# Table of Contents

INTRODUCTION ........................................................................................................ iii

PART ONE — FOUNDATIONS, OLD AND NEW ................................................. 1
1 Individualism: a recipe for warding off “spirits” .............................................. 2
2 Getting to the consensus view ........................................................................... 13
3 Seeds of doubt....................................................................................................... 27
4 Another puzzle: a competing consensus ........................................................... 42
5 Tools and terminology ....................................................................................... 54
6 Grounding and anchoring ................................................................................... 68
7 Case study: laws as frame principles ................................................................. 82
8 Two kinds of individualism ................................................................................. 96
9 Against conjunctivism ....................................................................................... 111

PART TWO — GROUPS AND THE FAILURE OF INDIVIDUALISM ....... 127
10 Groups and constitution ................................................................................. 130
11 Simple facts about groups ............................................................................... 149
12 The identity of groups ...................................................................................... 168
13 Kinds of groups ............................................................................................... 182
14 Group attitudes: patterns of grounding ............................................................ 198
15 Group action: more than member action ........................................................ 220
16 Group intention .................................................................................................. 240
17 Other theories I: social integrate models ........................................................ 255
18 Other theories II: status models ...................................................................... 270

LOOKING AHEAD ............................................................................................... 283

Bibliography ......................................................................................................... 287
PART ONE

FOUNDATIONS: OLD AND NEW
Individualism: a recipe for warding off “spirits”

In the middle of the 20th century, a battle raged between “individualists” and “holists” about the nature of the social world. At the center of these battles was the philosopher Karl Popper. Popper is best known, nowadays, for his work on scientific hypotheses and how they can be falsified. But in his heyday, he was much more famous for his book The Open Society and its Enemies, one of the bestsellers of 1945. In that book, he railed against what he took to be a cancer in social theory: the idea that societies have collective minds that direct their activities. The Open Society describes a menagerie of philosophical villains, all supposed opponents of human freedom. Karl Marx was probably the most wicked, but following close behind was G.W.F. Hegel.

Popper ascribes to Hegel a disturbing view of political systems. He argues that Hegel privileged states, like Germany, Britain, and the Roman Empire, over individuals. Hegel believed, according to Popper, that the interests of states are more fundamental than those of individuals, and that the lives of individuals are in service to the state, rather than the other way around. Behind this, says Popper, was a twisted metaphysics of the state: the idea that states are autonomous, thinking organisms, something above and beyond the people that comprise them. Instead of just taking states to be aggregates of people, this metaphysics ascribes conscious “spirits” to states:

The collectivist Hegel...visualizes the state as an organism... Hegel furnishes it with a conscious and thinking essence, its ‘reason’ or ‘Spirit’.
This Spirit, whose ‘very essence is activity’... is at the same time the collective Spirit of the Nation that forms the state.¹

By Popper’s account, the trouble begins with the distorted idea that the state is a real thing or substance, a whole that exists separately from the people that comprise it. Hegel’s first error, he argues, was to “reify” the state, treating it as a real and separate thing. Once that mystical metaphysics was underway, it became natural to see the state as having interests and values of its own.

It was clever for Popper to choose Hegel as an enemy. The link

¹ Popper 1945, p. 41
between Hegel and Marx made Popper’s book a bible for anti-communists, whose ranks were swelling in the late 1940s. It did not matter much that Popper’s depiction of Hegel was a caricature. Philosophers nowadays do not look kindly on Popper’s scholarship. Even so, we have to give Popper credit for capturing a genuine anxiety of the time, and also a genuine philosophical problem. The issue Popper gives voice to — that many social scientists make use of a worrisome metaphysics — makes sense. Popper was right that social science has found it hard to avoid speaking of society’s “spirit,” and that it was never particularly clear what that spirit was supposed to be. Over the years, many theorists struggled unsuccessfully to avoid social “spirits.” I will mention two, one from the nineteenth century and one from the twentieth.

1. Leopold von Ranke was arguably the premier historian of the nineteenth century. Ranke was the first truly modern historian, pioneering rigorous historical research. Moreover, he was a fierce opponent of Hegel. He rejected the idea that you could tell a single unified history of the world. The histories of different societies and times, he argued, do not belong in one narrative. And, to be sure, history does not unfold in any particular direction. Ranke did admit the possibility of moral and cultural progress, but he did not see a logic to the history of the world. Instead, he professionalized history, transforming it into a specialized field based on empirical investigation. The job of a historian, according to Ranke, is to craft explanations from the detailed investigation of historical particulars in their contexts.

Despite this, Ranke could not figure out how to elude “spirits” in historical explanation. Social epochs, for Ranke, are marked by particular tendencies, and tendencies differ from epoch to epoch. The goal of social research is to uncover those tendencies, as they manifest in particular events. Together, these tendencies reflect a state’s spirit, which is the driving force of human history, much as a person’s soul animates his or her life:

If we now ask, ‘What is it which enables a state to live?’ then it is the same as with the individual, whose life incorporates both body and spirit. So too with the state. Everything depends on spirit, which is the pre-eminent of the two.²

Even the scientifically-minded Ranke was drawn into this position. Nations, states, and societies perform actions. They go to war, they are party to trade deals, they rise and fall. These actions are not coincidences, but occur at least

² von Ranke [1836] 1981, p. 112
in part as a result of the general attributes of the nations, states, and societies in question. Ranke was interested in explaining such historical events. To do so, he appealed to tendencies, but then needed something to account for those tendencies. And so, despite disagreeing with Hegel on so much else, he had little choice but to appeal to “spirit.”

Ranke was not particularly happy about this. The work of history, he argued, is not the “speculation of philosophers.” Historical terms like ‘tendencies’ were empirically grounded, derived from the investigation of facts, as distinct from abstract concepts. Still, he was aware that even though he denied the universal progress of historical epochs, something was required to unify any given epoch and its tendencies. And so the awkward metaphysics of spirits remained.

2. A second illustration comes from Harvard sociologist Talcott Parsons, a figure from the dominant school of sociology in the middle of the twentieth century. In his work, Parsons observed that the actions we take as individuals — whether we are mothers, soldiers, writers, or fraternity brothers — are performed within a cultural structure. Consider a fraternity, for instance. When a fraternity brother hazes new pledges, he is participating in a system largely dictated to him. He does not choose, but rather is pressed into, the hazing traditions. When the brother has some “creative” idea, like making new pledges stand naked in a snowstorm, this creativity is really a minor variant on a prescribed theme.

The key to explaining the member’s action — i.e., his sending the freshmen outdoors — is to explain why the hazing system is in place. Here Parsons has a straightforward answer. Hazing serves a “pattern maintenance” function for the fraternity.3 Its function might be to bind members to one another through common hardship. Or it might be to make the fraternity more attractive and exclusive by creating artificial barriers to entry. In either case, his action helps maintain the fraternity’s patterns of behavior. In performing the ritual, the fraternity brother is playing a role in this structure. The “structure,” then, is a powerful tool for explaining why he does what he does.

Of course, “structure” is not the same thing as “spirit.” By the time Parsons was writing, the social theories of the nineteenth century had been

3 Parsons 1951, 1954
enormously improved upon. Even by Ranke’s time, social theory had disposed of a number of worrisome commitments, and by the end of the nineteenth century, many of the assumptions Ranke had made were also abandoned. For example, Ranke did not manage to shake off the idea that the agents moving the course of history were nations or states, acting something like organisms, with their own unified interests and directions. In the late nineteenth century, theorists gave up on these commitments. Moreover, the religious overtones of social progress had withered as science matured over the nineteenth century. Even the focus of social science on states themselves had faded with the rise of theories of class interests. By 1900, the term ‘spirit’ had fallen from fashion in the social sciences.

Yet even in the middle of the twentieth century, when Popper and Parsons were both writing, the lurking presence of some incomprehensible ectoplasm lived on. Parsons did not talk of “spirit,” but nonetheless he did not manage to find metaphysically secure ground, free of mystical social unities. To many theorists, Parsons’ theory is as imbued with mysticism as Ranke’s, especially in its apparent suggestion that societies somehow direct, or are autonomous from, individual action. For instance, the sociologist Harold Garfinkel, in a biting critique, accused the theory of insulting individual autonomy and individual intelligence. Every day, as individuals we talk with our friends, families, and therapists about how we should live our lives, Garfinkel pointed out. We seem to be guided by our own goals, thoughts, and imaginings. But Parsons’ theory seems to assign genuine agency to social structures, not to individual people. Garfinkel argues that Parsons’ theory portrays us as little more than puppets, who play our role as dictated by society. As he famously put it, the theory turns people into “judgmental dopes.”

Much like Ranke, Parsons found it indispensable, when giving social explanations, to appeal to some kind of social unity that was different or separate from individual people. The reasons for this are clear: how we act and what we do differs radically from one society to another, and from one time to another. These differences are a product of the social contexts in which we are embedded. Those social contexts have properties, they change, they affect us. When we talk about them, we are talking about something.

---

4 Garfinkel 1967
Both Ranke and Parsons seem to have “reified” the social world in order to construct explanations about social phenomena. Ranke found that historical explanations were best given in terms of social tendencies, and Parsons held that explanations for individual action were best given in terms of the social functions that those actions played. Was this sort of move sound, or was it mystical, as Popper argued?

The trouble with social entities

This debate is an instance of a familiar pattern in contemporary philosophy. It does not just arise in connection with individuals and society. Rather, it is one example of a common problem in *inter-level metaphysics*: the discipline which studies the nature of “high-level” or “macroscopic” things, and how they are related to “low-level” or “microscopic” things. (*Macroscopic* and *microscopic* contrast things at intuitively different levels of organization: for example, to contrast macroscopic things like economies and governments [the things treated in *macroeconomics*] with microscopic things like individuals, households, and firms [the things treated in *microeconomics*]. Or to contrast macroscopic things like climates and oceans, with microscopic things like clouds and waves. Or things like proteins and strands of DNA, with things like protons and electrons.)

Is there anything to society above and beyond individuals? The structure of this problem is similar to that of other problems in inter-level metaphysics. We see it in the relation between biology and chemistry: is there anything to life, or to living organisms, above and beyond chemicals? Or in the relation between minds and brains: is there anything to the mind, or to thinking, above and beyond the firing of neurons?

In each of these domains, there seem to be two different stances one could take. First, there is the *dualist* stance: Yes, there is something to living things over and above interacting chemicals, some “vital essence.” Yes, there is something to the mind, over and above interacting neurons, some “soul” or “thinking substance.” Yes, there is something to society over and above interacting individuals, some “spirit” or “social substance.”

---

5 It is problematic even to divide the sciences into “levels.” See, for instance, Thalos 2013; Wimsatt 1976, 1994.
Or there is the monist stance: No, there is nothing to living things, over and above interacting chemicals. Nothing to the mind, over and above interacting neurons. Nothing to society, over and above interacting individuals.

In the philosophy of mind, dualism is most closely associated with Descartes, who distinguished two different kinds of substance: mental substance and "corporeal" or physical substance. Corporeal substance operates according to mechanical rules. And according to Descartes, if there were only corporeal substance, there could be no subjectivity and no thinking; the physical character of the brain alone is not enough to determine the properties of the mind. As soon as Descartes introduced his dualism, nearly 500 years ago, it came under attack, and the attacks have continued more or less nonstop.

Even in Descartes' own day, it was clear that mental substances raise as many problems as they solve. For instance:

1. How can physical matter be guided by physical laws, and yet leave room for the mental to intercede in the course of events?
2. If the mental and physical are separate substances, what is the mechanism for the mental to interact with the physical?
3. How are minds individuated? What makes a mind what it is, and what distinguishes one mind from another?

These sorts of difficulties make it much less tempting to solve the problem by postulating the existence of some kind of separate mental substance.

As alternatives to dualism, philosophers have developed a range of "monist" theories of the mind. In these theories, there is only one kind of stuff — physical stuff — and mental facts are exhaustively determined by physical facts. The relation between physical facts and mental ones can be quite complicated. But in a sense, there is nothing more to the mental "over and above" the physical.

The idea that societies and other social entities are separate and autonomous substances is thus just one species of dualism. And dualism in social theory has the very same problems as it does in theories of mind. If the dynamics of economics, politics, and history proceed through the actions of individuals, how can social laws have any influence on human affairs? Even if they could, by what mechanisms would such things as "structures" influence
FOUNDATIONS: OLD AND NEW

individuals? And how can we tell what “structures” are in the first place? Despite the considerations motivating Ranke, Parsons, and their ilk, it seems like an enormous step backward for us to reify the social world.

**Methodological individualism as the alternative**

If “reifying” social entities amounts to mysticism, what other choice do we have in understanding them? The alternative that developed in the long running battle is “individualism.” Put very roughly (with the promise of refining it as we go), this is the view that the social world is made up of nothing more than individual people. In *The Open Society*, Popper approvingly cites John Stuart Mill, who insists that social phenomena are nothing more than the thoughts and actions of individual people:

> Thus ‘all phenomena of society are phenomena of human nature’, as Mill said; and ‘the Laws of the phenomena of society are, and can be, nothing but the laws of the actions and passions of human beings’, that is to say, ‘the laws of individual human nature. Men are not, when brought together, converted into another kind of substance…’

Popper does not follow Mill slavishly. In fact, much of his book is dedicated to criticizing Mill’s view that we should fold social theory into the discipline of psychology. But Popper does endorse Mill’s thinking about what society is. There is nothing more to facts about society, above and beyond facts about individual psychology.

One of Popper’s students, J.W.N. Watkins, became the dominant voice in favor of individualism in the mid-1950s. Watkins elaborates the claims of individualism in much more detail than Popper did. One of his sharpest polemics is against the sociology of Talcott Parsons. Watkins accuses Parsons of taking a view that is basically theological, and just layering a secular veneer on it. According to Watkins, Parsons does not distinguish his view much from the view that history is guided by divine providence. After all, Parsons does not object to the idea that the actions of individuals are guided by some social entity. All Parsons does, says Watkins, is replace the divine plan with something that seems more scientific. Parsons, like Hegel, is a “methodological holist”:

---

On this view, the social behavior of individuals should be explained in terms of their positions within its cultural-institutional structure, together with the laws which govern the system... This is what is called methodological holism.  

In contrast, Watkins proposes “methodological individualism”: It is people who determine history, however people themselves are determined. This factual or metaphysical claim has the methodological implication that large-scale social phenomena like inflation, political revolutions, etc., should be explained in terms of the situations, dispositions and beliefs of individuals. This is what is called methodological individualism.

Worries about dualism remain, Watkins points out, even if we replace talk of social spirits with social structures, and even if we moderate theories like Parsons’, so that people are not quite so dopey. So long as we talk about social structures and social functions and cultural systems, about bases and superstructures and frameworks of oppression, even about nations and institutions and corporations, we risk treating them as if they are real objects or agents, with intentions, plans and goals, and governed by their own laws or logic. From a metaphysical perspective, all these seem questionable, unlike the more sensible view of those who take the social world to be nothing more than individual people.

As the name suggests, methodological individualism is a view about the proper methodology of the social sciences. Methodological individualists take a certain attitude toward theories, explanations, or models, in the social sciences. They argue that these are best given in terms of individual people: a “holist” theory, like Parsons’, is a bad theory, a poor explanation.

In the 1950s, debate grew feverish between methodological individualists and methodological holists. In that period, a number of promising individualistically inclined theories flowered, with some especially exciting developments in economics. And individualists like Popper and Watkins were persuasive, not only in their insistence that holist theories were predicated on a mystical metaphysics, but also in their claims that holism threatened individual freedom.

---

7 Watkins 1955, pp. 179-80
8 This term was coined by Joseph Schumpeter in 1908, but only gained currency later on.
9 Watkins 1955, pp. 179-80
10 Particularly important were game theory (Nash 1950; von Neumann and Morgenstern 1944) and general equilibrium theory (Arrow and Debreu 1954).
On the other hand, there were plenty of social theorists who seemed to be giving useful and adequate explanations without putting them in terms of individuals. Huge projects in macroeconomics, for instance, were underway at the time, gathering and modeling aggregate measures of economies as a whole.\textsuperscript{11} And holists started coming up with examples of phenomena that seemed particularly resistant to individualistic explanations. In a 1955 article, the philosopher Maurice Mandelbaum described the case of an individual withdrawing money from a bank teller window.\textsuperscript{12} The behavior of the bank teller and the customer, Mandelbaum points out, depends on certain socially defined roles they play, in the circumstances they are in. It is impossible, he argues, to dispense with “societal facts” in explaining their actions. Such were the holist arguments against individualists, to which the individualists of course had further replies.

It was only very late in the course of these debates — once they had already started to run out of steam — that philosophers started to realize the deep confusion at their heart. Neither side had taken one distinction seriously enough: the distinction between individualism about the metaphysics of the social world, and individualism about explanations of social theories.

**Explanatory individualism vs. ontological individualism**

It took until 1968 for people to notice and start untangling this confusion.\textsuperscript{13} In a paper titled “Methodological Individualism Reconsidered,” Steven Lukes inserted a wedge between two different theses that had routinely been conflated in earlier debates.\textsuperscript{14} Methodological individualism, Lukes pointed out, was not just a single thesis: it was two. In fact, it consisted of one controversial thesis and one trivial thesis. The controversial part has become known as explanatory individualism, and the trivial part as ontological individualism.

Explanatory individualism is a thesis about social explanation. It is

\begin{itemize}
  \item \textsuperscript{11} E.g., Klein and Goldberger 1955; Tinbergen 1956
  \item \textsuperscript{12} Mandelbaum 1955, pp. 308-9
  \item \textsuperscript{13} Goldstein 1958 noticed that these needed to be pried apart. Unfortunately, he got tangled in a long discussion of ideal types and “anonymous individuals,” and failed to insert a wedge between the two theses. This was only exacerbated in a fruitless multi-part exchange between Goldstein and Watkins in the British Journal for the Philosophy of Science.
  \item \textsuperscript{14} Lukes 1968
\end{itemize}
the claim that social facts are best explained in terms of individuals and their interactions. That is, theories and models in the social sciences should be individualistic. They should model the properties of individual people and the interactions among individual people.

This, however, is a separate thesis from ontological individualism. Ontological individualism is a thesis about the makeup of the social world. It holds that social facts are exhaustively determined by facts about individuals and their interactions. Ontological individualism says nothing about theories or models or how best to construct explanations in the social sciences. All it says is that there is nothing more to societies, their composition and their properties, above and beyond individual people. Explanatory individualism is a stronger thesis than ontological individualism. Even if societies consisted of nothing more than people, it may be impractical or impossible to construct social explanations individualistically. Even if explanatory individualism is false, in other words, ontological individualism need not be.

Lukes, for his part, rejects the need for (and often, the possibility of) explaining social phenomena in terms of individuals. But he accepts the ontological thesis that social phenomena are fully made up of individualistic ones. In other words, he sides with Parsons and Mandelbaum on explanation, while agreeing with Popper and Watkins on ontology. Lukes begins with a statement of individualistic ontology:

Let us begin with a set of truisms. Society consists of people. Groups consist of people. Institutions consist of people plus rules and roles. Rules are followed (or alternatively not followed) by people and roles are filled by people. Also there are traditions, customs, ideologies, kinship systems, languages: these are ways people act, think, and talk. At the risk of pomposity, these truisms may be said to constitute a theory (let us call it “Truistic Social Atomism”) made up of banal propositions about the world that are analytically true, i.e. in virtue of the meaning of words.  

