter guides antecedently existing activities (*CSR*, pp. 27–8). But the baseball example suggests the difference between what Searle calls constitutive rules and regulative or even strategic rules are not a difference in kind; rather, their violation is only less threatening to the integrity of the goals of the institution (Y_1). A few constitutive rules in the form of rights and obligations may be violated, and one still may be considered a participant in a Japanese tea ceremony. Looking at chess, Searle grants this when he writes that, "The rules are *constitutive* of chess in the sense that playing chess is constituted in part by acting in accord with the rules. If you don't follow at least a large subset of the rules, you are not playing chess" (*CSR*, p. 28). But if it is possible to participate in an institution, even while in violation of a few of the constitutive rules, contrapositively, it is possible to not participate in such an institution, even when in full accordance with the constitutive rules, conceived in terms of rights and

obligations. If in playing baseball participants flout enough of the regulative or even strategic apparatus, it is no longer the case that they are playing baseball. This can be put in way that avoids talk of regulative or strategic rules altogether: if it were not the case that many of the participants were not always already skillful or competent, even if all such obligations were met, it is not the case that their behavior would tend to bring about Y_1. If an appropriate X (Y_2) is strictly conceived in terms of the satisfaction of rights and obligations, it is rarely the case that Y_2 would tend to bring about the shared projects, goals, or functions (Y_1) indicative of institutional reality as required by Wright's analysis of functional phenomena.

Perhaps I am reading "rights and obligations" too narrowly, excluding a vast uncodified deontology to which participants are subject. It may be the case that the shortstop is not in violation of a codified rule, but an uncodified rule, and so is remiss in some sense. But this is unsatisfactory: when enablements and requirements are rendered so richly, something like excellence more perfectly captures the normative component indicative of institutional reality. Perhaps a Japanese tea ceremony, with its myriad prescriptions, is the sort of exemplary case on which Searle models his account of institutional reality. In Searle's view, other institutions are importantly similar to the tea ceremony, except that the rules are less explicit and largely uncodified. But if other institutions are similar to the tea ceremony we are led to the almost indigestible conclusion that a cocktail party is just a more Byzantine version of the tea ceremony; both are bounded by rights and obligations, and yet the former burdens the participant with an almost inexhaustible set of permutations. Such obligations are not only unnecessary to something's being an institution, but that such highly codified institutions often do not even seem to be the paradigm case. It is in the cocktail party that the *phronimos* is most prominently on display; these activities, in the Aristotelian model of normativity, are

exemplary of our institutions. There are institutions that do not seem to be bordered by rules, the content of which is explicated in terms of rights and obligations.

Moreover, we have seen that Searle has difficulty accounting for this labyrinthine rule structure, so construed. The rule structure is not located in the conscious intentional structure of participants and Searle rightly resists situating that structure within an unconscious network of beliefs and attitudes. These rules, then, might be found in the nonintentional Background. But how this is possible is mysterious; indeed, Searle himself sometimes construes the Background as neurophysiological clay against which a "rule structure" impresses itself. But it is the possibility of this "rule structure" which worried us in first place.

We generally *are* subject to rights and obligations, but these rules are themselves not robust enough to constitute an institution. Indeed, while I have not argued for this stronger claim, my suspicion is that such a deontology is largely derivative of an ability to effortlessly go on together even in both familiar and novel situations—and this ability is what is embodied in the actions of the *phronimos*. Searle is right to suggest that the point of view of interactive institutional participants implicates a normative component inaccessible to the structural functionalist, but is wrong to construe this competence *strictly* in terms of rights and obligations.

I hope it is clear that my aim is *not* to insist that Searle has said something *false* about institutional reality. Rather, both Searle and Aristotle have articulated *ideal types*, each of which uncovers and obscures certain features of the infinite multiplicity which is institutional reality. The normativity indicative of institutionality can be brought under at least two ideal types, one of which is Searle's.

Ideal types are tools that give rise to puzzles and provide part of the means to their resolution. Recall that Weber reviewed two responses to crisis. First, an ideal type can be "stretched" to accommodate anomalous phenomena; we saw Searle respond

to Smith's counterexample by leaning more heavily on the constitutive formula's Y term—the status-function. Here, we see how an ideal type (or mnemonic) gives rise to a puzzle and also provides the means to its resolution. But sometimes the inquirer is confronted with a phenomenon that simply cannot be easily accommodated by a given ideal type. Ideal types take certain cases as exemplary, and make modifications to account for errant, peripheral cases.[53] This is the case with money construed as a medium of exchange: it is our familiarity with economic situations that resemble the village fair that make the orthodox model compelling in the first place. As we have seen, this ideal type is then stretched to cover increasingly dissimilar economic transactions and historical situations, and so becomes increasingly less perspicuous. The orthodox ideal type misrepresents the historical record, and badly misconstrues the workings of highly centralized economies (unlike the chartalist account). Likewise, Searle's appeal to rights and obligations may account for much of the normativity constitutive of our institutions, but there are cases in which such an appeal seems neither necessary nor sufficient for institutionality. A similar breakdown is also seen in Searle's attempt to account for this normative component in terms of intentionality, the Background, or a rule-structure.

Stretching eventually pushes the limits of perspicuity, or even intelligibility. When this happens we can supplement the understanding embodied in the original ideal type with an additional ideal type. This is Weber's second response to crisis. Indeed, Weber's infinitude suggests that all ideal types must break down eventually. In this case, radically uncodified institutional phenomena such as cocktail parties cannot be comfortably brought under the status-function model. But these cases do not falsify Searle's ideal type so much as they indicate its scope. They are an indication that a different tool might better serve the interests of the inquirer.