Lukes’ argument thus begins with a concession about the nature of social facts, about what they “consist of.” He points out, however, that this does not imply that explanations can be given individualistically. He notes that there are many forms of explanation, among which there are perfectly good ones that do not involve individuals at all. And Lukes uses Mandelbaum-style examples to show that in many cases, we should not expect to be able to give

---

15 Lukes 1968, p. 120
strictly individualistic explanations of many social phenomena. Even if facts about bank withdrawals are exhaustively determined by facts about large numbers of individual people, that does not mean that we can construct individualistic explanations of banking.

In the minds of most philosophers, the distinction between explanatory individualism and ontological individualism divides the uncontroversial issues from those that are to be debated. The ontological thesis is settled: there is no more to society, over and above individual people. What remains open is how best to construct social theories and explanations. Are the best social explanations individualistic? Once we sever the link between ontological individualism and explanatory individualism, we can endorse the former and debate the latter.

Having separated these two theses — a controversial thesis about explanation and an uncontroversial one about ontology — philosophers were on their way to a consensus about dualism in social theory: There is no need to be a dualist. Instead, we can be ontological individualists, and still debate explanatory individualism.

They were on their way, but still not quite there. The Lukes paper is not particularly precise. He only paints the distinction between ontological and explanatory individualism in broad brushstrokes. For a time, this was good enough for social theorists, but it soon became clear that we could do better.
Getting to the consensus view

For thirty or forty years now, ontological individualism has quieted our fears of social “spirits.” All sides seem to agree: We do not need to worry about the ontology of the social world. The social world is nothing but people and their interactions. Of course, we can still fight about methodology: ontological individualism does not imply explanatory individualism. Maybe explanations should be individualistic, maybe not. But the ontology is safe.

What, exactly, does it mean to endorse ontological individualism? How do theorists currently understand this claim? In the last generation, a broad consensus has developed, built on the shoulders of two other fields in philosophy — the philosophy of science and the philosophy of mind. From the 1960s to the 1980s, philosophers in those fields made great strides in understanding the relations between macro- and micro-levels.

Three advances in particular merit attention. First is the idea of theory reduction: what does it mean to “reduce” a theory at a macro-level to one at a micro-level? Second is the idea of reduction failure: why might it be impossible to reduce a macro-level theory to a micro-level one? And third is the idea of supervenience: even if we cannot reduce a macro-level theory to a micro-level theory, what does it mean to say that there is nothing to the macro-level “over and above” the micro-level?

These three advances allowed philosophers of social science to move far beyond Lukes’ vague distinction. The contemporary consensus is this: ontological individualism should be understood as a claim about supervenience, and this claim is obviously true. In Chapter 3, I will raise doubts about these conclusions, but first we must understand this claim and the key advances that led to it.

In discussing this material, I have an ulterior motive. As soon as explanatory individualism is contrasted with ontological individualism, many people fall back on a tempting view. Namely, that social facts are “emergent,” that they emerge from individuals and the interactions among individuals.
With this chapter, I want to underline that emergentism is little different from the prevailing, consensus view. When, in subsequent chapters, I challenge the consensus view, I also challenge emergentism. This book is just as much a challenge to the theory that society “emerges” from individuals and their interactions, as it is to other versions of the consensus view.

Reduction

Let’s begin with the idea of a “reductive” explanation. Consider the relation between “macro” and “micro” phenomena in a different sort of system. Consider the behaviors of a school of fish — a huge school of herring, for instance.

Atlantic herring live out most of the year in the North Atlantic, feasting on microscopic creatures. In early winter, these creatures sink into the depths of the ocean, so billions of herring migrate to the Norwegian fjords. (They stay there until January, when they turn back across the Atlantic to spawn.) The fjords are about as protected a place as the herring can find, but their stay is not altogether peaceful. Where herring gather, pods of killer whales follow, attacking in coordinated multi-flank maneuvers. To stay alive, the herring need strategies to protect themselves.

In November 1993, the biologists Leif Nøttestad and Bjørn Axelsen sent a small boat out into the fjords to study these strategies. Using a sonar imaging system, they tracked schools of herring as they responded to whale attacks. A single herring school is enormous, consisting of as many as 50 million fish, and stretching for a quarter mile. Imaging entire schools, Nøttestad and Axelsen found that a school behaves as if with one unified mind when whales approach.¹

When a whale swims at it, a school may split in two, half the fish going left and half going right. Or the school may take a sharp turn in one direction. Or, if it is a particularly large school, it may create a moving vacuole around the whale as it passes through, the herring gathering in propagating waves of density. Or the fish may cluster together in a tight defensive ball, herring packed shoulder to shoulder as if they were already in a jar of cream sauce. A school of herring exhibits a wide range of “defensive”

¹ Nøttestad and Axelsen 1999
patterns, parrying in clever ways with each thrust of the predators. With these countermeasures, the herring can largely escape mass slaughter.

How do the herring manage this melee, knowing as a group whether to cluster, or split in two directions, or rapidly reverse? What makes the fish at the edges of the ball join the cluster, where they are most likely to be eaten, rather than swim off on their own? What makes them so apparently altruistic, sacrificing themselves for the safety of the group?

Back in 1971, the biologist W.D. Hamilton had offered a rather depressing hypothesis to answer questions like these. In “Geometry for the Selfish Herd,” he argued that the members of a herd need not be altruistic for the herd to exhibit apparently coordinated defenses. A number of theorists in the 50s and 60s had used schooling behavior to defend theories of “group selection” — theories in which the evolution of individual traits is explained by the advantages they provide to the group of those individuals. They argued that the gregariousness of individuals, clustering as they do under threat, is mutually protective for the group. Hamilton argued for a simpler mechanism. When a member of a herd or school is threatened, Hamilton suggested, it does one simple thing. It tries to hide. Unfortunately, the only thing for a herring to hide behind is another herring. So every herring tries to dart behind the others. And as they do so, the packing of the school as a whole generates patterns, such as splitting in two, creating a vacuole, or forming a ball. One simple behavior generates a variety of macroscopic patterns in the aggregate.

To evaluate this hypothesis, Nottestad and Axelsen used their sonar to record detailed measurements of the whales and the schools of herring. They recorded the reactions of the schools and the countermeasures they took under what circumstances, comparing the results with computer simulations. To the dismay of many an altruist, they found that Hamilton’s theory was all they needed. No self-sacrifice, no coordination mechanisms, no instinct to cluster. Just a swarm of selfish herring, acting according to a simple rule. All the complex behaviors of the school “emerged” from a large number of individuals interacting with one another.

This sort of theory, should it turn out to be correct, allows us to talk

---

1 Hamilton 1971
2 E.g., Wynne-Edwards 1962, Pitcher and Parrish 1993
about schools of fish and their properties, without taking on any sort of dualism about schools. A Hamilton-style explanation allows us to speak of the school’s “strategies” without worrying that we are speaking of some independent metaphysical realm. The strategies become no more than a shorthand or abbreviation for the complex of individuals that comprise them. This kind of explanation is a “reduction” of the macro-level theory to the micro-level theory.

**Humans, insects, and anti-dualism**

When it comes to human societies, we tend to see ourselves less often as schools of fish than as swarms of bugs. Homer speaks of the Achaeans buzzing wildly like bees around their warships, and Plato takes bees as a moral model of organization and industriousness. Bernard Mandeville uses insects to illustrate the aggregation of sinful individuals into a virtuous society, while Thoreau compares industrial society to the breeding of workers in the abdomen of an ant queen. The nineteenth century philosopher Herbert Spencer argued that both ant colonies and human societies were instances of “superorganisms.” (This view is coming back into vogue in some circles.) And C.K. Ogden, the translator of Wittgenstein’s *Tractatus*, built his nominalist theory of language on the lessons of another book he translated, Forel’s *Social World of the Ants*.

Many aggregates have different properties from those of the individuals that compose them: a pool of water has waviness, viscosity, and so on, while these properties do not apply to individual water molecules. This is not always so obvious in the social sciences. Social scientists are fond of pointing out what we might call “aggregation reversals,” i.e., when a group of people displays the seemingly opposite property that was displayed by the individuals composing it. John Maynard Keynes, for instance, argued for the “paradox of thrift,” where an increase in savings by a large number of individuals in an economy can causally produce a reduction in the aggregate savings of the economy as a whole. Economists are also fond of quoting

---

4 Lattimore 1951, pp. 84-96; Plato 1969, pp. 267-9; Mandeville [1714] 1934, Thoreau [1854] 1966
6 Keynes 1936
Adam Smith’s famous statement, “It is not from the benevolence of the butcher, the brewer, or the baker, that we can expect our dinner, but from their regard to their own interest.” Public welfare, it is suggested, can arise from private selfishness. Mandeville, in his 1705 essay *The Fable of the Bees*, makes the point even more vividly. Where there is a class of vain and lazy nobles, there is demand for herdsmen, weavers, tailors, furriers, and more. If these vices of vanity and laziness were replaced by honesty, temperance, and toil, we might live in a society with greater individual virtue, but collapsing prosperity. In Mandeville’s view, the “diseases” of lust, sloth, avarice, and pride in individuals are essential for a healthy society.

This echoes Hamilton’s “Geometry of the Selfish Herd,” where selfishness at an individual level produces gregariousness in the herd. It is a striking fact that properties of individuals can produce their reverse in the aggregate. But the real insight of explanations like these is not that group benefits arise from individual selfishness. Rather, it is that complex and varied properties of aggregates can arise from simple properties of their members in combination with one another.

Looking down at swarms of pedestrians from atop a skyscraper or at troops of ants from a picnic blanket, we perceive macroscopic regularities that seem to manifest a kind of group coordination or intelligence. The appeal of the analogy between humans and insects is that it helps us to see social phenomena as mere abbreviations for the complex patterns that emerge from potentially simple interactions among individuals. And this defuses anxiety about “reifying” the social world.

**Nagel on theory reduction**

If Hamilton’s and Mandeville’s accounts are successful, they give individualistic explanations of group properties. Ideally, the aim of such accounts was to mathematically derive the “geometry of the herd” from the behavior of individuals. This was exactly the kind of work that mid-century

---

7 Smith [1776] 2006
8 Mandeville [1714] 1934
9 To be precise, Mandeville does not quite present what I have called an “aggregation reversal.” In his essay, it is a subset of the population being dissolute that triggers effects in the rest of the population.
philosophers of science saw as the central quest of the sciences. Explanations like this formed the heart of their program of “the unity of the sciences,” one of whose aims was to put dualism to rest once and for all.

The centerpiece of the program was the notion of theory reduction. Ernest Nagel articulated the classic model of reduction in his 1961 book The Structure of Science.\textsuperscript{10} Reduction is a relation between two scientific theories. A theory, in Nagel’s view, is a set of sentences expressing a set of causal laws, both experimental and theoretical.\textsuperscript{11} A theory about schools of herring, for instance, might consist of causal laws about the behavior of the school under attack from predators:

1. When the school is in configuration A, and a predator attacks from direction B at speed C, the school will pack into a ball.
2. When the school is in configuration D, and a predator attacks from direction E at speed F, the school will form a vacuole around the predator as it swims through the school.

Etc.

A theory of individual herring, in contrast, might consist of laws about the behavior of individuals when they encounter disturbances in their immediate environment:

1. When a herring perceives such-and-such a threat, and is in the presence of some other object, it hides behind that object.

Etc.

To reduce the herring-school theory to the individual-herring theory, we need a set of “bridge laws,” which satisfy two conditions. First, they connect the vocabulary of the first theory to that of the second theory. For every term of the herring-school theory (e.g., ball, vacuole, etc.), a corresponding term is defined in the individual-herring theory. For example, the term ‘ball’ in the school-theory corresponds to a term in the individual-herring theory, defined as a set of configurations of individual herring. The new term is defined so that whenever the individuals are in one of those configurations, the school is in a ball. Second, if we combine the bridge laws together with the laws of the individual-herring theory, it has to be possible to derive from them all the laws of the herring-school theory.

\textsuperscript{10} Nagel 1961. For a more detailed exposition, see Suppe 1977.
\textsuperscript{11} See Nagel 1961, pp. 79-105.
predator approaches school of herring at $t_1$

causal law

herring packed in a ball at $t_2$

defined as

individual herring in such-and-such positions, perceive a threat, react in such-and-such a way at $t_1$

causal law

individual herring swim to such-and-such new positions at $t_2$

defined as

Figure 2A The reduction of “macro” laws to “micro” laws

Hamilton’s geometry of the herd can be construed as a basic instance of this procedure. He begins with a simple macroscopic generalization about the packing of herds under threat. At the individual level, he proposes a mechanism for the interactions of members of the herd. Statements of these mechanisms compose the individual-level theory.

Then he gives a mathematical characterization, in individualistic terms, of the conditions for a collection of individuals to be “tightly packed.” This is the definition of the new individual-level term, which corresponds to a herring-school term, for use in the bridge law. With these theories, Hamilton can derive macroscopic generalizations from the bridge law together with the individual-level theory.

Figure 2A illustrates the relations between the various parts of the theories. In the figure, the horizontal dimension represents time: the dotted arrows connect causes at time $t_1$ to their effects at time $t_2$. Usually in such diagrams, the vertical dimension is ontological: the “macroscopic” facts above stand in some ontological relation to the “microscopic” facts below. That is, they are made of them, or determined by them, etc. This structure will show up many times later on. But in figure 2A, the relation between the macro and the micro is not exactly ontological — instead, it is linguistic: macroscopic terms are defined using microscopic ones.

Since 1961, philosophers have developed devastating objections to Nagel’s approach to theory reduction. The most influential objections to Nagel’s reductionism came from Thomas Kuhn, who argued in The Structure of Scientific Revolutions that different theories reflect different paradigms, are not objective, and are often untranslatable into one another. Others

---

12 Kuhn 1962
challenged the mid-century views that theories involve distinct theoretical and experimental laws, that they consist of laws at all, and that they are linguistic things. Still others focused on the multiple ways that science constructs explanations, denying the centrality of reduction to the project of science. And others remained committed to theory reduction, but objected to the mid-century approach on technical grounds.  

These reactions have led many people to believe that the entire reductive program was a disaster, a dark chapter in philosophy. This assessment is too bleak. Within a few years, the program gave rise to innovations that have reverberated through the sciences, and social sciences in particular. It is one of those encouraging cases in philosophy where a failed program led to a much deeper understanding of the issues.

**Putnam’s insight**

One of the key insights to come out of the program was a new response to anxieties about dualism. As I mentioned, theory reduction gives an answer to dualism. Schools of fish do not need to be treated as separate objects existing in a separate and real sphere, over and above the individual fish. Instead, any statement about schools of fish can be reduced to statements about individual fish. Thus reductionism gives a satisfying answer to the dualist challenge.

In the early 60s, Hilary Putnam — who earlier had been one of the most forceful advocates of “the unity of the sciences” — began to realize two things.  

First, that theory reduction was actually exceedingly difficult to carry out, and rarely if ever done. It was particularly not so easy to construct “bridge laws.” And second, that reduction was not really needed to respond to anxieties about dualism.

One of Nagel’s many accomplishments in *The Structure of Science* was to clarify the several disparate elements that must come together in a theory reduction. He noticed that bridge laws needed to satisfy a number of conditions in order for reduction to be successful. Among them was the condition he called “connectability”: that every term in the high-level theory T1 must correspond to its own definition in the reducing theory T2. Nagel

---


14 Putnam 1967, 1969
also pointed out that the connections between terms given in the bridge laws could have a variety of strengths.¹⁵

Putnam realized, however, that even though connectability might be necessary for reduction, it was more than what we need if all we want is to deny dualism. Even if reduction is impossible, we can still open a window to avoid dualism by denying connectability.¹⁶ After all, connectability is a tall order: it insists that every term in the high-level theory be defined in terms of the lower. As a practical matter, this has almost never been successfully accomplished. Explicit definitions of scientific terms are devilishly hard to produce.¹⁷ But we do not need such definitions in order to avoid dualism. Even if we cannot have bridge laws between theories T1 and T2 that satisfy “connectability” — and so cannot reduce T1 to T2 — we still might have reason to think that the laws described by T1 as a whole are taken care of by T2 as a whole.

Putnam argued that in many sciences, we should not even expect connectability to hold between theories at the high-level and theories at the lower level. Sometimes, Hamilton-style reductive explanations might be possible, but often they are not. This is not because the objects or phenomena at the high-level are metaphysically distinct, over and above the low-level stuff. It is only that the terms of the theory at the high-level cannot be defined in terms of the low-level theory. What might get in the way of these definitions? The fact that the high-level properties are “multiply realizable.”

Multiple realizability

Putnam’s idea — which even today is probably the most widely endorsed argument against reduction — was that certain high-level properties might be “multiply realizable” at the lower level. A screwdriver, for instance, is a thing that turns screws. It might, however, be made of all different kinds of physical materials. It might be made of steel, or wood, or hard plastic, and still function to turn screws. Suppose we have a theory about screwdrivers. For instance, a theory about how different-sized flathead screwdrivers function to

¹⁶ Davidson 1980 presents a different influential argument for denying reducibility.
¹⁷ Even the ones that appear somewhat straightforward are not, such as the reduction of thermodynamics to statistical mechanics. Cf. Callender 1999.
turn different kinds of screws. What is the physical theory that this screwdriver-theory corresponds to? A theory about the properties of steel? Of wood? Of plastic? There is no single correspondence between the term ‘screwdriver’ and any term in any one physical theory.

Jerry Fodor filled out this argument in more detail. In contrast to the neat figure 2A above, Fodor drew the following diagram, representing the messy relations between laws at different levels. There is not a simple correspondence between “high-level” properties and “low-level” ones. Rather, high-level properties are often realized as hugely complicated and messy properties at the lower level:

```
Figure 2B  Multiple realizability
The bottom part of figure 2B is a wildly disjunctive set of causal connections. It is not a low-level causal law. Because there is no single connection between high-level and microphysical terms, the connectability condition is not satisfied. Thus the multiple realizability of properties at the high level blocks Nagel-style reduction.

Still, none of this means that we need to postulate any separate and distinct “realm” of entities at the high level. Because high-level properties are multiply realizable, reduction fails. Nonetheless, any given high-level state is realized in some low-level state. Any given screwdriver is made out of some particular material, not “screwdriver substance.” Any given mental state is made out of some particular state of the brain, not “thinking substance.” And any given social state is made out of some particular state of the population, not “social substance.” Even if we cannot construct a theory reduction, we
```

---

18 Fodor 1974
can still reject dualism.

This Putnam-Fodor argument is controversial. It has been challenged on several fronts. Many people deny that most high-level properties can be analyzed in terms of functions such as serving to turn screws. This may undermine the Putnam-Fodor argument for the multiple realizability of these properties.\(^{19}\) Also, the notion of multiple realizability has itself come under serious criticism in recent years.\(^{20}\) And, finally, a number of theorists have taken innovative approaches to reduction, so that it does not require connectability.\(^{21}\) So it might be that the Putnam-Fodor argument does not actually rule out reduction, even though it rules out Nagel-style reduction.

In this complex back-and-forth, however, the basic point should not be lost. We can avoid dualism, or end up with an innocuous version of it, whether or not macroscopic theories can be reduced to microscopic ones.\(^{22}\) This is a key legacy of the mid-century program in the unity of the sciences.

The Putnam-Fodor arguments dramatically clarify the point Steven Lukes made in “Methodological Individualism Reconsidered.” Lukes distinguished explanatory individualism from ontological individualism, but did little more than point out that the two were distinct. Now we have a reason why certain reductive explanations may be impossible, even if there is nothing to the high-level phenomena over and above the low-level ones. It becomes clear how we might insert the wedge between explanatory individualism and ontological individualism.

**Anti-dualism and supervenience**

By the mid-1970s, a strategy for avoiding an unpalatable dualism was taking shape. Avoiding dualism does not require that we correlate every high-level property with a low-level property. We do not need to be able to translate social theories into theories about individual people. Instead, all we need is that the social properties are collectively “nothing over and above” the properties of individuals.

\(^{19}\) Theories of functions and functional properties have also changed radically since Putnam’s day. See Ariew et al. 2002.

\(^{20}\) See, for instance, Shapiro 2000.


\(^{22}\) This result is still controversial. See, for instance, Kim 1989.
The next fifteen years saw the development of a powerful new tool for making this “over and above” phrase much more precise: the notion of supervenience. Supervenience is a relation between two sets of properties. Take property set A to be all the social properties, and property set B to be all the individualistic properties. To say A supervenes on B, then, is to say an object cannot change its A-properties without there being some accompanying change in its B-properties. Or to put it more intuitively, *the B-properties fix the A-properties*. Once all the individualistic properties are in place, that fixes the social properties.23

To illustrate, consider the pictures on a TV screen. These supervene on the pattern of illumination of its pixels. The picture on the TV screen cannot change without some change in the pattern of illumination of its pixels. There is nothing to the TV picture over and above the illumination of individual pixels. Notice, however, that there need be no lining-up of individual A-properties with individual B-properties. We might have a TV image — say, an image of a ball or a tree — generated by many different patterns of illumination of pixels, on many different types of screens.

There is not just one supervenience relation, but a family of them. One member of the family is “local supervenience.”24 Local supervenience makes a claim about each object in the world. To say that A locally supervenes on B is to say that for each object, if you fix its B-properties, then you have fixed that object’s A-properties. More relevant for our purposes will be “global supervenience.” This is a weaker claim. It does not insist that the B-properties of each object fix that object’s A-properties. Instead, it only makes a claim about the spread of properties across the entire world. To say that A globally supervenes on B is to say that if you fix all the B-properties in the entire world, then you have fixed the A-properties in the world.25

Supervenience turns out to be a hugely useful family of diagnostic tools, like X-rays or MRIs, for evaluating the relationships between “macro”

23 See Kim 1984, 1987. Also, McLaughlin and Bennett 2005 is an excellent overview of varieties of supervenience.
24 This is often called ‘individual supervenience’, but that term would be confusing in the present context.
25 It has long been recognized that social properties fail to locally supervene on individualistic properties. This was first pointed out by Currie 1984. Though his conclusion is correct, however, his argument does not quite work. For a rigorous argument, see the appendix to Epstein 2011. In Chapter 8, I give a more precise formulation of global supervenience.
and “micro” properties. As a diagnostic tool, supervenience helps show how we may be able to dodge the threat of dualism.\(^{26}\)

It may be impossible to reduce social theory to a theory of individual people. But all it may take to avoid dualism is for a weaker relation to hold — a supervenience relation — between the whole set of individualistic properties and the whole set of social properties. This is the strategy endorsed by “non-reductive individualists.”

To address the problem of dualism in the social sciences, the supervenience of the social on the individualistic quickly became the consensus response. It gives theorists a way to be ontological individualists, even if they are skeptical about explanatory individualism.

**Contemporary ontological individualism**

Much of this work, on reduction, multiple realizability, and supervenience, was done in the philosophy of mind. But that work has dovetailed nicely with discussions of methodological individualism in the social sciences. For instance, it is standard nowadays to argue against explanatory individualism in social science on the grounds that social properties are multiply realizable.

In recent years ontological individualism has come to be interpreted as a supervenience thesis: the social properties globally supervene on the properties of individual people.\(^{27}\) The individualistic properties exhaustively determine the social properties, even if there is no way of producing a correspondence between a given social property and one or more individualistic properties. A representative statement is Harold Kincaid’s:

> Social wholes are both composed of individuals and determined by their actions… Individuals determine the social world in the intuitive sense that once all the relevant facts (expressed in the preferred individualist vocabulary) about individuals are set, then so too are all the facts about social entities, events, etc. Or, to put this idea in terms of supervenient

\(^{26}\) Supervenience is not a flawless diagnostic tool, but can offer excellent evidence. A crucial shortcoming is that it does not suffice for full grounding. (I discuss grounding in Chapter 5, and the limitations of supervenience in Chapter 8.)

properties, the social supervenes on the individual in the sense that any two social domains exactly alike in terms of the individuals and individual relations composing them would share the same social properties.\textsuperscript{28} Like Lukes, most people use some statement of ontological individualism as a quick prelude, on the way to discussing the obstacles to explanatory individualism. Sometimes it is accompanied by a reminder of Mandeville’s point that aggregates can have emergent properties that none of the individuals do.\textsuperscript{29}

This is the consensus view. Ontological individualism is distinct from explanatory individualism. Explanatory individualism is debatable. Some problems might be susceptible to individualistic explanation, while others are probably not. Ontological individualism, on the other hand, is much more straightforward. It is best understood as a supervenience thesis, and it is obviously true.

To be sure, ontological individualism was never intended to solve the central problems of social theory. It does not give us any indication of how to model the individual. It does not solve the problems of aggregation. It does not even suggest how to approach the longstanding debates between structure-centered and agent-centered explanations in the social sciences. The aim of the division between explanatory and ontological individualism was to free theorists from the anxieties of dualism: to clear the ring, so that more sensible fights over explanatory methodology could begin.

\textsuperscript{28} Kincaid 1986, 1998
\textsuperscript{29} E.g., Pettit 2003, p. 191
Despite the appearance of being settled, there are reasons to worry about ontological individualism. Social theorists often assume that, with a little thought, the wrinkles will be easy enough to iron out. But this attitude is too cavalier. Where theorists do make specific claims about the composition of social entities out of individuals, they tend to go wrong. Consider, for instance, Lukes’ “truisms,” his “banal propositions about the world that are analytically true”:

Society consists of people. Groups consist of people. Institutions consist of people plus rules and roles. Rules are followed (or alternatively not followed) by people and roles are filled by people. Also there are traditions, customs, ideologies, kinship systems, languages: these are ways people act, think, and talk.¹

Banal they may be, but are they analytic? Are they truistic? Are they even true?

Or consider the quotation from Kincaid in the previous chapter. It is fairly clear and explicit, about as detailed as any statement of ontological individualism. But on closer look, it is vague. It says, for instance, not a thing about what gets included in “individuals and individual relations,” “how things are with and between individuals,” or “all the relevant facts (expressed in the preferred individualist vocabulary) about individuals.” In the literature on the social world, almost no one talks about why social facts are supposed to supervene on individualistic ones, or even about what facts or properties count as individualistic in the first place.

In this chapter, I begin to argue what may seem like a radical claim: The contemporary consensus is mistaken. Ontological individualism is false. The social facts do not supervene on the individualistic ones. My aim in this particular chapter is to show this intuitively. Denying ontological individualism does not mean endorsing “ontological holism.” It does not

¹ Lukes 1968, p. 120
mean endorsing “emergentism.”\textsuperscript{2} What is wrong with ontological individualism is that it is a stronger thesis than many people have realized.

I will start by discussing the failure of an analogous thesis in a different science. It is an actual historical episode, one that is interesting in its own right: the nineteenth century “cell theory” of organisms. Then I will apply that case back to the social sciences.

\textit{Why be skeptical about ontological individualism? An analogy.}

As students, we all learned the principles of cell theory: \textit{All living things are made up of cells. All cells come from preexisting cells by division. The cell is the structural and functional unit of all living things.} Cell theory is one of the great accomplishments of modern science, and was surprisingly hard won. Although Robert Hooke discovered and named cells in the 1650s, it took two hundred years for people to recognize their biological importance, and to develop these simple principles.

As responsible as anyone for this accomplishment was Rudolf Virchow, a German biologist, anthropologist, doctor, and politician. Virchow was prolific in the way that only nineteenth century Prussians could be: a bibliography of his works runs 113 pages, his masterworks being the handbooks on pathology from the 1850s and 60s. In \textit{Cellular Pathology} of 1860, he catalogued the systems of the human body, presenting the cellular anatomy of each and describing the diseases and degenerations of the systems in terms of cellular transformations. He began the book with a clear statement of the primacy of cells as the building blocks of both plant and animal life:

\begin{quote}
Every more highly developed organism, whether vegetable or animal, must be regarded as a progressive total, made up of larger or smaller numbers of similar or dissimilar cells. Just as a tree constitutes a mass arranged in a definite manner, in which, in every single part, in the leaves as in the root, in the trunk as in the blossom, cells are discovered to be the ultimate elements, so it is with the forms of animal life. \textit{Every animal presents itself as a sum of vital unities}, every one of which manifests all the characteristics of life.\textsuperscript{3}
\end{quote}

\textsuperscript{2} Some commentators have misunderstood this, e.g. Hindriks 2013, p. 432; Guala and Steel 2011, p. 282.

\textsuperscript{3} Virchow 1860, pp. 13-14.
At the time, Virchow was engaged in a struggle against theories that now seem so old-fashioned it is hard to imagine them ever being taken seriously. Although scientists had mostly abandoned ancient theories of life, such as vitalism and humor theories, cell theory was not the only alternative. The physicalist school, for instance, regarded organisms as having “formative forces.” Others devalued the importance of the cell, regarding the protoplasm as the basic substance of organisms. Obviously, cell theory won this struggle. But in the wake of its success, we can miss the fact that the above passage from Virchow is not quite true. It sounds plausible. But if we actually take a look at organisms, whether possums or people, it is clear that they are not a sum of cells, not even mostly.

Any organism includes a lot of extracellular material. An average human body has fifteen liters of fluid, floating and pumping around in various places. There is interstitial fluid between cells, blood plasma, gastrointestinal fluid, cerebrospinal fluid, ocular fluid, joint fluid, and urine. Metabolites, ions, proteins, neurotransmitters, and hormones flow between cells. And then there are big chunks of solid anatomical stuff that are not made up of cells either. Bone matrix — something that plays a rather important function for humans — makes up about 15% of body weight, but is not composed of cells. Nor are teeth. Nor are eyes — which are mostly made up of a transparent gel. Even organs like the lungs are largely built of connective tissue, consisting of fibrous proteins and collagenous compounds.

Imagining an excellent simulation, on some futuristic computer, of all and only the cells in a human body, leaving out everything that was not a cell. Red and white blood cells would be represented, but without plasma for them to travel in. There would be neurons, but no way for them to communicate with one another. There would be muscles, but nothing to connect to and pull on. There would be no chewing, no seeing, no digesting or excreting. An excellent simulation of cells would be a terrible simulation of the body.

Virchow was, of course, aware of this problem. He noticed that many tissues were made up of more extracellular material than cells. In fact, he noticed that certain tissues had very few cells at all. Nonetheless, he was committed to the principles of cell theory. How was he to deal with

---

1 Drysdale 1874; Fletcher 1837. See Mayr 1982 for a fascinating discussion of views at the time.
2 E.g., connective tissue and “mucous” tissue (Virchow 1860, pp. 41-48).
extracellular material?

Theodor Schwann, a few years earlier, had proposed a solution to this very problem, but Virchow found it unacceptable. Schwann suggested that the extracellular material in an organism was the stuff of proto-cells, the “cytoblastoma destined for the development of other cells.” It was clear to Virchow that this could not be right. Much extracellular material was destined for nothing of the sort.

So Virchow took a different strategy. Although it is true that extracellular material lies outside of cells, Virchow argued that all the extracellular material in an organism is apportioned to the cells that govern it. Cells are proprietors of “cell-territories”:

I have, by means of a series of pathological observations, arrived at the conclusion that the intercellular substance is dependent in a certain definite manner upon the cells, and that it is necessary to draw boundaries in it also, so that certain districts belong to one cell, and certain others to another... Any given district of intercellular substance is ruled over by the cell, which lies in the middle of it and exercises influence upon the neighboring parts.

In other words, the matter between cells is divided up by boundaries, according to the cells into whose districts they fall. Virchow seems to have in mind something like the division of Lake Superior into the portion belonging to the United States and the portion belonging to Canada. Just as the two countries have “superintendence” over their respective portions of the lake, cells rule over their portion of extracellular material.

Virchow holds that extracellular material has implicit boundaries, according to the cells that have purview over it. Organisms do not only consist of the interiors of cells, but of the interiors together with the exterior parts that belong to them. Thus the principles of cell theory remain unsullied: organisms are exhaustively composed of cells conceived of as cell-territories.

All in all, it was a good idea for Virchow to be as rigorously committed to cell theory as he was. He emphasized that we should see the

---

6 It is an interesting question, why these biologists were so committed to cells exhaustively composing the body.
8 Virchow 1860, pp. 15-16.
9 Virchow 1860, pp. 15-16.
body as composed of innumerable vital parts. Virchow’s erroneous ontology was much better than the earlier erroneous ontology. But this “cell-territory” strategy is a stretch. Even if we were to believe Virchow about the governance of some extracellular material by cells in some cases, it could not apply in general. Which are the cells governing the gelatinous goo in the eye? Which cells govern the bone matrix, or the digestive fluids sloshing in the stomach? (Analogously, we might ask what countries have “superintendence” over the middle of the Pacific Ocean.)

To be sure, cells are important parts of the human body, maybe even the most interesting and dynamic parts. Certainly it would be a mistake to devalue their role, as some of Virchow’s nineteenth century opponents did. But the body is only partially composed of cells. There are many basic functional components apart from cellular ones.

Virchow’s claim is not unlike a botanist claiming that trees are composed of leaves. (In fact this claim, bizarre though it seems, was defended at length by Goethe in *The Metamorphosis of Plants.*10) No doubt, leaves are an important part of trees. Without leaves, most trees would be in trouble. But they are only part of the story. Botany is not just leaf-ology. Nor is anatomy just cytology.

Notice that the failure of anatomy to be exhaustively determined by cytology does not involve any remarkable claims about anatomical “spirits,” dualism, or ghostly ectoplasm. We would not expect anatomy to be exhaustively determined by dermatology or nephrology. There is more to the body than the skin and the kidneys. The source of supervenience failure is that the “cellular facts” are too limited to exhaustively determine the “anatomical facts.”

How bad ontology leads to bad science

Virchow’s flawed ontology damaged his scientific practice. Because he insisted on the exhaustive determination of anatomical facts by cellular facts, he committed himself to thinking of the body as divided into cell-territories.

---

10 “It came to me in a flash that in the organ of the plant which we are accustomed to call the leaf lies the true Proteus who can hide or reveal himself in all vegetal forms. From first to last, the plant is nothing but leaf, which is so inseparable from the future germ that one cannot think of one without the other.” (Goethe 1962, pp. 258-59)
And this led him onto some radically mistaken tracks when it came to certain theories.

Take, for instance, his theories of the formation and degeneration of bone. In Lecture XVIII of *Cellular Pathology*, Virchow proposes a theory of the formation of bone out of cartilage. Cartilage, according to Virchow, consists of cartilage-corpuscles, each of which is a “territory” consisting of a cartilage cell plus the “capsule” in which it is contained. A bone-corpuscle, then, is formed when a cartilage-corpuscle becomes calcified. Although the cell is transformed, the territory remains the same:

The old limits of the capsule still represent the real district which is under the sway of the bone-corpuscle... Within these limits we see the bone-corpuscle accomplish its peculiar destinies.

We see the reverse process occurring, according to Virchow, in bone necroses. In degenerative ostitis, for instance, bone-corpuscles transform back into other sorts of cells:

The bone first produced and proceeding from cartilage may undergo a transformation into marrow, then into granulation-tissue, and finally into nearly pure pus.

As scientists discovered in the intervening century, there are many ways bones can degenerate. But Virchow’s proposal describes none of them. When bone deteriorates, it does not change from corpuscle-sized regions of bone into marrow and other types of cell. His view of bone growth is likewise mistaken. There are, of course, cells involved in bone growth: osteoblasts act as “construction workers,” depositing layers of bone. Nonetheless, a bone is no more made up of osteoblasts than a roof is made up of roofers.

Virchow’s commitment to his cell-territory version of cell theory derailed his explanations. To explain bone necroses as transformations of cell territories into other cell territories is not just awkward or psychologically misleading, but is something between a distortion and a flat mistake. Ontology has ramifications, and ontological mistakes lead to scientific mistakes. Commitments about the nature of the entities in a science — how they are composed, the entities on which they ontologically depend — are woven into the models of the science. The errors of Virchow’s cytocentric

---

11 Virchow 1860, pp. 415-18
12 Virchow 1860, pp. 417-18
13 Virchow 1860, p. 422
approach to anatomy are easy to see, as are the scientific failures that resulted. Despite Virchow's expertise with a microscope, his commitment to cell theory led him to subdivide tissues into cells where there are none. And that led to poor theories about how anatomical features come to be, how they are changed or destroyed, and what they do.

**Bringing it back to social theory**

Virchow’s approaches, strategies, and failures have close parallels in the contemporary social sciences. Take, for instance, the implicit ontology built into one of the best known contemporary frameworks for social theory: the “boat diagram” developed by James Coleman.

In *Foundations of Social Theory*, Coleman observes that social theory typically intends to explain phenomena at the macroscopic level. He takes as a paradigm Max Weber’s argument that the development of capitalism in the Western world can largely be explained by the influence of Protestant religious doctrine. Coleman breaks Weber’s argument into three steps:

1. Protestant religious doctrine generates certain values in its adherents.
2. Individuals with certain values (referred to in proposition 1) adopt certain kinds of orientations to economic behavior. (The central orientations to economic behavior are characterized by Weber as antitraditionalism and duty to one’s calling.)
3. Certain orientations to economic behavior (referred to in proposition 2) on the part of individuals help bring about capitalist economic organization in a society.\(^{14}\)

Coleman points out that in this argument Weber is explaining the transition of society from one “social-level” feature to another “social-level” feature, i.e., the transition from Protestantism to Capitalism. But to explain this transition, Weber moves to the “individualistic level.” Weber’s narrative is not a good explanation, Coleman suggests, if he does not show how the transition was mediated by the activities of individuals. This means that Weber needs to demonstrate all three propositions: how the doctrine generates the relevant values at the individual level, how the individual-to-individual transition of values leads to a certain economic behavior, and how the individual economic behavior generates capitalism at the level of the

\(^{14}\) Coleman 1990, p. 8.
society. Coleman depicts this with the following boat-shaped diagram:

**Figure 3A The boat diagram of Weber’s Protestant Ethic**

The two nodes at the top of the diagram are macroscopic, social phenomena. The two at the bottom of the diagram are phenomena involving individual people. The numbered arrows roughly correspond to the three numbered propositions.