When we compare Searle's model of institutionality with that of the Aristotelians' we see this pattern. Searle's begins with

explicitly and intentionally codified institutions as exemplary, and then stretches the ideal type by positing uncodified rules and the Background to cover the remaining cases. An Aristotelian conception of institutionality, however, begins with those cases that Searle has the most difficulty covering—cocktail parties, dating, Wittgenstein's building activity; these are cases where the participants exhibit mastery in the form of spontaneous adaptation, but would not necessarily have the tools to be able to articulate the rules that were being followed. Indeed, there is a sense in which rule-following is the wrong way to describe such institutions—if people are defective cocktail-party goers, it is not because they have violated some uncodified constitutive rule. According to the Aristotelian ideal type, institutions are represented by skillful individuals, rather than constituted by rules. These institutions, which resist articulation in terms of the status-function with its associated rights and obligations, are those most effortlessly rendered by a conception of normativity that looks less to rules, and more to idealized participants—the *phronimos*. Skillfulness, rather than rule-following, is the dominant metaphor, targeting the socially created normative component indicative of institutional reality.[54]

Van Fraassen and the ideal types of institutional reality

I have so far argued that the status-function (or constitutive formula) is an ideal type because there are some institutions that seem to require a kind of normativity not characterizable in terms of rights and obligations. I now want to make the more radical claim that even in the case of those institutions that *can* be circumscribed by Searle's deontology, we should nevertheless regard the status-function as an ideal type. If a puzzle we may have can be solved strictly by appeal to rights and obligations, it is in virtue of the interests that gave rise to the puzzle in the first place. This counters the suggestion that, even with respect to

the most rule-bound of institutions, Searle's deontology is necessarily more fundamental than the Aristotelian model of normativity. Searle might argue, perhaps, that the latter principally concerns regulative or strategic rules, which are derivative of constitutive rules.

In suggesting that both Searle and the functionalist have articulated ideal types, I am making a point about how these two models can be expected to fit into a discursive activity. They both give rise to, and help solve, *different* puzzles. Both Searle and the functionalist are providing frameworks that can be used to answer why-questions about the behavior of participants. It principally is in the answering of why-questions that a discursive activity finds momentum, although in the event of crisis sometimes inquirers are forced to step back and ask what-questions; these are questions about the ideal type, model, or paradigm by which we can ask why-questions in the first place. Answers to what-questions only acquire significance in the context of an actual puzzle-solving activity (or, minimally, against the possibility of such an activity).

Different ideal types give rise to different sorts of puzzles, even when concerning the same phenomenon. I wish to appeal to Bas van Fraassen's theory of why-questions in order to further characterize the sense in which Searle and the Aristotelian have articulated competing ideal types by which we might characterize institutional reality.

Van Fraassen, in Chapter 5 of *The Scientific Image*, suggests that explanation is best understood interrogatively, as an answer to a particular question that arises over the course of inquiry. Van Fraassen's question-answering model of explanation readily coheres with Kuhn's puzzle-solving idiom.[55] Van Fraassen's formalized rendering of erotetic explanation usefully highlights several criteria by which we can distinguish between various explanations of the same phenomenon. Van Fraassen cites Norwood Hanson who writes:

> There are as many causes of x as there are explanations of x. Consider how the cause of death might have been set out by a physician as "multiple haemorhage", by the barrister as "negligence on the part of the driver", by a carriage-builder as "a defect in the brakeblock construction", by the civic planner as "the presence of tall shrubbery at that turning".[56]

In explanation, the salient feature picked out as the cause of the death-by-accident differs, depending on, among other things, the interests of the discussants. Because there are an infinitude of causes, we can imagine an infinitude of potential interests. But when we look at the requirements of explanation in our actual discursive and non-discursive practices, ordinarily most events do not in fact raise *any* question for us—and those that do usually raise no more than a few different sorts of questions. However, *when* several questions or puzzles are raised about the same event, these need to be treated as *distinct*. Van Fraassen goes on to explicate this point:

> It is important to notice that in a certain sense these different answers cannot be combined. The civic planner "keeps fixed" the mechanical constitution of the car, and gives his answer in the conviction that regardless of the mechanical defects, which made a fast stop impossible, the accident need not have happened. The mechanic "keeps fixed" the physical environment; despite the shrubbery obscuring vision, the accident need not have happened if the brakes had been better. What the one varies, the other keeps fixed, and you cannot do both at once.[57]

Normally, the form of the question or puzzle the inquirer is addressing, plus the context, is enough to give us a sense of what is being kept fixed. But, in this case both the form and context fail us, where all of these proposals purport to explain why it is the case that S died in a car accident, we need an additional device to disambiguate the question. This is just to say that questions are not

always transparent, and when they are not we need a way of getting clearer on exactly what is being asked.

In "Questions," C.L. Hamblin suggests that a question is equivalent to its possible answers. We understand the question, "In which continent is Luxembourg?," when we understand the sort of statements that might count as answers, even if we do not know which of these statements is correct: "Luxembourg is in Asia" or "Luxembourg is in Europe," etc.[58] Listing its possible answers, for Hamblin, is a means by which a question is uniquely identifiable.[59]

The barrister, civic planner, carriage-builder, and physician *seem* to be asking the same question: "Why did S die in an accident?" We can get a better sense of the question asked by the different inquirers by listing the possible explanations considered by each: the barrister is primarily interested in the driver's *mens rea* (or lack thereof), and rejects the possibility that it was an accident or that he killed himself intentionally. While he ultimately concludes that the driver did not act intentionally or according to a reasonable standard of care, these rejected answers give us a better sense of which question is asked, and why it is different from that of the civic planner. The civic planner cited the shrubbery, because he saw the turn was clearly marked, etc. Van Fraassen would state the point by noting that negligence and the shrubbery bear different relevance relations (R) to the fact that so-and-so died (P_K, X)—they address different domains of inquiry. While all these different relevance domains may be addressed separately by the same person, for each stage of the investigation things that were, and will be, and could be called into question must be kept fixed.

In what Wright calls the second continuum of relevance, there are an infinite number of conditions that *may* be cited as a cause of the accident.[60] Explanation would be impossible without a further specification of relevance. Contextual factors, like those cited in Hanson's quote, then, cap this infinitude: in this case, the city

planner, for example, is interested only in the small subset of those things under control of the city's planning department.