Coleman thinks Weber does a good job with arrow 1, explaining how Protestant religious doctrine leads individuals to have certain values, and with arrow 2, explaining how those individuals’ values induce them to behave in characteristic ways in the economic sphere. But he complains that Weber pays inadequate attention to arrow 3. Having paid such careful attention to how Protestant values lead individuals to save and invest, Weber neglects the problem of aggregating the economic behavior of individuals into a capitalist system. This, Coleman stresses, is often the crucial problem for social theory: explaining how the systematic behavior of individuals generates systemic macro-phenomena.

Coleman’s scheme for social explanation is more modest than Watkins’s individualistic strategy discussed in Chapter 1. Entities at the social level need not be eliminated, or analyzed in individualistic terms. Coleman starts his social explanations with a background of institutions that are already in place. Only the transitions marked by the arrows need to go through individuals.

*Applying boat diagrams to cell theory*

Nonetheless, Coleman’s approach presupposes a quite specific ontology. To see this, consider what happens if we apply Coleman’s diagram to the kinds

---

of cases Virchow addressed. The first case will be a phenomenon for which
cell theory is adequate, and the second, a phenomenon for which it is not.

For the first, consider the organism-level phenomenon: *Being
deprived of water causes a plant to wilt*. We might construct the following
diagrammatic explanation:

![Figure 3B](image)

*Figure 3B  A successful boat-style explanation at the cellular level*

There is an easy transition from the organism-level cause, *being deprived of
water*, to the cellular-level effect, that cells cannot replace the water they lose.
Arrow 2 represents causes that take place entirely at the cellular level, where
water loss causes shriveling. Arrow 3 represents the plant-level effect of
shriveling at the cellular level. This is the sort of explanation that Virchow
might have successfully produced.

Contrast this, however, with the sort of case that trips up Virchow’s
version of cell theory. Consider the organism-level phenomenon: *Being
deprived of fluoride causes a child to get cavities*.

![Figure 3C](image)

*Figure 3C  A failed boat-style explanation at the cellular level*

In this diagram, we again employ a causal transition from the organism level
to the cellular level, one within the cellular level, and one from the cellular
level back up to the organism level.

But this diagram is absurd. The problem, of course, is that tooth
enamel is not made of cells. So it is not only this explanation, involving
fluoride ions reaching tooth cells, that is spurious. It is unlikely that there will be any explanation at the cellular level for tooth decay. The diagram implies that what is predominantly causally important for tooth decay is cellular. This is something Virchow might have believed, but we know better.

Now, the advocate of the boat diagram might object that I am being too literal. Here are some natural reactions one could have:

1. When we speak of the “cellular level” in an explanation of tooth decay, we mean anything cell-sized. Even though there aren’t tooth cells, there are cell-sized structures, and we can explain tooth decay in terms of those structures.

2. When we speak of the “cellular level,” we don’t really mean the cellular level. We mean the microscopic level, the level of whatever parts the tooth is made up of. The point, after all, is that there is some lower level explanation for tooth decay, and that explanations in lower level terms are good ones.

3. When we speak of the “cellular level,” we mean not just cells, but anything that interacts with cells. And there are cells involved in tooth decay — epithelial cells, gum cells, nerve cells, cells of the bacteria the form plaque, and so on. All of those are at the cellular level, and so tooth enamel is as well.

These are all variants on a theme. The first reaction is particularly reminiscent of Virchow’s cell-territory response. It did not work when we said that explanations had to go through cells, so we expand what we mean by ‘cell’. But the other reactions are similar. They all ask us to take ‘cellular’ with a grain of salt.

To a limited extent, these reactions make sense. We do not want to be so uptight as to rule out pragmatic explanatory strategies like Coleman’s boat just because we do not like their labels. Still, it is fair to ask, if we can be as flexible as we like in what we include in the “cellular” level, then in what sense is the explanation an explanation in terms of cells? We can twist words, but the facts are the facts: when it comes to tooth decay, not much happens at the cellular level. What is interesting is happening at large-scale levels, such as pits forming in tooth enamel, and at very small-scale levels, such as at the level of interactions between enamel and acids. But teeth are not composed of cells. This means that figure 3C is misapplied, not that we should start
playing games with what counts as “cellular.”

To endorse a boat-style explanation is to endorse an ontology. If a good explanation of high-level phenomena has to go through the lower level, that presupposes that facts at the lower level exhaustively determine those at the higher level.\(^\text{16}\)

What goes for cellular boat diagrams goes for individualistic boat diagrams as well. Coleman praises Weber’s overall strategy in the *Protestant Ethic*, faulting it only for its failure to explain inference 3. Implicit in this, however, is a presupposition: the phenomena at the social level are more like wilting plants than like decaying teeth. But if social phenomena are not exhaustively determined by facts about individuals, then we should not expect an explanation that goes by way of individuals to work. His explanatory strategy carries with it a commitment to a particular ontology of the social world. If that ontology does turn out to be mistaken, then the explanatory strategy will have to be revised or abandoned.

*An intuitive failure of individualism*

Coleman’s explanatory strategy carries ontological presuppositions. If social facts are not exhaustively determined by facts about individuals, it does not make sense to insist that social explanations should conform to the boat diagram. But why think that the claim *social facts are exhaustively determined by facts about individuals* is analogous to the erroneous claim *anatomical facts are exhaustively determined by facts about cells?*

Here is a reason: consider facts about the Starbucks Corporation. On a typical day at Starbucks, pots of coffee are being brewed, baristas are preparing frappuccinos, cash registers are ringing, customers are lining up, credit cards are being processed, banks are being debited and credited, accountants are tallying up expenses, ownership stakes are changing in value, and so on. At least on the face of it, some of these facts about Starbucks fail to supervene on facts about the people and their interrelations. To be sure, the employees are critical to the operation of Starbucks. But facts about Starbucks seem also to depend on facts about the coffee, the espresso machines, the business license and the accounting ledgers.

\(^{16}\) It actually assumes much more: that the entire causal network that relevantly affects the system is also exhaustively determined by the lower level of entities.
FOUNDATIONS: OLD AND NEW

Consider what we might want to accomplish in a model of some changing property of Starbucks. Suppose we were to model it through some sort of unfortunate event. Suppose, for instance, there is a freak, late-night power spike at a number of Starbucks outlets, causing the blenders and refrigerators to break, the ice to melt, and the milk to spoil. Suppose that event is the last straw for a financially struggling Starbucks, underinsured as it is. So, when the power spikes and its key assets melt down, its assets no longer exceed its liabilities. Overnight, as the owners, employees, and accountants are asleep in their beds, Starbucks goes from being financially solvent to insolvent.

In this example, the transition to insolvency involves property and equipment, not individuals. It is analogous to the tooth-decay argument in this respect. At least at first blush, it is not individuals, or phenomena at the “individualistic level,” that explain this social-level transition. If this is right, a Coleman-style explanation for the transition in terms of social facts would be impotent.

In Figure 3D A failed individualistic explanation of Starbucks

Like the explanation of tooth decay, the explanation of Starbucks’ transition from one macroscopic state to another does not make sense, if given at an irrelevant or incomplete microscopic level. It is not, of course, that no people at all are involved in the ordinary course of Starbucks operations. Nor is it that there are no cells involved in the ordinary course of anatomical operations. But just as that fact does not entail that anatomical explanation should go through the cellular level, neither does the involvement of people in Starbucks entail that explanations of its states should go through the individualistic level.

Also analogous to the anatomical case is the point about simulation. Imagine an excellent simulation, on the same futuristic computer, of all and only the people involved in Starbucks — employees, customers, board
members, even vendors — but leaving out everything that is not a person. It might be organized something like this:

![Diagram of multiscale organization with employee subagents](image)

Figure 3E Multiscale organization with employee subagents

In this simulation, there would be baristas, but no coffee to drink. Customers, but no chairs or wi-fi connections. Cashiers, but no cash. However excellent this might be as a simulation of the people, it would be a terrible simulation of the company.

And also as in the tooth decay case, there are several Virchow-style reactions one might have. This is a common move in social theory: we bend what is meant by “individualistic.” When we speak of the “individualistic level,” we don’t really mean the individualistic level. We mean individual-sized things, whatever they are. Or we mean the microscopic level, the level of whatever parts society is made up of. Or we mean anything that causally interacts with individuals. In short, just as Virchow slides from cells to cell-territories, we slide from individuals to individual-territories, or else from individuals to things that are not in any respect individualistic.

Such are the moves social theorists have implicitly taken. Even Watkins did this, in the later part of his career. Back when he was debating Mandelbaum in the 1950s, he argued that social facts are entirely composed of the psychological states of individual people. Over time, however, he realized that psychological states were an inadequate base for determining facts about the social world. After all, social phenomena involve behaviors and actions, not just thoughts. It is absurd to think that social facts — like actions performed by Starbucks Corporation — are nothing more than the

---

17 From North and Macal 2007.
psychological states of individual people. At the very least, even the most strident individualist needs to admit non-psychological stuff, like bodily movements. Quietly, Watkins started to expand his notion of what counts as individualistic beyond the psychological.\(^\text{18}\)

But his moves were not enough. In the intervening years, epicycle upon epicycle has expanded, in one way or another, the set of individualistic facts. Some theorists emphasize that when a customer purchases a drink from a barista, the two people interact with one another; they do not just stand isolated from one another. Thus, they argue, we should take the individualistic facts to include certain relations between and among individuals.\(^\text{19}\) Others point out that social facts depend on the parts of the world we interact with. In microeconomics, for instance, we typically model economies not just as interacting individuals, but also include “bundles of resources” they own. Here the individual-territories include not just people, but the “bundles” they govern.

Other theorists point out that we need a more sophisticated theory of beliefs and other mental attitudes, beyond simple psychological states. Philip Pettit, for instance, has developed a theory he calls “individualistic holism,” where the mental attitudes of individuals are themselves social.\(^\text{20}\) Still other theorists take the basic building block to be at once individualistic and yet somewhat broader than the bodies and minds of individual people. Anthony Giddens, for instance, builds social facts out of “practices,” which are the patterns of activity of individuals in the world.\(^\text{21}\)

All of these theories take the whole of the social world to be carved into individual people and — in one way or another — their respective “territories.” Yet it is unclear why any of these should succeed.

Admittedly, some social phenomena do seem to be naturally divisible into individual people, or into individuals and their resources. For instance, a flea market is a bunch of tables piled with goods, each table manned by a vendor who owns the goods being sold. Individual customers walk around the tables, holding some money, and sometimes exchanging that money for

\(^{18}\) E.g., in Watkins 1959.  
\(^{19}\) This point is less significant than is often thought. See my discussion of Hodgson in Epstein 2014.  
\(^{20}\) Pettit 1993  
\(^{21}\) Giddens 1984
goods.

But many things in the world are unlike flea markets. Starbucks outlets, for instance. Or air force battles: these do not naturally break down into individual people. People are involved, but the basic units of action seem to be planes and aircraft carriers, not people. It seems more natural to see the battle as pieces of military hardware interacting with one another, with the people acting as resources apportioned to them, than the interaction of people. Or economies: many economists take these to be divisible into individuals, households, firms, and institutions, each with its own bundle of resources. Yet it is not obvious why this should be so, any more than for Starbucks.

This much is only an intuitive point, a seed of doubt about ontological individualism. Some theorists will regard the comparison with Virchow to be slander. Others will defend Virchow and the comparison. Still others will wonder whether there is not a different sense in which the social world is made from individual minds — namely, that it is a projection of our minds onto the natural world. (This is an issue I take up in the next chapter.) But in all this, one thing is clear: we cannot trust the prevailing dogma. We cannot trust it — not until we engage in more careful metaphysics.

A more careful metaphysics is best done from scratch. It is a waste of time to haggle endlessly over the meaning of ‘individualism’, or to trace the paths of the dozens of individualistic theories in circulation. Instead we ought to just cut through it. Given that ontological individualism is a claim in metaphysics, we might as well avail ourselves of the latest technology in metaphysics, rather than shy away from it. We should even add to that technology when needed.
Another puzzle: a competing consensus

The prevailing dogma cannot be trusted. Ontological individualism was a major advance over those old theories that mixed up claims about ontology with claims about explanation. And its enduring and widespread popularity should make us wary of a quick dismissal. But this does not mean it is true. Nor does the absurdity of social dualism make ontological individualism true. Ontological individualism needs to be examined suspiciously and closely, precisely because it has *individualism* at its core. So, it seems, we have a long road ahead of us.

At this point, however, I want to take a right turn. In the last three chapters, I have been speaking about the consensus, the settled view, of the nature of the social world. That was not quite accurate. Ontological individualism is almost universally endorsed. But it is not the only consensus claim about the social world. A second view is also almost universally endorsed. The philosopher Francesco Guala, in fact, has gone so far as to call this second view “The Standard Model of Social Ontology.”

The idea is this: the social world is a kind of projection of our thoughts, or attitudes, onto the world. We, as a community, *make* the social world by thinking of it in a particular way. The bills in my pocket are money because we all think of them as money. The President has the powers he does because we grant him those powers. America is a nation because we think of it as such. The social world, quite generally, is the social world in virtue of our beliefs about it.

Strangely, we rarely notice that this thesis differs sharply from ontological individualism. To many people, the Standard Model seems like a version of ontological individualism, a particular way to fill out the details. But it is not. It is a different claim about a different aspect of social metaphysics. Moreover, the Standard Model offers a different response to dualism about the social world. Just like ontological individualism, it denies

---

1 Guala 2007
an autonomous or separate sphere of the social. But the Standard Model’s denial of dualism exploits a different strategy than does ontological individualism. In fact, the Standard Model is at odds with ontological individualism, even though both views are endorsed by many of the same people. Ontological individualism does not logically contradict the “Standard Model,” but if one is right, it is very likely that the other is wrong.

Instead of one consensus view, we have two. And they are in tension with one another. With this, we find a big monkey wrench in the works. Theorists largely see themselves as agreeing on the basics of social ontology. But they do not have a consistent picture of what they are agreeing on.

There is, however, good news. By sorting out the conflict between these perspectives, we can cut the Gordian knot. We can quickly assemble a synthesis, a model of the social world with several parts that work together. Ultimately, the Standard Model will fall, just as ontological individualism will. But by seeing how they address complementary problems, we can pave a shortcut for clarifying both.

In this chapter, I present this Standard Model of Social Ontology, and explain how it differs from ontological individualism. I do so using two versions of the model: John Searle’s and, going back a few centuries, David Hume’s.

**Searle’s version of the “Standard Model”**

In *The Construction of Social Reality*, Searle proposes a theory of what he calls “institutional facts.” This is a broad category that includes many of the things in the social world. Universities, governments, restaurants, and money are all examples of institutional facts. Searle contrasts institutional facts with “brute facts.” A dollar bill is an institutional fact, and a piece of paper with green printing on it is a brute fact.

According to Searle, members of a community create institutional facts in their community, by imposing “statuses” on material objects. His simplest example is a boundary line around a village. Searle tells the following story. In ancient times, the inhabitants of some village built a high wall to keep out invaders. It worked because the wall was high. The wall physically

---

functioned to keep the invaders out. Over time, however, the wall deteriorated. Eventually, it was only a line of stones in the sand surrounding the village. But the villagers and their neighbors had grown accustomed to having the wall there. Despite the fact that the line of stones no longer physically functioned to keep people out, the villagers and their neighbors continued to treat the line of stones as a boundary, just as if it were the wall. At this point, the line of stones had taken on a symbolic function. Even though the line of stones provided no physical barrier to movement, the villagers and their neighbors have imposed the functional “status” that it once had as a physical object: the status of being a boundary.³

This is an intriguing story, and with it Searle introduces the centerpiece of his theory: the constitutive rule. The constitutive rule expresses the status that the villagers impose on the physical object, that is, on the line of stones in the sand. The constitutive rule for boundaries, according to Searle, is: “The line of stones counts as having the status of functioning as a barrier in the village.” Saying that X counts as Y in C basically means that, in context C, we treat X as if it performed the function we associate with Y, and hence we give it a certain status. The “X-term” denotes the object to which the status is assigned. The “Y-term” denotes the status assigned to it. We have institutional facts because we have constitutive rules in place in our society.

From this toy example, Searle moves to a more realistic one: paper money. In creating money, according to Searle, we assign a status to pieces of paper that have been printed in a certain process. Dollar bills, for instance, are just pieces of paper issued by the Bureau of Engraving and Printing. But when a piece of paper is issued in this way, we assign it a very important status. Dollar bills, according to Searle, have the following constitutive rule:

(CR) Bills issued by the Bureau of Engraving and Printing (X) count as dollars (Y) in the United States (C).⁴

The form of the constitutive rule, however, is only part of the theory. The constitutive rule expresses what object or objects receive what status. But what is it that puts constitutive rule CR in place, in a community? What

³ Searle 1995, p. 39
⁴ Searle 1995, p. 28. It is questionable whether Searle gets the conditions for being a dollar bill right, but CR is fine for illustrating his view.
makes CR a constitutive rule for dollars? This is the second part of Searle’s theory.

Constitutive rules are not facts of nature. Instead, Searle argues, what puts a constitutive rule in place in a community is that we collectively accept it. That is, constitutive rule CR is in place in the United States because the people in the United States all have a particular attitude toward CR. “Something is money,” Searle explains, “only because we think of it as money.”

The other centerpiece of Searle’s theory, then, is an account of “collective acceptance.” Searle argues that for a community to “collectively accept” something is not for just a bunch of individuals to have the attitude “I accept the rule.” In Searle’s view, for the members of a community to collectively accept something is for each of the community members to have a special kind of attitude toward it: a “we-accept” attitude.

To summarize, Searle’s theory consists of two parts. One part is about constitutive rules, which have the form \( X \) counts as \( Y \) in \( C \). The second part is about what puts the constitutive rule in place in our community. What puts a constitutive rule in place, according to Searle, is that we collectively accept that rule. Collective acceptance is the glue that binds a constitutive rule to a community.

**Another version of the standard model: Hume’s theory of convention**

Searle’s is not the only version of the “Standard Model of Social Ontology.” To broaden our view a bit, it is helpful to look back to an older theory, historically even more influential than Searle’s. In Book III of *A Treatise on Human Nature*, David Hume presented a theory of government, justice, money, property, promises, and languages. Hume argued that these things are created by social conventions.

Consider, for instance, his theory of promises. Hume begins with a distinction: a person’s promise to do something, he points out, is different

---

5 Searle modifies this slightly in Searle 2010.

6 This analysis of collective attitudes is controversial. Still, nearly everyone agrees that collective acceptance of a rule is not just a matter of each person having an individual-acceptance attitude toward it. I discuss collective attitudes in Chapters 14-16.
from a mere resolution to do something. If I promise to paint your house, that involves my resolution to paint your house, but it also involves more. If I resolve to paint your house, I am not *obliged* to paint it. Whereas if I promise to paint your house, I have incurred an obligation to paint it.

According to Hume, social convention is what adds obligation to promises. On Hume’s account, we have a social convention of the following form: words uttered according to a certain formula incur an obligation. That is, when somebody utters a phrase of the form ‘I promise to S’, that utterance is a promise.⁷

Before Hume, political theorists had applied the notion of “convention” to certain kinds of laws, and occasionally to language as well. Hume expands it to a much broader array of social phenomena. Property, money, justice, etc., are all products of social convention.

What is a social convention? It is here that Hume makes the biggest contribution. Historically, theorists had thought of convention as a matter of agreement — either explicit or tacit. Hume, however, severs the connection between convention and agreement. A convention, for Hume, is instituted by some regular behavior, together with beliefs on the part of members of the community that the behavior is in their mutual interest. The recognition by each member of the community that performing the activity will be mutually beneficial, and the expectation that other members of the community will perform the activity as well, provides reason to perform the activity. A convention, says Hume in the *Enquiries*, is:

> a sense of common interest; which sense each man feels in his own breast, which he remarks in his fellows, and which carries him, in concurrence with others into a general plan or system of actions, which tends to public utility.⁸

For a convention to be in place, in Hume’s view, is thus a matter of certain beliefs being held by members of the community, accompanied by regular behaviors.⁹ In the case of promises, the regularities and beliefs put in place the convention that when somebody utters a phrase of the form ‘I promise to S’, that utterance is a promise.

---

⁷ See Hume [1740] 1978, Book III, Part II, Section V.
⁸ Hume [1777] 1975, p. 257
⁹ Hume does not explicitly say that those beliefs can be tacit, but makes it clear that he thinks they can. (See Hume [1740] 1978, Book III, Part 3, Sec. 2.) Moreover, tacit conventions were discussed long before Hume; see, for instance, Pufendorf [1673] 2007.
So Hume’s theory can be divided into two parts, just as I described in connection with Searle’s. First, as with Searle’s constitutive rule, Hume has a principle or rule that is established: When somebody utters a phrase of the form ‘I promise to S,’ that utterance is a promise. Second are the facts about our beliefs and actions that put the convention in place. In this case, English speakers believe that it is mutually beneficial to take utterers of the formula ‘I promise to S’ to incur obligations to perform S, and they behave accordingly.

Hume’s account gives us the basics of a theory that at once illustrates the “Standard Model,” and at the same time departs from Searle’s theory. In Searle’s view, the constitutive rule is entirely put in place by the possession of certain “acceptance” attitudes by members of the community, toward the constitutive rule. Hume’s theory differs in two notable ways. First, the attitudes are simpler on an individual level: each member has a belief and an expectation about the behavior, but does not have to have an attitude directed toward the rule. The rule emerges from a matrix of common beliefs, without the rule itself having to be the object of attitudes. Second, Hume’s theory of what puts a convention in place involves more than attitudes. In Searle’s theory, constitutive rules are put in place only by attitudes. Hume, on the other hand, takes conventions not just to involve our attitudes, but also material facts about regularities in our practices.

Summarizing the Standard Model

In an insightful discussion of the topic, Francesco Guala highlights three characteristics of this “Standard Model of Social Ontology.”10 First, according to this model, the social world is made by our attitudes. In Searle’s view, for instance, we have collective attitudes about what makes something money, and those attitudes are the very things that put the constitutive rule in place.

Second, the model holds that the social world is performative: “If social entities are made of beliefs, they (unlike natural entities) must be constantly re-created (or ‘performed’) by the individuals who belong to a given social group.”11 One point here is that social things can be created by linguistic performances or declarations. The official utterance “I declare you man and wife” does not just describe a marriage, but creates it. Another point

---

10 Guala 2007
11 Guala 2007, p. 962
is that the social world must be actively maintained. In Hume’s theory, a
convention is sustained just as long as members of the community continue
to have the appropriate beliefs and practices regarding it.\(^\text{12}\) Searle has a
similar view. In order for a constitutive rule to be maintained in our
community, we need to continue to accept that rule. That continued
acceptance does not need to be at the forefront of our minds.\(^\text{13}\) We can accept
things tacitly, as well as explicitly. But without continued acceptance, the
constitutive rule is no longer in place in the community.

Third, the model takes the social world to be the product of collective
intentionality. This is the more general topic, and both Searle’s theory of
collective acceptance and Hume’s theory of shared belief are instances of it.
All the theories in the Standard Model hold that we need to have some
collective attitudes or intentions in order to constitute the social world.

Taking all these together, the Standard Model provides a particular
picture of the nature of the social world: the social world is made and
maintained by us, by our mental attitudes. And not just our attitudes as
individuals, but as a community.

\textit{The tension between the two consensus views}

Many things are appealing about this “Standard Model.” But how does it
relate to the theories of the social world in the last three chapters? The way
this model approaches the making of the social world is distinct from the way
ontological individualism does. Ontological individualism is about one
relation between individuals and the social world, and the Standard Model is
about an entirely different relation. Consider the following sets of examples:

1. a. A \textit{mob} of drunken hockey fans storming down Howe Street in
Vancouver, breaking windows and overturning cars
   
   b. The \textit{flow of commuters} in the Boston metropolitan area,
      moving in and out of trains, subways, and buses
   
   c. The \textit{Jewish people}, expelled from Spain in 1492 and migrating
      into Europe and North Africa

2. a. A \textit{handicapped parking spot}, marked by a blue and white sign

\(^{13}\) Searle 1995, pp. 117-19, 127-47.
ANOTHER PUZZLE: A COMPETING CONSENSUS

b. A tea party arranged by a group of children, with stuffed animals arranged around a table, each with a miniature tea cup and saucer

c. An unkosher animal, such as a lobster or pig

All of these involve social objects in some sense “made” by people: a mob, a flow of commuters, a people, a parking spot, a tea party, an unkosher animal. But the two sets of examples involve people in different ways. The things in set 1 are composed of people. The mob is composed of drunken fans, the flow of Boston commuters is composed of travelers, and the Jewish people is composed of various people with a certain religious/ethnic background. The things in set 2, on the other hand, are not composed of people. The parking spot is a section of pavement, with some paint marks. The tea party is a bunch of stuffed animals and cups. (The animals, not the children, were sitting at “tea.”) The unkosher animal is a real animal.

The sets also differ in how they are conceptualized, or “borne in mind.” The parking spot is authorized by the city, and marked out by city workers. The tea party is orchestrated by the children, who tacitly agree on their game. The laws of kosher food are set out in Leviticus. The drunken mob, on the other hand, is more like the school of selfish herring. Although each individual is intelligent (or, somewhat intelligent) few if any conceptualize the mob as a whole. The same holds true for the commuters and the Jewish people. These people may or may not conceive of themselves as a people or group.

These are two very different relations between the people and the social phenomenon they “make.” In a very loose and casual sense, we might say that the fans constitute the mob, and the children constitute the tea party. But these turn on two different meanings of ‘constitute’. Consider how we might depict the relation between facts in the two situations. Type 1 examples can be depicted quite simply:

\[ \text{The mob ran down Howe Street} \]

\[ \text{Bob, Jane, Tim, Joe, Linda, ..., and Max ran down Howe Street} \]

Figure 4A Depicting Type 1 examples

This pair of facts is related in the way that ontological individualists, like Watkins, Lukes, and the supervenience theorists, were suggesting. The
properties of the mob, like *running down Howe Street*, supervene on the properties of Bob, Jane, and the others. There is nothing to the social facts over and above the facts about the individual people.

Type 2 examples, however, cannot be depicted so simply. Here is a first pass at drawing the relations between a fact about the tea party, and the non-social facts that “make” that social fact:

*Figure 4B Depicting Type 2 examples*

In the second diagram, the roles of people are different from their roles in the first. (I have highlighted the people in bold letters, to make it clear where they are.) In the first diagram, there is nothing to the mob “over and above” Bob, Jane, and the others. In the second diagram, we can see two different relations at work. The upward-pointing arrow does the same thing as in the previous diagram: it connects the facts about the things that “compose” the tea party to the fact about the tea party. The tea party is not composed of people. The rightward-pointing arrow represents a different relation. Lea, Micah, and Jeremy are not the constituents of the tea party, but their game is the “glue” that holds together what is going on in the box. Lea, Micah, and Jeremy’s mental attitudes set up the game, put in place the “constitutive rule” for what it takes to be at tea. But the things that satisfy the X-term of that constitutive rule are not mental attitudes. They are stuffed animals.

Ontological individualism holds that social facts supervene on facts about individual people. The property *being a mob* supervenes on properties of individual people clustering together. The same cannot be said for *being a tea party* or *being a dollar*, according to the “Standard Model.” It does not supervene on mental attitudes alone.

What does *being a dollar* supervene on? Among the things it supervenes on are the properties of pieces of paper. Given that the constitutive rule for dollars is what it is, for something to be a dollar requires that it be printed in green ink, on a particular kind of paper. Moreover, it supervenes in part on the properties of the Bureau of Engraving and Printing.
Another Puzzle: A Competing Consensus

According to Searle’s version of the Standard Model, the fact that the community collectively accepts rule CR puts in place that rule. That is, it sets up the conditions for something to be a dollar. Even after the rule has been collectively accepted, however, we do not yet have dollars. To get those, we need the X-part of the constitutive rule to be satisfied. We need green pieces of paper, and issuance actions by the Bureau of Engraving and Printing.

The Standard Model, in other words, does not say that social facts are collective intentions. It says that collective intentions set up conventions, or constitutive rules. These conventions or constitutive rules are propositions about what sorts of non-mental, non-personal stuff constitute social things. This, in short, is the tension between the two different consensus views. Ontological individualism holds that social facts supervene on facts about individuals. The Standard Model holds that facts about individuals set up the conditions for something to count as a social fact. They offer two distinct approaches to the social world.

Another conflict is in how the respective views respond to dualism. Ontological individualism addresses dualism using the strategy described in Chapter 2: social facts supervene on individualistic facts. The Standard Model also addresses dualism, but differently. It takes the social world to be a projection onto the non-social facts. On the Standard Model, the social world is “brute” facts treated in a certain way.

That does not mean that Standard Model theorists take the social world to be a fiction or an illusion. Two things are meant to prevent that. First, social facts are the product of collective, not just individual, attitudes. (On the Standard Model, if I alone accept that Bob is President, then that is a fiction, but somehow, if we as a group accept that Bob is President, then that is a social fact.) Second (and more persuasively), on the Standard Model the social world is not an attitude in our heads, but the actual stuff in the world to which a certain status or convention has been assigned. Take away the line of stones, and there is no boundary. The boundary is not an illusion; it is just the line of stones taken against the background of our constitutive rules or conventions.

Where we stand

If the Standard Model is so different from ontological individualism, why
have they been conflated with one another? The main reason, in my view, is that they have both remained so loose and unclear. In social theory, we have been cavalier about the metaphysics and not taken the details seriously enough. To be sure, there are other reasons as well. Both models appear to be responses to dualism. Both are in a sense “individualistic.” Both seem to apply to certain widely discussed cases. But for all these similarities, the views are different enough that if we talked about them precisely, their differences would pop out.

Unclarity is endemic in all sides of the literature. Consider, for instance, Searle’s term ‘institutional fact’. Examples of institutional facts, according to Searle, are money, marriages, restaurants, nations, national boundaries, etc. This is very confusing terminology. Money is not a fact — it is a social object, or maybe a social kind. *I have a dollar in my pocket:* that is a fact, a social fact. *A dollar exists:* that is a different social fact. *The bill in my pocket constitutes a dollar:* yet a different social fact. But *dollar:* that is a thing or a kind of thing, not a fact. And *being a dollar:* that is a property. If we want to inquire seriously about the making of social facts, objects, or properties, we need to be clear about what we talking about.

Even the best discussions in the literature are loose about the details. Consider, for instance, the passage from Harold Kincaid I quoted above. Although it seems to be written very precisely, it really is quite confusing. Here are three important ways it is unclear:

*i. It is unclear about what it means for something to determine something else.*
Kincaid says, “Social wholes are composed of individuals and determined by their actions.” What does this sentence mean? If “social wholes,” whatever they are, are composed by individuals, then what about them is “determined by their actions”? Does Kincaid mean that they are caused to be what they are by actions? Or does he mean that they are somehow composed by actions? Yet if they are composed by actions, then why did he say that they are composed of individuals? Then in the next sentence he says that they are determined by individuals. Why the flip between composed and determined, and between individuals and actions?

*ii. The role of language is unclear.* Kincaid speaks of facts being expressed “in the relevant individualist vocabulary.” What does vocabulary have to do with it? If some fact is the case, what does it matter whether it is
expressed in one vocabulary or another? For instance, if the fact *I am six feet tall* is “set,” then it makes no difference to the fact being set if we express it as “I am six feet tall” or as “I am 72 inches tall.” Does it?

iii. *It is unclear about the sorts of social and individualistic things we are talking about.* In the quoted passage, Kincaid moves from a point about “social entities, events, etc.” to a point about “supervenient properties.” Why the change to properties, and is the thesis about properties meant to be equivalent to the preceding thesis about social entities and events? Just what sorts of social things are determined by what sorts of individualistic things? Facts? Events? Properties?

My point is not to pick on Searle and Kincaid: their discussions of the social world are among the best out there. But if we are to move past intuition and to seriously understand society, we cannot be so loose about the building blocks we choose to work with. Most crucially, we need to avoid mixing up talk about the world and talk about language. And we need to be clear about what things we are talking about — facts, events, properties, relations, objects, etc.
This chapter offers a few basic tools of metaphysics, including the following:

- The distinction between facts, propositions, and sentences
- Possible worlds and possible facts
- Properties and relations
- Social facts and social kinds
- The grounding relation

This is a long list. Rather than giving a complete or systematic presentation, I will try to offer a reasonably precise way to make and assess claims about the nature of the social world. There are many models we might use for accomplishing this, but as much as possible I will stick to the standard interpretation of the standard tools. (The discussion of “grounding” is a minor exception, for reasons I will discuss.)

**A three-part model: facts, propositions, and sentences**

Let's start with a couple of facts: *The Earth is round,* and *Bill Clinton was President of the U.S. in 1994.* These are facts about the world. We use language to talk about them, but they are facts about a planet, a shape, a person, and a social property, not facts about language.

A three-part model for organizing these points distinguishes (a) the world, (b) abstract representations of the world, and (c) language, or ways of speaking about the world. According to this model, the sentence ‘Bill Clinton was President of the U.S. in 1994’ expresses a proposition. The same proposition can be expressed with other sentences as well. For example, ‘The President of the U.S. in 1994 was Bill Clinton’, or ‘Bill Clinton était président des États-Unis en 1994.’ Because it is a fact that Bill Clinton was President of the U.S. in 1994, the proposition representing that fact is true.

Facts are things in the actual world. Propositions are abstract representations of the world. Some propositions are true and some are false. The true propositions are the ones that correspond to the facts, and the false
propositions are the ones that do not. Each fact corresponds to a true proposition.\footnote{This is standard, but there are many ways of analyzing facts and propositions. See Richard 1990; Neale 2001.}

Sentences are expressions in particular languages, but propositions do not depend on language. Before there were people, there was no language, and no sentences to express propositions. But there were still true and false propositions.\footnote{Many people are bothered by the idea that there could be propositions without us. Again, all this talk about propositions can just be regarded as a useful model for thinking about the world. Lots of metaphysicains have developed models that circumvent propositions, but they tend to be much more complicated than the standard ones are, and not as powerful.} The proposition expressed by the sentence ‘The Earth is round’ would have been true even if people had never existed, because it is an abstract representation of the fact that the Earth is round. Which would be true even without us. Some propositions are about language, like the proposition expressed by the sentence, ‘A’ is the first letter of the alphabet.’ But many propositions are not.

Propositions are not linguistic, nor are most of them about language. Still, we do use language to express propositions. Furthermore, we also use language to denote facts and propositions, i.e., to specify which fact or proposition we are talking about. When I am being precise, I will use italics for propositions, and bold italics for facts. For instance:

(5.1) The proposition \textit{The Earth is round} is true.

(5.2) The fact \textbf{The Earth is round} obtains.

Notice that (5.2) does not say that the fact \textbf{The Earth is round} is “true.” Facts are not true or false, any more than a chair or a lake is true or false. An object exists and similarly a fact obtains, or is the case. A proposition corresponding to a fact is true if and only if that fact obtains.

\textbf{Possibilities}

Some propositions are true, and some are false. For example, the proposition \textit{Bill Clinton was President of the United States} is true, and the proposition \textit{Chelsea Clinton was President of the United States} is false. Equally, we can distinguish (actual) facts from possible facts. Possible facts are things that might obtain in the world, some of which actually obtain and others of which do not. For example, the possible fact \textit{Bill Clinton was President of the U.S.}
actually obtains, whereas the possible fact *Chelsea Clinton was President of the U.S.* does not. But both are alike in terms of being *possible facts*.

Some people are queasy about talk of possibilities, and about treating possible facts similarly to how we treat facts. But this model is convenient. Possible facts correspond to propositions, and each possible fact that obtains in the actual world corresponds to a true proposition. This is particularly helpful for thinking systematically about the sciences. When we build models in the sciences, we are not only interested in the way things actually are, but how they *would be* if things changed. How would unemployment change if we raised interest rates? How would educational attainment change if we increased standardized testing? These are questions about ways the world might be. Similarly, we might model the circumstances in which *Chelsea Clinton is President of the U.S.* obtains. In such a model, we could consider what other facts have to obtain in order for it to obtain.

Another convenience is to talk about “possible worlds” as a whole, and to compare them to one another. A good deal of debate in metaphysics is dedicated to the question of how to understand possible worlds. Are they fictions? Concrete objects? Linguistic constructions? Complex properties? Shadowy “ways things might be”? For our purposes, it does not much matter which of these we adopt. We can avail ourselves of talk like “In some possible worlds, Chelsea Clinton is President of the U.S.,” and “In all possible worlds, two plus two equals four,” without committing ourselves to a particular theory of possible worlds, or possible facts.³

It is now standard to cash out the ideas of necessity, contingency, and impossibility, in terms of possible worlds. Whether a proposition is true or false often depends on the way the world is, although some propositions are true or false regardless of the way the world is. For instance, the proposition *Bill Clinton was President of the U.S.* is true. If things were different in the world, that proposition would be false. Such a proposition is *contingent*. The proposition *2+2=4*, on the other hand, is true however the world might be. This proposition is *necessary* — i.e. true in every possible world. And the proposition *2+2=5* is false however the world might be. It is, in other words, *impossible* — false in every possible world.

³ This should be qualified: see the arguments against “linguistic ersatzism” in Lewis 1986, pp. 142-65. Also see Sider 2002.
Properties and relations

Just as the person Bill Clinton is not the same thing as the name ‘Bill Clinton’, we also need to distinguish properties from the predicates we use to denote them. The expression ‘being President of the U.S.’ is a predicate. It is a linguistic item. That predicate denotes a property, i.e., the property being President of the U.S.

The predicate ‘is taller than’ expresses a relation between two objects, not a property of a single object. The predicate ‘is taller than’ is called a binary (or two-place) predicate, and the relation is taller than is called a binary (or two-place) relation. Thus the sentence ‘Bill is taller than Hillary’ consists of two names and a binary predicate, and the proposition Bill is taller than Hillary consists of two objects — Bill and Hillary — and a binary relation that holds between them — the is taller than relation. Sometimes I will flank the relation with letters, to make it clear that is a binary relation: for instance, the relation A is taller than B.

There are also three-place predicates and relations, four-place predicates and relations, and so on up. For instance, if Carol is sitting between Bob and Alice, then the three-place sitting between relation holds of the three objects Carol, Bob, and Alice.