But this solution comes bearing its own potential demons: we need to remain especially alert to the possibility of inadvertently slipping between various relevance domains, even as they purport to answer what nominally appears to be the same why-question (why did S die in the accident?). It is this concern which is behind van Fraassen's observation that "What the one [relevance domain] varies, the other keeps fixed, and you cannot do both at once." A question is malformed because the civic planner is no more (or less) interested in the culpability of the driver, than he is in fuel availability, the distribution of driver's licenses, or even Newton's first law; while these are conditions that led to the accident, they simply are not part of the causal chain that the civic planner is principally interested in. They are largely irrelevant or incidental, except insofar as they help inform the relevance domain that the civic planner *is* interested in. Newton's first law may be useful to the city planner—especially if he determines that, say, it was the wet concrete rather than tall shrubbery that explains the death—but this can only help supplement an answer to the question that he *is* addressing. More typically, the law of inertia or the driver's *mens rea* are, for the civic planner, incidental, irrelevant, or unexplanatory—even if they are among the conditions that brought about the death.

With this in mind, I would like to return to the three models of normativity we surveyed. Upon entering a university auditorium, we see a woman speaking in front of a large audience. While the context is quite rich, we can still imagine formulating a why-question about the situation: Why P_K?—why is this woman speaking in front of an audience? This question is indeterminate in that it flags a number of different why-questions or puzzles we might have about the same activity. The relevance relation aims to disambiguate the questions by drawing our attention to different interests inquirers may have.

The question could be about the function or purpose of the activity (Y_1). Perhaps our suspicions are confirmed that this woman is trying to teach the audience something. The fact that her instruction is falling on deaf ears does not preclude its having a certain function: it turns out that this is an Introduction to Philosophy class and she foolishly included Chapter 5 of van Fraassen's *The Scientific Image* on the syllabus. But, this is not the only functional story that would explain the woman's behavior. It could have been the case that her goal was not to teach, but to entertain: perhaps we have happened upon a Shakespearean monologue. The functionalist's question is one why-question or puzzle we might have about the activity.

Why is the woman standing in front of the audience? The purposes of the activity might be obvious, and yet we still might have puzzles about the rights and obligations by which the action is permissible. This is *one* question we can have about the appropriate means (Y_2) by which Y_1 is brought about. In van Fraassen's terms, the purposes of the activity are kept fixed so that an inquiry into the rights and obligations of the actor is possible. Why is she instructing the audience? Because, as a teacher, she has been specifically enabled or even obligated to do so in virtue of being a Faculty Fellow. This is a question about the status-function of the activity—what are the rights and obligations that permit this sort of activity. Or perhaps she is instructing the audience, not in virtue of her status as a Faculty Fellow, but in virtue of being an Associate-In (in the University of California, this status does not require a doctoral degree), a guest lecturer, or a performance artist. In each of these cases, the function is to teach, but the configurations of rights and obligations that help bring about this function are distinguishable. While the why-question nominally remains the same, we can distinguish the functionalist's question from that of Searle's by way of the relevance domain.

Finally, the status-function might be either obvious or irrelevant: perhaps it is clear she is teaching in virtue of her being a

Faculty Fellow, or we are not particularly concerned about her status. But we might have a question as to the skillfulness or excellence by which she is able to convey the subject matter. Is she defective in being a teacher? Why is she talking to the audience? Perhaps she feels that lecturing in front of the audience is less distracting than lecturing while wandering throughout the audience. Or perhaps she understands that addressing her audience is more effective than facing the chalkboard or reading off her notes.

This skillfulness is not dependent on her actually bringing about the goal in question, but is, rather, manifest in the means by which that goal is pursued. That is, perhaps she had stepped in for another instructor who fell ill just before the term began (so that the very fact that van Fraassen is on the syllabus in the first place cannot be taken as an indication of unskillfulness), and her failed efforts to convey van Fraassen's theory of why-questions to freshmen would be nevertheless recognized by competent onlookers as heroic.

The question as to whether she is satisfying certain codified or uncodified obligations may be quite different from the question as to her skillfulness in teaching. Both involve normative components that cannot be exhausted by a functionalist rubric, but the possibility of failure is articulated in quite different ways. There is at least a sense in which what makes someone a teacher cannot be explicated in terms of a cluster of rights and obligations. Here, teaching is a success term in a way that it is not when the institution is construed under the status-function, when excellence is constitutive of the institution. Indeed, with enough skill, someone might be remiss on even basic obligations and yet still be recognized as a teacher. For example, a skillful teacher might even violate the requirement of showing up to the lecture, if in not doing so she is able to communicate a point about the importance of attendance to a class, where excessive absences are a problem. Perhaps assigning a grade to coursework is a basic obligation of an instructor. But we can imagine a professor who refrains from assigning a grade to

a student, if he or she had worked closely with that student; to have done so may have trivialized the exchange, which perhaps would be a kind of initiation into the life of a professional academic. While many institutions are in fact bordered by perhaps inviolable rules (three strikes always counts as an out), rules of this kind are not necessary to an activity's being institutional. I recommend that we might rather look to the capacity for spontaneity and innovation in the form of skillfulness as a help-ful criteria by which to distinguish institutional from noninstitutional activity.

Similar to the barrister and city planner, Searle and the Aristotelian are simply asking different questions about the same type of action. The difference is a matter of what the inquirer keeps fixed.[61] And this is a function of the interests of the inquirer; the driver's *mens rea* is central to the inquiry of the barrister, but peripheral (i.e., held fast) to that of the civic planner. We can prioritize one phenomenon over another depending on whether or not it is relevant to the interests of particular inquirers. To canonically prioritize certain phenomena requires that we reify certain of these interests.

The important thing to see is that the puzzle about rights and obligations that underlie an activity is distinguishable from a question about its excellence. Although rights and obligations are often related to an activity's function, when it comes to questions we might have concerning the activity, one is not necessarily relevant to the other; establishing that the person is not remiss may be largely irrelevant, incidental, or unexplanatory, when it comes to a question concerning skillfulness. At best, according to the Aristotelian ideal type, to flout codified or uncodified rights and obligations is *one* way to exhibit incompetence.