Objects actually have (or instantiate) some properties, and possibly have others. Chelsea Clinton, for instance, has the property being the daughter of Bill and Hillary, and possibly has the property being President of the U.S. In any possible world, any given object either does or does not have any given property. Additionally, it should be stressed that how a property is instantiated does not change over time. Of course, the properties that a given object instantiates can change over time. For instance, the color of a particular piece of clothing may fade over time, e.g. from bright red to pale red to pink. But what it takes for something to have the property being bright red remains invariant over time.

To put this point differently, we might associate with a property a set of “instantiation conditions.” For a given property P, the instantiation conditions are simply other properties, such that necessarily, an object has P if and only if that object has properties R and S and T, for example.4

4 I do not assume that all properties have such sets of associated properties, nor that we can always know when a property does.
Instantiation conditions like these do not change over time or possibilities. Even a property that involves a particular time, such as being President of the U.S. in 1994, has fixed instantiation conditions.

I will use this same model of properties for social properties, just like any other property. For instance, Bill Clinton changed from having the property being Governor of Arkansas to being President of the U.S. But the conditions for his having either of these properties did not. “But suppose,” one might object, “the Arkansas legislature changed the law about how Governors are elected. In that case, wouldn’t it be true that the instantiation conditions for the property being Governor of Arkansas have changed?” Not according to the model of properties we are using. To make sense of this, we can understand the instantiation conditions to be something like satisfying whatever conditions the Arkansas legislature sets out for being Governor. Or we can simply take there to be two different properties having different instantiation conditions, the property being-Governor-of-Arkansas-before, and the property being-Governor-of-Arkansas-after.

This model is, of course, an idealization. We do identify the before- and after-versions of such properties, as being a change in a single property. But there are two good reasons for pushing that burr under the carpet. First, in order to make sense of it, we would need a more sophisticated apparatus than I can discuss in this book. We would need a model of how a given property can be “anchored” and “re-anchored,” over and over again, while remaining the same property. (Anchoring is a topic I introduce in the next chapter.) That, however, would require us to revise more metaphysics than is practical or needed for now. So, for our purposes, properties have unchanging instantiation conditions, both over all times and across all possibilities.

Second, there are advantages to this model of properties — i.e., as having unchanging instantiation conditions over time. When we want to assess whether a given object has a given property, we want to do it in a single way, regardless of the time or circumstances we are assessing. Otherwise there is no way to compare whether things have changed. Consider, for instance, a proposition about some change over time, something like Clothing has become more brightly colored over the course of the last century. We do not want the truth or falsity of that proposition to depend on
changing conditions for what it takes to be brightly colored. Instead, we take the property *being brightly colored* to have fixed instantiation conditions, throughout that period.5

So long as we are being clear about properties and relations, we should also be clear about what sorts of things they apply to — i.e. their *relata*. For instance, I have already mentioned the supervenience relation. The literature sometimes gets quite confused because it is vague about the relata of this relation, i.e., what supervenes on what. As I pointed out, supervenience is best (and most commonly) understood as a relation between sets of properties. But this is often muddled, with people talking about it as relating sets of facts, events or even sets of predicates. Although these may seem innocuous, they are not, as I will discuss later on.

**Social facts**

In the passage I quoted earlier, Kincaid says: “Once all the relevant facts (expressed in the preferred individualist vocabulary) about individuals are set, then so too are all the facts about social entities, events, etc.” But when is a given fact a fact *about* individuals? Kincaid seems to suggest that it has something to do with being expressed in “the preferred individualist vocabulary.” But this cannot be right. Facts are unaffected by the way we describe them, in the same way that the moon is unaffected by the ways we describe it — e.g. calling it ‘la luna’. Fortunately, the model I have described points us in a more fruitful direction. Consider an example from the last chapter:

(5.3) **Bob, Jane, Tim, Joe, Linda, … and Max ran down Howe Street.**

(5.4) **The mob ran down Howe Street.**

Are (5.3) and (5.4) the same fact? It is understandable why one might think they are. Both obtain, and at least in some limited sense, there is nothing more to the mob than those people. But there are at least two reasons that (5.3) and (5.4) denote different facts. One is that there is a possible world

---

5 That does not, however, mean that when I talk about properties I mean only intrinsic properties. (On the distinction between intrinsic and extrinsic properties, see Kim 1982; Langton and Lewis 1998; Lewis 1983; Yablo 1999.) Both intrinsic and extrinsic properties have fixed instantiation conditions.
in which (5.3) is the case and (5.4) is not, and another possible world in which the reverse is the case. (Consider a world in which Bob and the others are running down different parts of Howe Street, maybe in different directions. Though they are all running down Howe Street, they are not a mob. And consider a world in which the mob is running, but Bob leaves the mob.) If they denoted the same fact, one could not obtain without the other.

Here, however, I also want to highlight another reason these are two different facts. In the model I have described, facts correspond to propositions, and the corresponding propositions are different. One is a proposition having Bob, Jane, and the others as constituents, and the other is a proposition having the mob as a constituent. There are two different propositions, and hence there are two different facts: one about individuals and one about a mob. Notice that this has nothing to do with language, or the vocabulary in which anything is expressed. The way we distinguish facts from one another corresponds to the way we distinguish propositions from one another, not the way we distinguish sentences from one another.

As a rough guide, we can take a social fact to be a fact that corresponds to a proposition that has any social entity as a constituent. It might have social objects as constituents, or it might have social properties as constituents, or both. Of the four facts listed in Table 5a, the first three are all plausibly social facts, and the fourth is not:

<table>
<thead>
<tr>
<th>Social vs. non-social facts</th>
<th>Social object</th>
<th>Social property</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>The mob was impoverished.</em></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td><em>The mob was cold.</em></td>
<td>x</td>
<td>-</td>
</tr>
<tr>
<td><em>John was impoverished.</em></td>
<td>-</td>
<td>x</td>
</tr>
<tr>
<td><em>John was cold.</em></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Some people talk only about social objects, and some only about social properties. But if we are talking about social facts, we cannot limit ourselves

---

6 Here I am speaking of propositions as though they were Russelian, i.e., having objects and properties as constituents. Similar points can be made with different conceptions of propositions as well, but this is a convenient one.

7 This presumes, of course, that the mob is not identical to the individuals. I discuss this point in Chapter 10.
Social kinds

It is common nowadays for philosophers to speak of “social kinds” or “human kinds.” This is a term introduced by analogy to “natural kinds,” which have long been discussed in metaphysics and philosophy of language. The intuitive idea of a natural kind is a category of objects grouped naturally, rather than arbitrarily or by fiat. Typical examples given of natural kinds include kinds in physics, such as electron and charge, kinds in chemistry, such as gold and water, and kinds in biology, such as the various species. A gold ring, a gold bar, and a gold nugget, for instance, all group together into a category because of their chemical composition, which does not depend on human choices or interests.

It is controversial which kinds are natural, and even whether there are any natural kinds at all. But, these debates notwithstanding, natural kinds seem to have some distinctive characteristics. John Stuart Mill, for instance, observed that they form the basis for inductive inferences in the sciences. By investigating certain gold things, testing and analyzing their characteristics, we can draw inductive inferences about other gold things. Another characteristic natural kinds seem to have is that they are essential to their members. If something is a piece of gold, then it is essentially a piece of gold: it could not be changed into lead without being destroyed. This is a more controversial thesis, and the relation between natural kinds and essentialism remains hotly debated.

The notion of a social kind is somewhat looser. It is convenient to distinguish social kinds from social properties more generally, largely because social kinds seem to figure in the social sciences similarly to how natural kinds figure in the natural sciences. Social scientists commonly use terms like ‘class division’, ‘religiously sanctioned inequality’, ‘economic factor’, ‘material circumstances’, ‘public good’, ‘commodity’, and so on. These terms are grammatically similar to natural kind terms, and the things they refer to seem to work in inductions, just as natural kinds do. On the other hand, it is certainly wrong to think of these categories as being independent of human

---

8 As I discuss in Chapter 8, this trips up Jaegwon Kim’s treatment of fact supervenience in Kim 1984.
activity, and it is not clear that they have the other distinctive characteristics that natural kinds do.

In speaking about social kinds, then, it is useful to think of them as the categories we might use in the social sciences, but remain open-minded about the sorts of categories these might be. Maybe the social kinds are the same as the social properties. Maybe they are a subset of the social properties. Or maybe they are a different thing altogether. To make progress, we do not need to start with a comprehensive understanding. We can just regard “social kind” as a generic way of referring to categories like these.

Social kinds — like social properties — have fixed instantiation conditions. Or, more appropriately, we might say that kinds have fixed membership conditions. The conditions under which something is a member of a social kind are the same across all times and possibilities. The reason for this is the same as the one I gave above. Social kinds serve a variety of functions: we employ them for recognizing things, classifying things in various situations, finding and correcting departures from norms, drawing inductive inferences, and accomplishing other practical matters.

As such, they are applicable across a universe of different situations: we can look at any object whatever, in any situation, and assess whether that object is a member of the kind teacher, tire, hem, or hipster. That does not mean that social kinds are not put in place by local contexts in the actual world. The kind hipster, for instance, is put in place by a range of idiosyncratic facts about our current society. But its potential instantiation is not limited to that current situation. The membership conditions for hipster, in other words, are what they are for a panoply of reasons. But the conditions are the conditions, and we can look around at all possible objects in all possibilities to see if those conditions apply, not just at the objects in our local context.

Grounding
Consider again the following two facts:

(5.3) Bob, Jane, Tim, Joe, Linda, ... and Max ran down Howe Street.
(5.4) The mob ran down Howe Street.

As I pointed out earlier, (5.3) and (5.4) are different facts. Nevertheless, they
are intimately related. On Tuesday evening, both obtained at 10:00 and were the case until 10:25. Then, at 10:26 neither obtained. Then at 11:08 they obtained again, and at 11:37 did not obtain.\(^9\) In fact, every time those people ran together down Howe Street, the mob did. An amazing coincidence!

Of course, the coincidence is not so amazing. These are not two arbitrary facts, but are related to one another in a particular way. Fact (5.4) obtains at 10:00 because (5.3) does. The two facts are metaphysically related to one another. They are not the same fact, nor is (5.3) quite sufficient for (5.4) to obtain. If Bob and the others disperse, then even if (5.3) is the case, (5.4) might not be. So to be more precise, (5.4) obtains at 10:00 in part because (5.3) does.

We use the word 'because' in many different ways. Often, it connects causes and effects. We say that the barn burned down because Mrs. O’Leary’s cow knocked over the lamp. The lamp is causally related to the fire, not metaphysically related to it. Knocking over the lamp “makes” the fire in a causal sense, not a metaphysical one. The flames, on the other hand, are metaphysically related to the fire. The flames do not cause the fire; in a sense, they are the fire.\(^{10}\)

Facts (5.3) and (5.4) are not causally related to one another. The fact that Bob and the others ran down Howe Street was not the causal reason that the mob did, but the metaphysical reason.

To assign a word to this “metaphysical reason” relation, we say that fact A grounds fact B. The fact The barn door, walls, and roof are burning grounds the fact The whole barn is burning; the fact I am married grounds the fact I am not a bachelor; and the fact a million herring turned in such-and-such directions grounds the fact the school split in two. All of these pairs of facts are metaphysically related in the sense that the first fact “metaphysically makes” the second fact the case. Grounding will be a central part of the discussion from here on, so I will say a bit more about a few of its characteristics.

i. **Fundamentality**

Grounding is usually understood to involve a kind of priority, or

---

\(^9\) I am being a bit casual about the role of time.

\(^{10}\) For more on the distinction between causal and constitutive relations, see Haslanger 1995, Bennett 2011, Schaffer 2012, Koslicki 2012.
fundamentality. The more fundamental fact grounds the less fundamental fact. For instance, the fact *A million herring turned in such-and-such directions* grounds the fact *The school of herring split in two*, but not the other way around.

This is perhaps the most controversial part of the notion of grounding, since many people worry about the idea that some parts of the world are “more fundamental” than others. In developing a model for thinking about social facts, however, it is hard to imagine how we could do without a notion of fundamentality. The fact *Bob and Jane ran down Howe Street* is part of what we look for in seeking the metaphysical explanation for the fact *The mob ran down Howe Street*. But not the other way around: the fact *The mob ran down Howe Street* does not metaphysically explain the fact that Bob and Jane did.

**ii. Partial vs. full grounding**

In some of the examples of grounding I have mentioned so far, the grounding facts are metaphysically sufficient for the things they ground. In the following case, for instance, the first fact is metaphysically sufficient for the second:

(5.5) *I am married.*

(5.6) *I am not a bachelor.*

Two facts, however, can stand in a relation where one is *part of* the metaphysical explanation for the other, but is not quite metaphysically sufficient. For instance, the following pair of facts:

(5.7) *Bob and Jane ran down Howe Street.*

(5.8) *The mob ran down Howe Street.*

Fact (5.7) is part of the metaphysical explanation for (5.8), because Bob and Jane were participants in the actual event. On its own, however, (5.7) is not enough. Two people do not make a mob. In this case, we say that the former fact *partially grounds* the latter fact. (Sometimes I will talk about fact G “fully grounding” fact F. But there is no difference between *G fully grounds F* and *G grounds F*. I add the word “fully” just to make the contrast with partial grounding clear.)

Looking at some of the other cases, it actually takes some thought as to whether they are examples of full or partial grounding. For instance, look
again at (5.3) and (5.4). Suppose that (5.3) is the case, but that all the people are actually running in different parts of Howe Street — some at the end, others a mile behind them, and others at the beginning. Then, even though (5.3) obtained, (5.4) might not. Another fact should be added to (5.3), in order to fully ground (5.4). For instance, the fact that all those people were clustered together. Or else the fact that other running people filled in the gaps between Bob, Jane, Tim, and the others. On its own, (5.3) only partially grounds (5.4). Together with one of those other facts, it fully grounds (5.4).

Here I will mention a point of departure between the way I will use the grounding relation, and the way it is usually understood in the literature. Most people take full grounding to involve necessity. That is, if fact F fully grounds fact G, then it is necessary that if F obtains, then G obtains. In my view, this is not the best way to understand full grounding. Instead, I will distinguish full grounding from an even stronger relation, the relation $A$ determines $B$. Determination, as I will discuss later, is basically just full grounding plus necessitation. There are practical reasons for making this distinction, but it is easier to explain this in the context of examples I will present later on.

So for now, I will just flag the point. If F fully grounds G, then F is metaphysically sufficient for G. That does not mean that in every possible world where you have F, you also have G. Still, (full) grounding is a pretty strong relation, and it is most important to keep grounding distinct from partial grounding.

### iii. Evidence for grounding

How do we tell if one fact partially grounds another? How do we tell if one fact fully grounds another? There is no infallible method that works in all cases. But there are ways of working it out. One way is just to think things through. Is fact (5.3) metaphysically sufficient for (5.4)? Or is there something to (5.4) over and above (5.3)? Sometimes we can come up with

---

11 For a detailed discussion of the controversy, see Skiles Manuscript. Also related are: Audi 2011; Correia 2005; Dancy 2004; deRosset 2010; Fine 2012; Leuenberger 2014; Rosen 2010; Witmer et al. 2005; Zangwill 2008. I favor the “contingency” view, but none of the substantive claims in the book turn on it. In particular, the division between grounds and anchors (see Chapters 6 and 9) is an independent point.

12 For a fuller discussion of the intuitive notion of grounding, see Audi 2011 and Rosen 2010. For a more complete overview of the details, see Fine 2012.
FOUNDATIONS: OLD AND NEW

case justifications for various grounding claims. Another method is to apply certain tests, or diagnostic tools. One especially valuable test, for instance, is to examine the ways that various facts vary in lockstep with one another.

Above I pointed out that (5.3) and (5.4) co-vary with each other over time. By that I mean that when one obtains the other also does. And when one does not obtain, the other does not. This co-variation is decent evidence that there is some relation between the two facts. On the other hand, there can be lots of reasons for co-variation even without grounding. They might be causally connected: where there’s smoke, there’s fire, but the presence of smoke and the presence of fire are causally related, not metaphysically. Or two facts might co-vary by accident.

Nonetheless, different sorts of co-variation can be useful tools for diagnosing different sorts of grounding relations. When we consider the co-variation of facts, we can consider how facts change over time, but also how facts change over different possibilities. This is what supervenience is about. Supervenience is built using the idea of co-varying properties: a set A of properties supervenes on a set B of properties just in case any change in the A-properties must be accompanied by a change in the B-properties.\textsuperscript{13} As such, it is a diagnostic tool for assessing whether facts of the form \(x\) has such-and-such an A-property are grounded by facts of the form \(x, y, z, \ldots\) have such-and-such B-properties. (More accurately, it is a test for metaphysical dependence. But dependence can also be evidence for determination, so supervenience is a useful tool for diagnosing both. I discuss these topics in more detail in Chapter 8.)

Moving on
To make headway, it is crucial to work with a simple and powerful toolkit, to be precise, and to apply the tools consistently. With regard to the nature of the metaphysical tools themselves, I tend to approach them with a lightly accepting temperament. For instance, do propositions really exist? Some people insist they do, while others scoff. For our purposes, neither hill is

\textsuperscript{13} More precisely, a change in the pattern of A-property instantiations must be accompanied by a change in the pattern of B-property instantiations.
worth dying on. To develop a model for making sense of the social world, I use propositions and these other tools freely, without worrying here about the commitment to a rich ontology.

It may seem ironic or even hypocritical to be casual about the "reality" of propositions, while refusing to be casual about the “reality” of governments, money, and other social entities. But this little hypocrisy is worth the payoff, at least to get us going. These tools of metaphysics are powerful for their precision, and for how much detailed work has gone into assembling them into a cohesive model. They enable us to think about the social world in much clearer ways than without them. And we have to use some toolkit or another, so we might as well start with a powerful one. We use the best tools at our disposal to investigate things they might help with. And then with those new insights, maybe over time we can scrutinize the tools themselves with more success.

Now let’s return to the two models discussed in the last chapter, applying these new tools to assemble the models into a unified picture. That will allow us to discuss the anchoring relation and the notion of a frame principle.
Case study: laws as frame principles

The grounding/anchoring model has many parts: social facts, the grounding relation, grounding facts, frames, frame principles, the anchoring relation, and anchors. It is time for a concrete application, to see how these things work in practice. Most importantly, I want to clarify the notion of a frame principle, and what role it plays in the model.

An illuminating example, to do this, is the law. Laws can be understood as frame principles. They give the grounding conditions for certain social facts, and they have anchors.

Understanding laws as frame principles helps clarify both the grounding/anchoring framework and also the nature of law. It sheds light, for example, on the difference between the law and the law code — i.e., the written documentation that tries to record what the law is. And it also helps clarify different theories of the “sources” of the law. All this is a nice immediate payoff of the model.

In this chapter I mostly discuss H.L.A. Hart’s theory of law, laid out in his 1961 classic The Concept of Law. I show how it fits into the grounding/anchoring model, and how it illustrates the way frames can be “nested” within other frames. There are two different kinds of laws in Hart’s theory — what he calls “primary rules” and “secondary rules.” Each of these can be understood as frame principles, with different kinds of anchors.

Hart’s theory of law

In The Concept of Law, Hart laid the groundwork for a new understanding of the nature of law. Hart’s book recast and defended “legal positivism,” and its framework became the structure for subsequent approaches to positivism and its opponents.

Theorists of the nature of law divide into two rough camps: the legal positivists and the natural law theorists. Legal positivists hold that facts about the law are facts about sociology, not morality. What counts as a law does so
in virtue of facts about our social beliefs and practices. This means that there can be grossly unjust laws. What the law is is one thing, and what the law should be is another.