Conclusion

I have brought two complaints against Searle's conception of institutional reality. First, I have argued that there is an alternative

conception of normativity under which institutional phenomena can be characterized. The *phronimos* fails differently than someone who violates implicit or explicit rights and obligations. I contend that, first, the spontaneity and plasticity indicative of the *phronimos* resists articulation under a rubric of rights and obligations and, second, it is not always the case that rights and obligations even represent boundary conditions on institutional phenomena. The Aristotelian conception of normativity can stand on its own.

Second, I have made a general point about how we ought to understand the significance of these two depictions of institutional reality. Both Searle and the Aristotelian have articulated ideal types, which are tools of inquiry.

In articulating an alternative to the status-function, I *could have* gone on to contend that the Aristotelian conception of normativity more accurately reflects the structure of institutional reality; in this view, to follow Gilbert, Aristotle "carves nature at her joints".[62] But, this is to do metaphysics in the sense attributed to Searle in the first chapter, and so runs afoul of difficulties outlined in Chapters 2 and 3.

These two criticisms of Searle can be distinguished from each other. If one were interested in doing metaphysics, one could make the stronger claim that Searle's deontology gets the normative component of institutional reality *wrong*, whereas Aristotle gets it *right*. Indeed, I have suggested that a case could be made for thinking that a deontology of rights and obligations is importantly derivative of Aristotelian competence or excellence. Institutions are not built up or everywhere bounded by rules, but it is only against the background of competence embodied in the *phronimos* that constitutive rules acquire significance. If we were constitutionally different, we might need a restriction against eating baseballs.

If the project were metaphysics, then the *anomalies* that challenge Searle's account are reconstrued as *counterexamples* that *falsify* the account. On the ideally typical conception, however,

these anomalies gesture to the *limits* of a given dialogical tool. These limits are not fixed—an ideal type may be indefinitely "stretched" or modified to accommodate errant applications. But something like the law of diminishing returns sets in as we lose track of the picture that made a particular ideal type compelling in the first place; indeed, as we have seen with Searle's notion of the Background, the picture can be pushed beyond the point of intelligibility.

Moreover, it is not clear what should motivate an excursion into metaphysics. Following the atomist "under-labourer" conception of philosophy,[63] both Searle and Gilbert ("Presumably in a logical ordering, ontology precedes methodology.") suggest that methodological endeavors such as the social sciences somehow require or are benefited when we are clear about the ontology which underlies the object of investigation. Searle takes himself to be answering "questions [that] concern what might be thought as problems in the foundations of the social sciences" (*CSR*, p. xii). He asserts that the founding fathers of the social sciences—Weber, Simmel, and Durkheim—"were not in a position to answer the questions that puzzle me, because they did not have the necessary tools" (*CSR*, p. xii). Indeed, the founding fathers' purported failure to shore up sociology's substructure is what drives Osborne to characterize the *Construction* as "an aesthetic achievement, a satisfying tidying-up operation, just a nice, careful putting of things into proper order".[64]

But what if, as Weber suggests, institutional reality is not amenable to concise articulations of the kind offered by Searle? Indeed the considerable difficulties such theories confront suggest that while we can talk about this reality, no such canonical descriptions are in fact available to us. Our confidence that the status-function mirrors the form of fact, or cuts nature at her joints, seems shaken. Moreover, these discursive activities, which motivate the metaphysical inquiry, in fact do not appear to need metaphysical buttressing. Osborne surveys the ideal types advanced by working

sociologists and concludes that "one can see that these methodologies are pragmatic; they establish the contours or terrain of a research programme, and they make little sense outside that programme".[65] But he worries about the contribution of Searle's text to actual sociological inquiry: "it is quite literally *indifferent* as sociology".[66]

I do not agree with Osborne's contention that Searle's contribution is "irrelevant" to empirical sociology. The status-function is one of at least two portraits by which institutional phenomena can be rendered. Just as the neoclassical ideal type of money could *not* count names on a clay tablet as money, but the chartalist account would, Searle's ideal type excludes certain activities from the sphere of the institutional which are otherwise admitted by Aristotle. Moreover, even when an activity has an identifiable deontology of rights and obligations, certain competencies are excluded as accidental or unconstitutive of the institution. But Osborne is responding to Searle's *characterization* of his project, not the project per se. And Osborne is right to wonder why discursive activities need the kind of metaphysical underpinning articulated by Searle and Gilbert. The methodological endeavor does not stand in need of a crutch. If we cannot look to our practices to provide the motivation for metaphysical inquiry, it is difficult to see what the question is that these authors take themselves to be answering.

The ideal type reading is deflationary; it at once motivates these authors' theories by seeing how they might in fact be used to drive actual investigations, while simultaneously avoiding the sticking points endemic to ontological inquiry. As we have seen, inquiry proceeds piecemeal, as various paradigms, models, and ideal types generate puzzles and questions, which are subsequently addressed in different ways; so long as this discursive process continues fruitfully, the ideal type is taken to be valuable. Osborne puts it this way: the ideal types of the social sciences "were constructed according to the needs and circumstances that were strictly *local*".[67] Ideal types are wrenched from our extant understanding of how things

work and reapplied to novel phenomena. While this extant understanding is responsive to real patterns or actual regularities,[68] it need not mirror the form of fact in any strong sense; indeed, if Weber is right, we cannot expect *this* of our models. Social scientific inquiry, or any investigation into our institutions, does not depend on a foundation, following Searle's characterization of the constitutive formula, any more than natural science, following van Fraassen, depends on a realist interpretation of associated concepts and theories. If, however, what Searle means by "ontological" something akin to "ideally typical" (sometimes, in this vein, we speak of "ontologies" or "epistemologies"), then Searle has chosen misleading terms by which to characterize his project. In this case, he and I are in broad agreement, and I hope this book has taken steps to dispel a possible misreading of his text.