Natural law theorists, on the other hand, hold that an unjust law is no law. We have a duty to follow the law, but cannot have a duty to be immoral. These two rough approaches have ancient roots, and over the generations have been refined into a variety of nuanced positions. Hart’s work influenced the course of these positions, not just among legal positivists, but among natural law theorists as well.

Hart divides the law into primary rules and secondary rules. Primary rules are those that direct and appraise behavior. Consider, for instance, the law on first degree murder. The Massachusetts General Law (M.G.L.) lists the following conditions:

Murder committed with deliberately premeditated malice aforethought, or with extreme atrocity or cruelty, or in the commission or attempted commission of a crime punishable with death or imprisonment for life, is murder in the first degree. (M.G.L. c. 265 §1)

The following section lists the punishments:

Whoever is guilty of murder committed with deliberately premeditated malice aforethought or with extreme atrocity or cruelty, and who had attained the age of eighteen years at the time of the murder, may suffer the punishment of death pursuant to the procedures set forth in sections sixty-eight to seventy-one, inclusive, of chapter two hundred and seventy-nine. Any other person who is guilty of murder in the first degree shall be punished by imprisonment in the state prison for life. (M.G.L. c. 265 §2)

These are all statements of primary rules. They involve the conditions for a certain kind of legal attribute, and given that, further conditions for sanctions.

Secondary rules identify, change, and enforce the primary ones. These govern how primary rules themselves are put into place and enforced. The most important secondary rule in a legal system is the system’s “rule of recognition.” This determines which other rules are recognized as legally valid, i.e., what the primary rules are, in that legal system. Part of the rule of recognition in a modern legal system, for instance, pertains to how the legislature enacts statutes. In Massachusetts, these are set out in Articles I and II of the State Constitution. Another part of the rule of recognition is the role judicial interpretation plays in shaping the law. Other secondary rules in a
system include the rules under which other rules may be changed, and the rules for who has the power to apply them.

The crux of Hart’s argument for positivism is a theory of what puts secondary rules in place in a society. Hart gives sociological — and non-moral — conditions for a secondary rule to be in place. Two conditions must be satisfied. First, legal practitioners, such as judges and lawyers, must have a convergent set of practices or behaviors. That is, they have to actually conform to the rule, when they act. Second, they all must take a certain attitude toward their practice.

More specifically, according to Hart, a rule R is present among a group P whenever there is a regularity in behavior such that (1) most people in P regularly conform to R, and lapses from conforming to R are criticized, and (2) R is “accepted” in P, i.e., R is treated as a standard for the behavior of people in P, and the criticism for lapses is regarded as justified.¹ These facts put secondary rules in place in a legal system, and the secondary rules contain the conditions for a primary rule to be in place.²

This is only a cursory description of primary and secondary rules, and I will fill out some details as I go. But even from this much, it may start to be clear how the grounding/anchoring model will apply. When we speak of “putting rules in place,” we are speaking about anchoring. And the rules themselves give the grounding conditions for particular legal facts.

A primary rule as a frame principle
I will start with in the same way I began discussing grounding and anchoring: with a particular fact. Consider, for instance, the fact Whitey Bulger is a first degree murderer. (Whitey Bulger, if you’re not from Boston, is the famed local mobster who, at the time of this writing, is heading to trial.) This legal fact is the case because of things Whitey Bulger actually did. It is grounded by

---
¹ Hart 1961, pp. 55-60. See also Coleman and Leiter 1996; Green 1996; Greenberg 2004; Marmor 2010, 2012; Raz 1975; Shapiro 2011.
² Strictly speaking, Hart’s secondary rules are a mix. The rule of recognition is put in place by these two kinds of facts, and other secondary rules (understood as rules about rules) are put in place the way primary rules are. That is, by satisfying the conditions set out in the rule of recognition. In this discussion, I will focus on the rule of recognition in speaking about secondary rules. I am grateful to Simon May for clarifying these points, and for detailed discussion of Hart and the law.
CASE STUDY: LAWS AS FRAME PRINCIPLES

historical facts about Whitey: facts about his having killed people and facts about his mental state in the course of doing so.

Massachusetts law contains specific conditions for having the property being a first degree murderer: if someone kills a person with deliberately premeditated malice aforethought, then the killer is a first degree murderer. All it takes for someone to be a first degree murderer is that the person satisfies those conditions. Since Whitey murdered several people, there are many facts that make it the case that Whitey Bulger is a first degree murderer. Among them, for instance, is the fact Whitey Bulger killed Bucky Barrett with deliberately premeditated malice aforethought.

The connection between these facts is, of course, not just happenstance. The fact that Whitey killed Bucky grounds the legal fact about Whitey: it is the metaphysical reason for Whitey being a murderer. This means that a careful statement of the law will look similar to the frame principles I discussed in the last chapter. That is, something closer to: For all x, if x kills a person with deliberately premeditated malice aforethought, then that grounds the fact that x is a first degree murderer.

This is just one example of a primary rule. Laws come in many forms. It is typical, for instance, for some primary rules to set out the conditions for a person to have some legal attribute, like being a murderer. And for other primary rules to set out the sanctions that accompany that attribute. We saw this in the two quoted sections of the Massachusetts General Law. Moreover, not all primary rules have the form For all x, if x has property P, then that grounds the fact that x has property Q. In many cases, laws just give the conditions for grounding particular facts, not kinds of facts. They are also sometimes categorical facts, such as A has property Q, as opposed to conditionals. (A nice thing about using the law as a case study is that it provides a wealth of examples for broadening the forms of frame principles.)

What exactly is the legal code?
The passage from the Massachusetts code quoted above lists the conditions

---

3 I am leaving out a few conditions, as can be seen in the passage from M.G.L. Also, certain exceptions are made for legal killing.
for murder in the first degree. This passage does a reasonably good job capturing the conditions for being a first degree murderer.\(^4\)

This passage from the Massachusetts code, however, is not a perfect statement of the law — that is, the frame principle. The law is distinct from the legal code, although the legal code plays important roles in the law. To see what the legal code does, we need to distinguish anchors from frame principles.

One issue with the passage I quoted (i.e., M.G.L. c. 265 §1) is the one I mentioned a moment ago, about grounding. On the surface, this passage looks as though it is an identity statement. It says of such-and-such behavior that it is first degree murder. But statements like these are better understood as giving the grounding conditions for having a social property. (Searle rightly stresses that being a first degree murderer is a status. Statuses are not identical to their conditions. Being a dollar is not the same thing as being printed at a certain bureau.) Thus this passage does not quite capture, or at least does not explicitly capture, what we want: the grounding connection between facts.

Interestingly, even apart from the issue of grounding, the Massachusetts law is still not exactly what is written in the code. First, although the legal code is an attempt to record the law, it is not a definitive attempt. The law can diverge from what is written in the legal code. Second, the legal code itself is a part of what anchors the law. I will explain these two roles, and then it will be easy to identify them in a grounding/anchoring diagram.

**i. The divergence between the law and the legal code**

The obvious source of divergence between the code and the law is that Massachusetts is a common law jurisdiction. This means that judges are bound by precedent, and thus that historical decisions by the judiciary anchor, in large part, the law itself. If there is a judicial tradition that diverges from the enacted statutes, that tradition carries weight in determining what the law is.

In addition, the legal code is also distinct from the statutes in force.

\(^4\) There is, of course, a difference between being a murderer and being found guilty of murder. The latter depends on people’s judgments, and the former does not.
CASE STUDY: LAWS AS FRAME PRINCIPLES

The Massachusetts code, for instance, is the best attempt of the House and Senate Counsel to represent the law. But not everything that is enacted ends up in the code. The courts routinely rely on unencoded legislation. Moreover, not everything in the code is formally enacted. And even the laws that are not enacted have force in determining what the law is.

Finally, you might notice that the passage above does not actually spell out the conditions for being a murderer. It gives the conditions for being a first degree murderer in terms of being a murderer. But the legal code lacks further specification of the legal conditions for being a murderer. Those conditions are something else that is anchored in part by practices in the judiciary. Interestingly, to track down the anchors for these conditions, we must look to things like the instructions that judges give to juries in homicide cases.5

ii. The legal code as partially anchoring the law

None of this means that the legal code itself is merely an inert, imperfect record. Rather, the legal code carries weight in anchoring what the law is. What is actually written in the legal code, in other words, is a part of the overall package that anchors the law.

It makes sense that the legal code should play these two roles. It is a practical document, an instrument for applying the law. In applying the law, we are interested in knowing the legal facts that obtain. Is Whitey Bulger a first degree murderer, or not? Under what conditions does a particular legal fact obtain? Is it legal to avoid taxes in such-and-such a way, or to copy such-and-such a document for personal use? These questions concern facts in our current frame. They are questions about the law as it has been anchored. The anchoring is complex, and in order to work out what the law is, we sometimes have to go back and look at the whole network of interacting anchors. But most of the time that complicated network is largely superfluous. Most of the time, we can just look at what we have recorded — the legal code itself — as our best stab at the grounding conditions. That is close enough.

That explains why the legal code is an imperfect record of the law. But why should the written code itself be among the anchors for the law?

FOUNDATIONS: OLD AND NEW

Because the legal code is part of the network of legal practice. Much of what is important in anchoring the actual law is what elements of the legal system do. And just as lawyers and judges are legal participants, the code is effectively a legal participant as well. Thus the code itself carries weight not just as a fallible record, but as a part of legal practices.

The grounding/anchoring diagram

Figure 7A is a rough picture of how Whitey Bulger is a first degree murderer is anchored and grounded. Grounding that fact is the fact Whitey killed Bucky Barrett with deliberately premeditated malice aforethought. The conditions for being a murderer are facts about the law, i.e., the frame principle for facts of the form \( x \) is a first degree murderer. M.G.L. c. 265 §1 is the House and Senate Counsel's attempt to record this law or frame principle. The law, in turn, is anchored by a variety of facts, including the history of trials and judicial decisions, actions of the legislature and the governor, and what the Counsel in fact recorded in the M.G.L.

Figure 7A Anchoring a law

It makes sense that we tend to think of the legal code as just being the law, even though it is not. After all, for simple things like first degree murder, the legal code does a reasonable job expressing the law. It is easy to overlook that the legal code is an imperfect statement of the law, and even easier to overestimate the anchoring role the legal code plays. But it is not an either-or choice: that the law is anchored exclusively by the text, or not at all. That false choice only reflects an impoverished theory of anchoring.

Secondary rules and their anchors

The bottom of figure 7A represents a particular anchoring. One particular set of anchors — a particular set of legislative actions, judicial decisions, jury
CASE STUDY: LAWS AS FRAME PRINCIPLES

instructions, the legal code, and so on — anchor a particular primary rule.

This anchoring is just one instance of a more general rule: a secondary rule. The rule of recognition gives the conditions for a primary rule to be legally valid. That is, it gives the anchoring conditions for a primary rule. The rule of recognition, then, is the general statement of which the bottom of 7A is an instance.

**The rule of recognition**

A rule of recognition needs to adjudicate among competing factors, in anchoring a valid law. Many factors are involved in this anchoring, including court hierarchies, various branches of the legislature and the executive, legal codes, and so on. Hart takes there to be one unified rule of recognition for a legal system, but we should expect this to have an extraordinary complicated antecedent. The bottom of 7A only gestures at the anchors that put law L in place.

Still, even without worrying about all the complexities, we can say a bit about the form of a rule of recognition. Start, for instance, with a statute enacted by a legislature. As I pointed out, that statute is not the law. So in the statement of the rule of recognition, we need to distinguish these. A rule of recognition may have something like the following form:

(7.1) For all y, if a statute y’ is enacted by the legislature and signed by the Governor, and y is the product of y’ together with the relevant influences of judicial decision, jury instructions, trial results, and statements in the legal code in such-and-such a way, then together those facts anchor y.

In this statement, I distinguish the law y from a statute y’. The idea is that y’ is but one component of the anchors of a law y, leaving room for the other anchors to have their effects in molding y. (Of course, this statement of the rule is still just a gesture. It also assumes that laws originate in statutes, which is not the way much of common law works.)

**Anchoring the rule of recognition**

As Hart recognized, the key question for the nature of law is not so much what the rule of recognition is, but what puts it in place. How, in other words, is this rule anchored? Here is where Hart’s “theory of practices” comes

---

Greenawalt 1986; Shapiro 2009, 2011, pp. 84-6
in.

The theory of practices is Hart’s theory of the conditions under which the rule of recognition is in place in society. The two conditions are (1) the conformance of behaviors and practices to the rule, and (2) the acceptance of the rule among the practitioners. This theory is closely related to Hume’s earlier theory of convention (and to later versions, such as David Lewis’s), and also to Searle’s subsequent theory of collective acceptance. Hart’s theory is another variant on the “Standard Model of Social Ontology.” Like Searle, Hart requires that members of the community have attitudes toward the secondary rule itself. Hart partly takes secondary rules to be anchored by attitudes involving R as a standard of behavior. And like Hume, Hart also adds explicit conditions on the actual behaviors: the behaviors must be regularly conformed to, and lapses actively criticized.

Just like these other views, we can depict Hart’s theory of practices as a component of the grounding/anchoring model as a whole. This is shown in figure 7B.

Figure 7B Hart’s anchoring of a secondary rule
At the left side of the figure are the anchors, according to Hart’s theory. That is, the facts about convergent practices and attitudes regarding R. These facts anchor the frame principle R, the rule of recognition.

Putting the pictures together
There is, of course, a relation between figures 7A and 7B. Figure 7B is different from all the other figures I have drawn, in one important way: what is anchored is not a frame principle governing the grounding of particular facts, but rather a frame principle governing the anchoring of frame principles. The anchoring depicted in figure 7A is governed by the frame principle anchored in 7B. That is, figure 7A is nested within the frame depicted in 7B.
This is what Hart is getting at in distinguishing primary and secondary rules. The secondary rules are the ones that set up the law-making principles for the primary rules. Primary rules are enacted within a frame governed by secondary rules. When, for instance, a law like killers with deliberately premeditated malice aforethought are first degree murderers is enacted, that involves the lawmakers satisfying the grounding conditions expressed in the secondary rule. Those secondary rules are, in turn, anchored (in Hart’s view) by the facts discussed in his theory of practices. Hart’s theory, modeled as a nested grounding/anchoring model, is depicted in figure 7C.

![Figure 7C Nested frames depicting the anchoring of secondary and primary rules](image)

Figure 7C Nested frames depicting the anchoring of secondary and primary rules

Again, on the far left of the diagram (marked as 2) are the “practices” — the convergent behaviors and attitudes toward those behaviors. Together, those facts anchor the rule of recognition. That rule gives the criteria for legal validity, i.e., the facts that must obtain in order for a law to be anchored in the frame.

As I did in the later diagrams in Chapter 6, I have left out all of the other possibilities in the frame. (Imagine how complicated the diagram would be if other possibilities were also depicted.) Still, it should not be forgotten that the secondary rule expresses conditions for any possible set of lawmaking facts in the frame. In a possibility where the relevant parties go through the relevant steps to enact and mold A, that anchors the fact that A is a law; and in a possibility where the relevant parties go through the relevant steps to enact and mold B, those steps anchor the fact that B is a law.

Now looking within the large frame, we see one actual set of facts
(1) anchoring a primary rule. These facts include the actions of the Massachusetts Legislature and the Governor, and also judicial practices, as well as facts about what is actually recorded in M.G.L. The rule of recognition governs the fact that these factors anchor the law on first degree murder. Then within the small frame is the actual grounding fact about what Whitey Bulger did (2), and the legal fact that his actions ground.

In some ways a complicated picture, but overall a fairly straightforward framework. One thing it clears up is the essential distinction between a frame principle and a set of anchors. This seems obvious, but is frequently elided. David Hume and David Lewis, for instance, often use the term ‘convention’ to refer both to the anchoring facts and to the frame principle. Likewise, Hart is vague about the application of the term ‘rule’, sometimes using it to refer to the regularities in behavior and attitudes. These are not the rule, but the rule’s anchors, i.e., the reasons that the rule is in place.

As I mentioned, the model also helps clarify some important distinctions in the subsequent literature on legal positivism. We can quickly use it to pick out various positions.

**The Hart-Dworkin debate**

Legal positivism holds that laws are legally binding in virtue of their social pedigree. It holds that the moral content of law is not relevant to its being legally binding. Not long after Hart’s book was published, Ronald Dworkin began a series of critiques in which he argued that moral considerations are indeed relevant to a law’s being legally binding. Judges routinely invoke moral considerations in their decisions, Dworkin points out. For instance, they overturn contracts on the basis of fairness. Consequently, says Dworkin, moral considerations figure into what the law is. Some philosophers rejected this, insisting that where judges brought in moral considerations, they were straying outside the boundaries of the law. This view has come to be known as “exclusive” legal positivism. Other philosophers, however — Hart included — replied that positivism can accommodate moral criteria for legal validity. It is not the criteria of legal validity that need to be strictly social and free of

---

7 Dworkin 1978, 1986
moral considerations. Rather, it is the source of those criteria. This view has come to be known as “inclusive” legal positivism.

A long description of the distinction between exclusive and inclusive legal positivism is unnecessary, since it can easily be shown using figure 7C. The differences between exclusive legal positivism, inclusive legal positivism, and various versions of anti-positivism, are differences among what theorists take to be the anchors for primary and secondary rules. ⁸

In figure 7C, ❶ marks the anchors of the primary rules, and ❷ marks the anchors of the secondary rules. Views on the sources of legality can be distinguished by their different commitments to what sorts of facts can be included in these categories. Different views are listed in Table 7a, along with what they take to be the anchors of primary and secondary rules.

Table 7a

<table>
<thead>
<tr>
<th>Theories of the “sources” of law</th>
<th>❶ Anchors of primary rules</th>
<th>❷ Anchors of secondary rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclusive legal positivism</td>
<td>Strictly social</td>
<td>Strictly social</td>
</tr>
<tr>
<td>Hart’s version of inclusive legal positivism</td>
<td>Mixed (no restrictions)</td>
<td>Convergent beliefs and practices about the rules</td>
</tr>
<tr>
<td>Inclusive legal positivism</td>
<td>Mixed (no restrictions)</td>
<td>Strictly social</td>
</tr>
<tr>
<td>Dworkin’s “legal interpretivism”</td>
<td>Both social and moral facts</td>
<td>Both social and moral facts</td>
</tr>
<tr>
<td>Certain religious natural law theories</td>
<td>Views differ on these</td>
<td>Divine sources, basic human purposes</td>
</tr>
</tbody>
</table>

Exclusive legal positivism insists that morality has no place in anchoring laws, either in the criteria for legal validity (i.e., in the anchors for the primary rules) or in the sources of those criteria (i.e., in the anchors for the secondary rules). When a judge invokes moral considerations in determining whether a law is legally valid, the judge is using moral facts in category ❶. This, according to the exclusive legal positivist, is an extralegal move.

⁸ Leiter 2003 and Shapiro 2007 are excellent presentations of these debates.