But to read pragmatics back into Searle's no doubt correct observations about institutional reality also radically changes the significance of "foundation," for "foundation" cannot mean some category or model that is presupposed by *any* depiction of our institutions. Foundations become localized to a particular question that arises within a given discursive activity: as van Fraassen notes, that which is "kept fixed" changes as we become interested in different aspects of the same phenomenon. In looking at some behavior, whether we are inquiring into constitutive rights and obligations, or into various models of excellence, or even into its ends, depends on what gaps or breaks we happen to have in our understanding; and this, following van Fraassen, depends on robust, local, contextual factors. So, in suggesting that the constitutive formula or status-function are ideal types I have, in a sense, turned the professed relationship between Searle and these founding fathers, and Weber in particular, on its head. I have used Weber to put things that Searle has said into their proper order—a satisfying tidying-up operation.

Weber's meta-sociological remarks are, themselves, foundational in the sense that they help say something about what we

can and cannot expect from our theories. This is done—not by looking at and describing the basic inventory of facts with which the sociologist has to work—but by looking at the discursive practice in which these models acquire significance, and for that we have to look to the role these concepts play in our discursive activities. To appropriate Searle's vocabulary, Weber is interested in discursive foundations, the conditions by which understanding is engendered. Just as ontological foundations aim to clarify the significance of certain kinds of potentially puzzling *objects* (i.e., the irreducibility of institutional facts), Weber articulates certain limitations in our ability to talk about the world that we should not expect to breach.

But the atomist reading of the *Construction*, a version of which Searle ultimately seems to adopt, constitutes one such trespass. We should not expect a canonical articulation of social and institutional reality, given, among other things, the infinite multiplicity of the object of discourse. And even when institutional reality is most generally articulated, institutional reality is susceptible to a variety of incommensurable characterizations. If the view that the constitutive formula/status-function is ideally typical is not Searle's, it is my contention that it is the right view, preserving that which is interesting about the *Construction* while avoiding the flaws of the atomist reading.

Perhaps one of the reasons why Searle is moved to observe that "the questions I am addressing in [my] book have not been satisfactorily answered in the social sciences," is because the "great philosopher-sociologists" had already rejected that kind of question.

Notes

1. See also Searle, *Mind, Language, and Society: Philosophy in the Real World*, pp. 123–4.
2. Searle, *Speech Acts: An Essay in the Philosophy of Language*, pp. 51–2.
3. Dreyfus, "The Primacy of Phenomenology over Logical Analysis," p. 2.

4. Hacking, "Searle, Reality and the Social,"p. 85.
5. Searle and Smith, "The Construction of Social Reality: An Exchange," p. 301.
6. Ibid.
7. Smith, *John Searle*, pp. 16, 19, 20.
8. Searle and Smith, "The Construction of Social Reality: An Exchange," p. 300.
9. Ibid., p. 305.
10. Ibid., p. 301.
11. Weber, " 'Objectivity' in Social Science and Social Policy," pp. 84, 104.
12. Ibid., p. 106. Italics mine.
13. Ibid.
14. Ibid., p. 104.
15. Ian Hacking, *The Social Construction of What?* (Cambridge, Mass: Harvard University Press, 1999), pp. 103–6.
16. Ian Hacking, *Rewriting the Soul: Multiple Personality and the Sciences of Memory* (Princeton, NJ: Princeton University Press, 1995).
17. Notice that we can also talk about indifferent- and interactive-noninstitutional-agentive-functions, where, say, a person rather than a rock is used as a paperweight.
18. As suggested, because there are interactive and indifferent status-functions, Searle need not be threatened by Smith's counterexamples—free-standing Y's. These could only concern indifferent X's. Thus, if we narrow our focus to the interactive status-functions, which are anyway "fundamental," the constitutive formula and the status-function are identical.
19. Wright, *Teleological Explanations: An Etiological Analysis of Goals and Functions*, p. 81.
20. Christopher Marquis, "Absent from the Korea Talks: Bush's Hard-Liner," *New York Times*, September 2 2003.
21. Geoffrey K. Pullum, *The Dan Brown Code* (Language Log, 1994 [cited]); available from http://itre.cis.upenn.edu/-myl/languagelog/archives/000844.html.
22. Wright, *Teleological Explanations: An Etiological Analysis of Goals and Functions*, p. 59.
23. Alasdair C. MacIntyre, *After Virtue: A Study in Moral Theory*, 2nd ed. (Notre Dame, Ind.: University of Notre Dame Press, 1984), p. 149.
24. Here Searle means to exclude honorific cases, which have a status but no associated rights and obligations.

25. John Haugeland, "Dasein's Disclosedness," in *Heidegger: A Critical Reader*, ed. Hubert L. Dreyfus and Harrison Hall (Oxford, UK; Cambridge: Blackwell, 1992), p. 34.
26. Talcott Parsons, *Social Systems and the Evolution of Action Theory* (New York: Free Press, 1977), p. 102.
27. Alean Al-Krenawi and John R. Graham, "Conflict Resolution through a Traditional Ritual among the Bedouin Arabs of the Negev," *Ethnology* 38, no. 2 (1999): pp. 166–7.
28. Wright, *Teleological Explanations: An Etiological Analysis of Goals and Functions*, p. 69.
29. Interestingly, Margaret Gilbert would add moral and prudential considerations to the list of mechanisms that can be contrasted with the kind of deontology that underlies institutional facts. See Gilbert, *On Social Facts*, pp. 162–3, 392–6, Margaret Gilbert, "Walking Together: A Paradigmatic Social Phenomenon," *Midwest Studies in Philosophy* 15 (1990): pp. 4–6.
30. D.F. Hake, T. Donaldson, and C. Hyten, "Analysis of Discriminative Control by Social Behavioral Stimuli," *Journal of the Experimental Analysis of Behavior* 39 (1983).
31. Wittgenstein, *Philosophical Investigations*, §6.
32. Winch, *The Idea of a Social Science and Its Relation to Philosophy*, p. 58.
33. Norbert Elias *et al.*, *The Civilizing Process: Sociogenetic and Psychogenetic Investigations*, rev. ed. (Oxford; Malden, Mass.: Blackwell, 2000), p. 61.
34. Ibid., p. 50.
35. Searle may contend that Erasmus' instructions do not concern constitutive rules at all, but are rules that regulate an existing *social* activity. If Searle would wish to say eating meat with three fingers (X) is not a constitutive rule, it is presumably because X performs its function solely in virtue of its physical structure. Status-functions, Searle argues, do not (*CSR*, pp. 41–2). Eating with three fingers, which performs its function in virtue of the underlying physical structure, thus flags a non-institutional social fact. I argued that the physical structure of X does not provide a reliable guide for distinguishing non-institutional social facts from institutional facts; we should rather look to the presence of the required normative component, which governs not just the end, but the appropriate means by which that end is brought about. Eating meat with three fingers does satisfy this criterion, as eating meat with two hands would be considered inappropriate even if X accomplished the end of feeding oneself. Moreover, even if Erasmus has

provided only a regulatory guide to eating, so qualified, it is still illustrative in the right way.

36. In fairness to Searle's dog, there is quite a bit of evidence that suggests such animals are quite capable of appropriate spontaneity. It is something like this adaptive capacity which prompts professional dog trainers to speak of a dog's sense of "responsibility," insisting that they act with "intelligence and repose". See Vicki Hearne, *Adam's Task: Calling Animals by Name* (New York: Knopf: distributed by Random House, 1986), p. 14.
37. Ludwig Wittgenstein, *Remarks on the Foundations of Mathematics*, trans. G.E.M. Anscombe, Rush Rhees, and G. H. von Wright, 3rd, rev. ed. (Oxford: Blackwell, 1978), p. 143.
38. MacIntyre, *After Virtue: A Study in Moral Theory*, p. 67.
39. Ibid., pp. 151–2.
40. In another sense, it is perfectly fine to talk about the expert's or *phronimos*' behavior in terms of rule-following. Wittgenstein §50, Winch, and MacIntyre talk about competence in terms of rule-following, while denying that these rules can be formulated. "Rule-following" does, among other things, capture the normative aspect, and the possibility of failure, indicative of even the least codified of our institutions.
41. Hubert L. Dreyfus, "Could Anything Be More Intelligible Than Everyday Intelligibility? Reinterpreting Division I of Being and Time in the Light of Division II," in *Appropriating Heidegger*, eds James E. Faulconer and Mark A. Wrathall (Cambridge; New York: Cambridge University Press, 2000), p. 162.
42. MacIntyre, *After Virtue: A Study in Moral Theory*, p. 150.
43. Ibid., pp. 150, 61–2.
44. Gilbert, *On Social Facts*, p. 408.
45. Ibid., p. 18.
46. It is precisely this difference that prompts one reviewer of an anthology of papers on social facts and collective intentionality to remark, "Most writers in the Gilbert-Tuomela tradition take it so much for granted that one should look for necessary and/or sufficient conditions for 'veridical' we-intentions and we-beliefs, that they can't imagine that Searle is not trying do this too." Ingvar Johansson, *Review of "Social Facts & Collective Intentionality"* (Notre Dame Philosophical Reviews, 2003 [cited]); available from http://ndpr.icaap.org/content/archives/2003/3/johansson-meggle.html.
47. Gilbert, *On Social Facts*, p. 162.

48. Ibid., pp. 162–3, 392–6, Gilbert, "Walking Together: A Paradigmatic Social Phenomenon," pp. 4–6.
49. Gilbert, *On Social Facts*, p. 163.
50. Margaret Gilbert, *Living Together: Rationality, Sociality, and Obligation* (Lanham: Rowman & Littlefield, 1996), p. 365; Gilbert, *On Social Facts*, pp. 198, 205, 409.
51. MacIntyre, *After Virtue: A Study in Moral Theory*, p. 152.
52. Ibid.
53. Compare this feature of the ideal type to Kuhn's discussion of the preparadigm schools. Looking at electrical research, "One early group of theories, following seventeenth-century practice, regarded attraction and frictional generation as the fundamental phenomena. This group tended to treat repulsion as a secondary effect and also to postpone for as long as possible both discussion and systematic research on Gray's newly discovered effect, electrical conduction." This school is contrasted with those who regarded attraction and repulsion as elementary manifestations of electricity. Both of these schools are further contrasted with those who came to see electricity on the model of a fluid, rather than effluvium. This group could easily account for conduction (that is what electricity does), but, likewise, had difficulty accounting for electricity's attractive and repulsive effects. Ideal types, like these preparadigm schools, cluster around exemplary cases, so that, typically, one ideal type is most confident just at the point where competing ideal types begin to falter. See Kuhn, *Structure of Scientific Revolutions*.
54. Note that for highly codified institutions, unskillfulness, or a failure at being good, analogously seems to misrepresent the way in which participants might fail.
55. Matti Sintonen, who also construes scientific explanation erotetically, explicitly takes himself to be "translat[ing] problem-solving talk into the interrogative idiom". Puzzles are solved in the same way that questions are answered. See Matti Sintonen, "Explanation: In Search of the Rationale," in *Scientific Explanation, Minnesota Studies in the Philosophy of Science, Volume Xiii.* (Minneapolis: Univ of Minn Pr, 1989), pp. 279ff.
56. Hanson in van Fraassen, *The Scientific Image*, p. 125.
57. Ibid., p. 126.
58. C.L. Hamblin, "Questions," *The Australasian Journal of Philosophy* 36, no. 3 (1958): p. 162.

59. According to van Fraassen, another means by which a question is identifiable is by way of its contrast-class. See van Fraassen, *The Scientific Image*, pp. 126–9, 141.
60. Larry Wright, "The Concept of a Reason," *unpublished* (2005): pp. 8–10.
61. Van Fraassen lists three broad relevance domains or kinds of request specification, which become further subdivided as the respect-in-which a reason is requested is buttressed by additional context (i.e., Hanson's dead driver): first, the request might be for events "leading up" to the event in question (negligence, perhaps). Second, if these are already well known, then the request may be for a standing condition which made that event possible in the first place (tall shrubbery). Finally, the question may be about the function of the event in question, in which case the underlying means mechanism is probably irrelevant. One fascinating implication of Searle's and the Aristotelian's models of normativity is that it adds, at least when the explandum is a kind of human activity, two new varieties of request specification. These other request specifications may be answered, or at least treated as fixed, and yet with respect to human activity we may still have questions about one or the other of these normative components.
62. Gilbert, *On Social Facts*, p. 442.
63. Winch, *The Idea of a Social Science and Its Relation to Philosophy*, p. 4.
64. Osborne, "The Limits of Ontology," p. 97.
65. Ibid., p. 99.
66. Ibid., p. 98.
67. Ibid., p. 99.
68. van Fraassen, *The Scientific Image*.

Bibliography

Al-Krenawi, Alean, and John R. Graham. "Conflict Resolution through a Traditional Ritual among the Bedouin Arabs of the Negev," *Ethnology* 38, no. 2 (1999): pp. 163–74.

Baumol, William J., and Alan S. Blinder. *Economics, Principles, and Policy*. New York: Harcourt Brace Jovanovich, 1979.

Budish, Eric B., and Lisa N. Takeyama. "Buy Prices in Online Auctions: Irrationality on the Internet?" *Economics Letters* 72, no. 3 (2001): pp. 325–33.

Burger, Thomas. *Max Weber's Theory of Concept Formation*. New York: Duke University Press, 1987.

Dickinson, Michael. "Solving the Mystery of Insect Flight," *Scientific American* (2001): pp. 49–57.

Dowd, Kevin. "Aristotle on Money," in *What Is Money?* edited by John N. Smithin, vi, p. 276. London; New York: Routledge, 2000.

Dreyfus, Hubert L. "Could Anything Be More Intelligible Than Everyday Intelligibility? Reinterpreting Division I of Being and Time in the Light of Division II," in *Appropriating Heidegger*, edited by James E. Faulconer and Mark A. Wrathall, pp. x, 214. Cambridge; New York: Cambridge University Press, 2000.

—— "Phenomenological Description Versus Rational Reconstruction," *Revue Internationale de Philosophie* 55, no. 1 (2001): pp. 181–96.

—— "The Primacy of Phenomenology over Logical Analysis," *Philosophical Topics* 27, no. 2 (1999): pp. 3–24.

Dupre, John. "Are Whales Fish?" in *Humans and Other Animals*, vi, 272. Oxford: New York Oxford University Press; Clarendon Press, 2002.

Economist, The. "Behaviourists at the Gates," *The Economist* (2003).

—— "Rethinking Thinking," *The Economist* (1999).

Elias, Norbert, Eric Dunning, Johan Goudsblom, and Stephen Mennell. *The Civilizing Process: Sociogenetic and Psychogenetic Investigations*. Rev. ed. Oxford; Malden, Mass.: Blackwell Publishers, 2000.

Fitzpatrick, Scott M. "A Massive Undertaking: Examining Stone Money in Its Archaeological Context," *Antiquity* 76, no. 1 (2002): pp. 331–2.

Foucault, Michel. *The Archaeology of Knowledge.* [1st American] ed. New York: Pantheon Books, 1972.

Garver, Newton. "Philosophy as Grammar," in *The Cambridge Companion to Wittgenstein*, edited by Hans Sluga and David G. Stern. Cambridge: Cambridge University Press, 1996.

Gilbert, Margaret. *Living Together: Rationality, Sociality, and Obligation.* Lanham: Rowman & Littlefield, 1996.

—— *On Social Facts.* Princeton, NJ: Princeton University Press, 1992.

—— "Walking Together: A Paradigmatic Social Phenomenon," *Midwest Studies in Philosophy* 15 (1990): pp. 1–14.

Grossbard, Amyra. "Towards a Marriage between Economics and Anthropology and a General Theory of Marriage," *American Economic Review* 68, no. 2 (1978): pp. 33–7.

Hacking, Ian. *Rewriting the Soul: Multiple Personality and the Sciences of Memory.* Princeton, NJ: Princeton University Press, 1995.

—— "Searle, Reality and the Social," *History of the Human Sciences* 10, no. 4 (1997): pp. 83–92.

—— *The Social Construction of What?* Cambridge, Mass: Harvard University Press, 1999.

Hake, D. F., T. Donaldson, and C. Hyten. "Analysis of Discriminative Control by Social Behavioral Stimuli," *Journal of the Experimental Analysis of Behavior* 39 (1983): pp. 7–23.

Hamblin, C.L. "Questions," *The Australasian Journal of Philosophy* 36, no. 3 (1958): pp. 159–68.

Haugeland, John. "Dasein's Disclosedness," in *Heidegger: A Critical Reader*, edited by Hubert L. Dreyfus and Harrison Hall, p. 303. Oxford, UK; Cambridge: B. Blackwell, 1992.

Hearne, Vicki. *Adam's Task: Calling Animals by Name.* New York: Knopf: distributed by Random House, 1986.

Holzer, Henry Mark. *Declaring War* FrontPageMagazine.com, September 5, 2002 [cited]. Available from http://www.frontpagemag.com/Articles/Printable.asp?ID=2742.

Ingham, Geoffrey. "Modern Money," in *What Is Money?* edited by John N. Smithin, pp. vi, 276. London; New York: Routledge, 2000.

Johansson, Ingvar. *Review of "Social Facts & Collective Intentionality".* Notre Dame Philosophical Reviews, 2003 [cited]. Available from http://ndpr.icaap.org/content/archives/2003/3/johansson-meggle.html.

King, Neil. "Fans Howl in Protest as Judge Decides X-Men Aren"t Human," *Wall Street Journal Online*, January 20, 2003.

Kuhn, Thomas S. *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: University of Chicago Press, 1977.

—— *Structure of Scientific Revolutions*. 3rd ed. Chicago: University of Chicago Press, 1962.

Lin, Chung cheng, Tze wei Chen, and Ching chong Lai. "The Economics of Honeybee Swarming," *Regional Science and Urban Economics* 33, no. 5 (2003): pp. 581–94.

Longino, Helen E. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press, 1990.

Lynn, Michael, and Andrea Grassman. "Restaurant Tipping: An Examination of Three 'Rational' Explanations," *Journal of Economic Psychology* 11, no. 2 (1990): pp. 169–81.

MacIntyre, Alasdair C. *After Virtue: A Study in Moral Theory*. 2nd ed. Notre Dame, Ind.: University of Notre Dame Press, 1984.

Marquis, Christopher. " Absent from the Korea Talks: Bush's Hard-Liner," *New York Times*, September 2, 2003.

Miller, N.J. "Contributions of Social Capital Theory in Predicting Rural Community Inshopping Behavior," *Journal of Socio-Economics* 30 (2001): pp. 475–93.

Morgan, Mary, and Margaret Morrison. "Models as Mediating Instruments," in *Models as Mediators*, edited by Mary Morgan and Margaret Morrison. Cambridge: Cambridge University Press, 1999.

Osborne, Thomas. "The Limits of Ontology," *History of the Human Sciences* 10, no. 4 (1997): pp. 97–102.

Page, Sam. "Searle's Realism Deconstructed," *The Philosophical Forum* 35, no. 3 (2004): pp. 249–74.

Palmer, Arnold. *Playing by the Rules: All the Rules of the Game, Complete with Memorable Rulings from Golf's Rich History*. New York: Pocket Books, 2002.

Parsons, Talcott. *Social Systems and the Evolution of Action Theory*. New York: Free Press, 1977.

Pitcher, George. *Truth*. Englewood Cliffs, NJ: Prentice-Hall, 1964.

Pullum, Geoffrey K. *The Dan Brown Code* Language Log, 1994 [cited]. Available from http://itre.cis.upenn.edu/-myl/languagelog/archives/000844.html.

Ramsey, William. "Prototypes and Conceptual Analysis," in *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, edited by Michael R. DePaul and William Ramsey, pp. xvi, 335. Lanham, Md.: Rowman & Littlefield, 1998.

Russell, Bertrand. "On Denoting," in *Classics of Analytic Philosophy*, edited by Robert R. Ammerman, pp. x, 413. Indianapolis: Hackett, 1990.

—— *The Philosophy of Logical Atomism*. Edited by David Francis Pears. LaSalle, Ill.: Open Court, 1985.

Schumpeter, Joseph Alois. *History of Economic Analysis*. New York: Oxford University Press, 1954.

Searle, John R. *Consciousness and Language*. New York: Cambridge University Press, 2002.

—— *The Construction of Social Reality*. New York: Free Press, 1995.

—— "How to Derive 'Ought' from 'Is'," In *Concepts in Social and Political Philosophy*, edited by Richard E. Flathman, pp. x, 532. New York: Macmillan, 1973.

—— *Intentionality*. Cambridge: Cambridge University Press, 1983.

—— *Mind, Language, and Society: Philosophy in the Real World*. 1st ed. New York, NY: Basic Books, 1998.

—— "Neither Phenomenological Description nor Rational Reconstruction: Reply to Dreyfus," *Revue Internationale de Philosophie* 55, no. 1 (2001): pp. 277–95.

—— *Rationality in Action*. Massachusetts: MIT Press, 2001.

—— "Reply to Thomas Osborne," *History of the Human Sciences* 10, no. 4 (1997): pp. 108–11.

—— *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press, 1969.

—— and Barry Smith. "The Construction of Social Reality: An Exchange," in *John Searle's Ideas About Social Reality: Extensions, Criticisms, and Reconstructions*, edited by David R. Koepsell and Laurence S. Moss, pp. vi, 313. Malden, MA: Blackwell Publishing, 2003.

Shachmurove, Yochanan, Gideon Fishman, and Simon Hakim. "The Burglar as a Rational Economic Agent," *CARESS Working Paper* (1997).

Sintonen, Matti. "Explanation: In Search of the Rationale," in *Scientific Explanation*. Minneapolis: Univ of Minn Pr, 1989.

Smith, Barry. *John Searle*. Cambridge; New York: Cambridge University Press, 2003.

—— and John Searle. "The Construction of Social Reality: An Exchange," in *John Searle's Ideas About Social Reality: Extensions, Criticisms, and Reconstructions*, edited by David R. Koepsell and Laurence S. Moss, pp. vi, 313. Malden, MA: Blackwell Publishing, 2003.

Smithin, John N. "Introduction," in *What Is Money?* edited by John N. Smithin, pp. vi, 276. London; New York: Routledge, 2000.

Stroud, Barry. "The Background of Thought," in *John Searle and His Critics*, edited by E. LePore and R. van Gulick, pp. 245–58. Mass: Basil Blackwell, 1991.

United States Golf Association. *Decisions on the Rules of Golf, 2004–2005.* Far Hills, NJ: USGA, 2004.

Urmson, J. O. *Philosophical Analysis; Its Development between the Two World Wars.* Oxford: Clarendon Press, 1956.

Van Fraassen, Bas C. *The Scientific Image.* Oxford, New York: Clarendon Press; Oxford University Press, 1980.

—— *The Empirical Stance.* New Haven, CT: Yale University Press, 2002.

Weber, Max. " 'Objectivity' in Social Science and Social Policy," In *The Methodology of the Social Sciences.* New York: Free Press, 1949.

Winch, Peter. *The Idea of a Social Science and Its Relation to Philosophy.* 2nd ed. London: Routledge, 1990.

Wisdom, John. "Metaphysics and Verification," *Mind* 47, no. 188 (1938): pp. 452–98.

Wittgenstein, Ludwig. *Philosophical Investigations.* Translated by G.E.M. Anscombe. 3rd ed. New Jersey: Prentice-Hall, 1958.

—— *Remarks on the Foundations of Mathematics.* Translated by G.E.M. Anscombe, Rush Rhees and G. H. von Wright. 3d, rev. ed. Oxford: B. Blackwell, 1978.

Woolf, Virginia. *The Voyage Out.* London: Duckworth, 1915.

Wray, L. Randall. "The Property Theory of Interest and Money," in *What Is Money?* edited by John N. Smithin, pp. vi, 276. London; New York: Routledge, 2000.

Wright, Larry. "The Concept of a Reason," *unpublished* (2005).

—— *Critical Thinking.* Oxford: Oxford University Press, 2001.

—— *Teleological Explanations: An Etiological Analysis of Goals and Functions.* Berkeley: University of California Press, 1976.

Index