⁹ It does some violence to Dworkin’s view to place it in the table, since he rejects the distinction between primary and secondary rules.
Inclusive legal positivism allows moral considerations to enter into the determination of legal validity. Although the secondary rules are anchored by strictly sociological facts, that does not exclude moral facts from being anchors of primary rules. It may be our practice, for instance, to allow judges to weigh moral considerations in determining whether a law applies to a particular case. That allowance is part of the content of our rule of recognition because it is our common practice. The anchors of the secondary rules remain strictly social, despite the anchoring conditions in the secondary rules themselves having moral content.

Whichever position turns out to be most attractive, in large part the debate can be seen as one about anchoring. Theories of the criteria of legal validity are theories of the sorts of facts that anchor primary rules. And theories of the sources of legality are theories of the sorts of facts that anchor secondary rules.10

Model building and the grounding inquiry

In practice, most lawyers and judges are not concerned with the anchoring inquiry, i.e., working out the criteria for legal validity or the sources of those criteria. Nor are most lawyers and judges concerned with the grounding inquiry, i.e., working out what the laws (the frame principles) are. Mostly they are interested in the determination of fact. They take the law to be understood. Their work is to establish whether the particular grounding facts obtain. In the Whitey Bulger trial, for instance, there will be no discussion about the conditions for being a first degree murderer. The question will be whether Whitey satisfies those conditions. Did he kill? Was it with malice aforethought? This is neither the anchoring inquiry nor the grounding inquiry. It is what we might call the “actual fact” inquiry.

A related part of the practice of law is concerned not with the determination of actual facts, but instead with exploring possible facts. Much of the practice of tax law, for instance, is concerned with exploring different possible structures, to see whether they satisfy the conditions for being taxable. If Verizon buys out the stake held by Vodaphone in a certain way, will that be a taxable transaction? How about in another way? This is an

10 A recent proposal is Shapiro 2011.
invention of possible facts, and the application of frame principles to evaluate whether various other facts would ground the possible legal fact **The buyout is taxable**. This project we might call the “model building” project.

However, to do either of these projects well — to determine the actual facts or to build models of possible facts — we need to have a good account of what the laws are. That is, for somebody to have done the grounding inquiry well.

For legal “model builders,” the anchoring inquiry is not usually relevant. They are simply interested in having people get the law in our frame right — i.e., the grounding inquiry — and in using the relevant frame principles as the basis for thinking through possibilities. Only in certain cases would a model builder be interested in the anchoring inquiry. For instance, if someone were interested in modeling the effect of lobbying efforts on various legal participants, then she would need to consider how those participants figure into anchoring the law.

The same point applies to frame principles in general, and to the social sciences in general. The bulk of projects in the social sciences are “actual fact” inquiries and “model building” projects. This is why social scientists are often familiar with individualism of the sort discussed by Popper and Watkins, and not with the “Standard Model” and its advocates, such as Searle. When the aim is to determine what the actual social facts are, or how they change with changes in the grounding facts, what matters is getting the grounding conditions right. For most modeling it is more important to know what the frame principles are, than to know what sorts of facts put the frame principles in place.

In the next chapter, I turn to the distinction between ontological individualism — i.e., individualism about the grounds of social facts — and individualism about the anchors of social facts. Subsequently, I return to the reasons for making a sharp distinction between grounds and anchors, rather than just seeing anchors as a kind of ground.
Against conjunctivism

Ham sandwiches are not kosher. Ham is made from pigs, and pigs aren’t kosher animals. Why are pigs not kosher? Because according to the laws of “kashrut,” listed in Leviticus, for a land animal to be kosher, it needs to have a cloven hoof and to chew its cud. “The swine, because it parts the hoof and is cloven-footed, but does not chew the cud, is unclean to you.” (Lev. 11:7) These are the grounds for the fact that ham sandwiches are not kosher.

I have sharply distinguished grounds from anchors. The fact Ham sandwiches are not kosher is grounded by the following four facts: Ham sandwiches are made of ham, Ham is made of pigs, Pigs are land animals, and Pigs do not chew the cud. Different facts anchor the frame principle expressed in Leviticus 11:7. The anchors might be facts about beliefs in the community, facts about practices over time, or facts about divine commands. The anchors, however, are not among the grounds. The frame principle gives the conditions for grounding the social facts, and the anchors set up these grounding conditions.

Although this distinction seems natural — as it should! — it represents a sharp break from the prevailing orthodoxy. In this chapter, I want to explain and confront the dominant view, which I will call “conjunctivism.” This is the view that the grounds for a social fact include the anchors, in addition to what I am calling the grounds. Anchors, according to the conjunctivist, are just another kind of ground. Conjunctivism seems attractive, for reasons I will discuss. However, it gets the grounds wrong for many social facts.

The claim of conjunctivism

The conjunctivist makes a strong claim about social facts: every social fact has two different kinds of grounds. The grounds for a social fact include those I have been calling grounds, and also include the facts I have been calling anchors. Many followers of Searle, for instance, think that the
fact **Whitey is a murderer** is grounded by two different facts:

(9.1) **Whitey killed a person with deliberately premeditated malice aforethought**, and

(9.2) **People in the U.S. collectively accept that people who kill with deliberately premeditated malice aforethought count as murderers.**

According to these philosophers, the first fact alone does not suffice to ground the fact that Whitey is a murderer. More generally, they think that a fact \( z \) is \( Y \) is grounded by the conjunction of two facts:

(9.3) \( z \) is \( X \), and

(9.4) **We collectively accept that Xs count as Ys in context C.**

Implicitly, the conjunctivist believes that we are **impotent** to introduce social facts that are grounded non-conjunctively. No social fact, for instance, could be grounded only by (9.1). According to the conjunctivist, the grounds of all social facts conform to a two-part template.

In this chapter, I argue that conjunctivism is implausibly restrictive. It gets the grounding conditions wrong for many social facts.

Still, rejecting conjunctivism is no small matter. A sharp distinction between anchors and grounds involves a major change in how we model possibility. When we anchor a frame principle, we set up the grounding conditions — the full grounding conditions — for a given type of fact. The frame principle gives grounding conditions for all possibilities in the frame.\(^1\)

And those grounding conditions do not include the anchors themselves. In standard models of possibility, there is just one universe of all the possible worlds. Rejecting conjunctivism, however, suggests that we model possibility differently. A frame should be understood as containing a full universe of possible worlds. When we anchor new frame principles, we move to a different frame: that is, to a different universe of possible worlds.\(^2\)

---

\(^1\) Typically, at least. As I mentioned in Chapter 5, I am skeptical about necessitarianism about grounding. This issue, however, plays no role at all in the present chapter. (The only role the rejection of necessitarianism plays in this book is to allow me to distinguish grounding from determining.) All the frame principles I consider are frame-necessary. I am grateful to Dilip Ninan for raising this issue.

\(^2\) To be more precise, this can be modeled using multi-dimensional modal logics, with multiple universes, or using multi-frame logics, with multiple Kripke frames. See, for instance, Grossi 2007; Grossi et al. 2005, 2006, 2008. Grossi and collaborators do not discuss anchoring, but provides a series of relevant multi-frame logics.
AGAINST CONJUNCTIVISM

This is a big revision to standard models of possibility, and it is natural — and probably wise — to be skeptical. The case against conjunctivism is convincing, in my view, and rethinking our model of possibility is powerful and clarifying. Still, I expect it to be controversial, and so I have put it at the end of this part of the book. Even readers who favor conjunctivism do not need to stop reading here. None of Part Two turns on the denial of conjunctivism. Everything to follow is compatible with both conjunctivism and its denial.

The conjunctivist, however, should also be aware of another cost to her view. Conjunctivism makes the defense of individualism even harder. In denying conjunctivism, I am doing the individualist a favor. After all, the conjunctivist takes the grounds of a social fact to include both its anchors and its grounds. For the conjunctivist, there is no distinction between ontological individualism and anchor individualism. To be an individualist, then, one cannot just defend one sort of individualism or the other, but instead needs to defend both.

Why conjunctivism seems appealing

Let us suppose that Searle's theory of anchoring dollars is correct. And suppose that the piece of paper I called 'Billy' is a dollar. Isn’t it a requirement for Billy to be a dollar that the anchors obtain? After all, consider the following counterfactual: If we did not collectively accept the constitutive rule, then Billy would be just a piece of paper, not a dollar.

This counterfactual seems like a slam-dunk argument for conjunctivism. The counterfactual certainly seems to be true. And if so, it seems to contradict the idea that the anchors are not among the grounds. A reasonable way to understand this counterfactual is exactly the opposite of what I have said: the facts about collective acceptance are among the grounding conditions for the fact Billy is a dollar. This suggests that the grounding/anchoring model is deficient for capturing how Billy is a dollar is grounded. Instead, it suggests that figure 9A is a better picture:
The key difference between figure 9A and the grounding/anchoring model is that in 9A, the frame principle does not just articulate what the grounds are for being a dollar. Instead, the frame principle itself is part of the grounds for the fact Billy is a dollar. The frame principle, in turn, is grounded by the “collective acceptance” facts. We can still use the term ‘anchors’ to denote certain grounds of frame principles. But these anchors are just more grounds.

Accepting this alternative picture, the thesis of ontological individualism becomes a thesis about both grounds and anchors. On this alternative, when we say that the grounds of social facts must be individualistic, that means both the grounds for frame principles and the other grounds must be individualistic.

The argument in favor of conjunctivism seems clear enough. To see why it does not work, I first want to look at the counterfactual. At first glance, it seems like decisive evidence, but under scrutiny it turns out not be evidence at all. Then I will consider the version of conjunctivism depicted in figure 9A. That picture turns out to be flawed, and probably cannot be fixed. Subsequently I will return to the point that conjunctivism gets the grounding conditions wrong for many social facts. And finally, having argued against conjunctivism, I discuss why it does not make sense to take ontological individualism to be a claim about anchoring.

The counterfactual

If we did not collectively accept R, then Billy would be just a piece of paper, not a dollar. This counterfactual seems to be obviously true. And that seems to imply that collective acceptance of the rule is among the conditions for being a dollar.

Counterfactuals like these, however, are slippery pieces of data. In
fact, we can interpret this kind of counterfactual in many different ways, some of which are compatible with collapsing the anchors into the grounds, and some of which are compatible with keeping the anchors separate from the grounds. It does not provide evidence one way or the other. ³

A standard way to interpret the counterfactual is using “possible worlds semantics.”⁴ On this interpretation, the counterfactual means the following: take a look at the possible world that is nearest to the actual world, but in which we do not collectively accept the constitutive rule R. In that world, Billy is not a dollar, but is merely a piece of paper.

On the grounding/anchoring model, however, there is a different natural way to interpret a counterfactual. We not only have possible worlds, but also possible frames.⁵ So the antecedent of a counterfactual can shift us in two ways: to the nearest possible world in which the antecedent obtains, within a frame; or to the nearest possible frame in which the antecedent is an anchor, while we remain in the actual world in that frame.⁶ The example of money may make this difficult to see, so let us illustrate with a cleaner example.

Consider the fact \textit{x is a senior citizen}. The grounding condition for this fact is \textit{x is at least 65 years old}. This is anchored perhaps by law, by convention, by collective acceptance, or by practices. Now consider the following counterfactuals about Max, a 66-year old:

\begin{enumerate}
  \item[(9.5)] If Max were 60 years old, he would not be a senior citizen.
  \item[(9.6)] If we changed the laws, conventions, and practices such that the conditions for being a senior citizen are being 75 years or older, Max would not be a senior citizen.
\end{enumerate}

On the conjunctivist model, both of these counterfactuals have natural interpretations. The antecedent of (9.5) takes us to the nearest world in which Max is 60 years old. In that world, the anchors still obtain, so the conditions for being a senior citizen are the same. And so Max is not a senior citizen, and the counterfactual comes out true. The antecedent of (9.6) takes us to the

\footnotesize
\begin{itemize}
  \item ¹ I am grateful to Jody Azzouni for extensive comments.
  \item ² See Lewis 1973; Stalnaker 1968
  \item ³ For discussion of a closely related point, see Einheuser 2006. Einheuser gives a fragment of a two-dimensional semantics for “counterconventional conditionals.” This fascinating paper makes significant progress in distinguishing context-shifting and convention-shifting.
  \item ⁴ An elaborated multi-frame model will need to account for how worlds are identified across different frames.
\end{itemize}
nearest world where the laws are different. In that world, Max is 66 years old but the second condition for Max being a senior citizen does not obtain. And so Max is not a senior citizen in that world, and the counterfactual comes out true.

However, both of these counterfactuals have natural interpretations on the grounding/anchoring model as well. The antecedent of (9.5) is naturally read as world-shifting. It shifts us to a different world in the frame, i.e., the nearest one in which Max is 60 years old. Since it is within the frame, being a senior citizen has the same grounding conditions as it actually does, and so Max is not a senior citizen, and the counterfactual comes out true. The antecedent of (9.6) is naturally read as frame-shifting. It shifts us to the nearest frame in which the frame principle is anchored differently. In the actual world with frame principles anchored as in that shifted frame, Max remains 66 years old, but the conditions for being a senior citizen are different. So Max is not a senior citizen, and the counterfactual comes out true. Thus neither (9.5) nor (9.6) says anything in favor of either view.

Now consider one more counterfactual, whose reading is not so clear:

(9.7) If there were no body of laws, conventions, and practices, Max would not be a senior citizen.

To my ear, counterfactual (9.7) is ambiguous. On one reading, the grounding conditions for senior citizen — our grounding conditions — are left the same, and all that has changed is that Max is living in a different environment (one without certain laws, etc.). In that case, Max still satisfies our grounding conditions and is a senior citizen, and the counterfactual is false. On the second reading, we are considering a case in which the grounding conditions for being a senior citizen change, and in that context there is no kind senior citizen at all. On that reading, the counterfactual is true.

A reasonable interpretation of the ambiguity of (9.7) is this: counterfactuals involving social facts can be employed either to shift us to other possibilities within a frame, or can mark a shift of frame. In this counterfactual, either shift is about as natural as the other. The reason (9.5) and (9.6) are unambiguous is that it is clear what the antecedent does: the first does not shift the frame, and the second does.

But this is not the only reasonable interpretation. Counterfactuals
can be modeled in many ways, and the case I have described can just as easily be seen as a shift in the set of relevant possibilities as it can a shift between frames. The fact that (9.7) has two readings, in other words, is compatible with both conjunctivism and its denial.

Altogether, the point is not that the counterfactual is evidence in favor of the separation of anchors from grounds. Rather, it is that the counterfactual is not evidence one way or the other. It is compatible with both views, and cannot be used to decide between them.

A dilemma for constitutive rules
Thus the counterfactual does not favor either conjunctivism or its denial. Why reject conjunctivism? Let us start with the version I depicted in figure 9A. This seems like a natural picture for Searle’s theory.

In figure 9A, one of the two grounds of *Billy is a dollar* is constitutive rule R. That rule says that if x is a bill issued by the Bureau of Engraving and Printing, that *fully grounds* the fact that x is a dollar. Unfortunately, the diagram contradicts that rule. Satisfying the rule’s conditions only *partially* grounds the fact that x is a dollar. The other partial ground is the rule itself. So 9A cannot be the right picture.

Perhaps the answer is to change the diagram, leaving the rule out. This would give us a different figure:

![Diagram](image)

*Figure 9B The constitutive rule left out*
But this move does nothing to fix the problem. In this diagram, we take only the anchors, not rule R itself, to be among the grounds of *Billy is a dollar*. Yet even if we do this, the rule *still* shows up in 9B: it is what is collectively...
accepted. So even if we move to the new figure, we still need to fix the rule.\(^8\) If we do not, then the people in the U.S. are collectively accepting something false. (Again, \(x\) is a bill printed by the Bureau of Engraving and Printing only partially grounds \(x\) is a dollar.)

Unfortunately, the rule cannot be fixed. We cannot revise it, so that it gives us the full grounding conditions. If we include the full grounding conditions in the constitutive rule — that is, if we change it so that we collectively accept the correct conditions for being a dollar — then we fall into an infinite regress.\(^9\) To see this, let us try to revise the rule, so that it is both the thing we collectively accept, and also the correct rule for being a dollar.

There is not just one condition for something to be a dollar, but two. Therefore, rule R cannot be what we collectively accept. Instead, we need to replace R with R\(-\text{partial}):

\[
\text{(R\(-\text{partial}) \quad \text{If } x \text{ is a bill issued by the Bureau of Engraving and Printing, that partially grounds the fact that } x \text{ is a dollar.}
\]

R\(-\text{partial}) does not give the full grounds for \(x\) is a dollar: it only gives partial grounds. If we are to collectively accept a rule that gives us the full grounding conditions for \(x\) is a dollar, we need to collectively accept a rule that has conjunctive grounding conditions:

\[
\text{(R*) \quad \text{If } x \text{ is a bill issued by the Bureau of Engraving and Printing and we collectively accept R\(-\text{partial}, then together those ground the fact that } x \text{ is a dollar.}
\]

We have to accept R\(*, if we take ourselves to accept the correct conditions for being a dollar. Because only R\(* gives the conjunctive grounding conditions.

However, that is still not enough. After all, now we are collectively accepting R\(*, and that acceptance is also part of the grounding conditions for being a dollar. Therefore, R\(*) is not correct: it only gives the partial grounding conditions for \(x\) is a dollar. To fix this, we need to replace R\(*) with R\(*\)-partial — that is, substituting ‘partially ground’ in the place of ‘ground’:

\[
\text{(R\(*\)-partial) \quad \text{If } x \text{ is a bill printed by the Bureau of Engraving and Printing only partially grounds } x \text{ is a dollar.)}
\]

---

\(^8\) This same point applies to nearly all versions of the Standard Model, not just Searle’s. It applies to views in which attitudes are taken toward the rule itself. Hume’s theory is not included, but Hart’s is, as are other prominent theories in the recent literature, such as Raimo Tuomela’s “collective acceptance thesis” and Frank Hindriks’ revised constitutive rules. Hindriks 2009; Tuomela 2007, p. 187

\(^9\) Searle 1995, p. 52 discusses another potential regress. The present difficulty is not the same.
AGAINST CONJUNCTIVISM

(R*-partial) If \( x \) is a bill issued by the Bureau of Engraving and Printing \( and \) we collectively accept R-partial, then together those \( partially \) ground the fact that \( x \) is a dollar. And then the rule giving the full grounding conditions involves a bigger conjunction:

(R**) If \( x \) is a bill issued by the Bureau of Engraving and Printing \( and \) we collectively accept R-partial \( and \) we collectively accept R*-partial, then together those ground the fact that \( x \) is a dollar. It is R** that gives the conditions for being a dollar, not R*. But once again we find ourselves in the same situation. If we want to insist that we accept the rule for what it takes to be a dollar, then we also have to be accepting R**. And according to the conjunctivist, \( that \) acceptance has to be part of the conditions for being a dollar. So R** needs to be replaced with R**-partial. And accepting \( that \) has to be part of the conditions for being a dollar. Which takes us to R*** and R***-partial. And then to R**** and R****-partial. And so, ad infinitum.\(^{10}\)

To break the regress, we have two options. We can drop the claim that we collectively accept the conditions for being a dollar, or else we can drop conjunctivism. To put it differently: the conjunctivist has a Hobson’s choice. Either the rule we accept is false, or else it is limited to capturing a partial set of conditions for being a dollar. For the conjunctivist, the constitutive rule cannot accomplish what the “Standard Model” takes it to: namely, to give the conditions for a social fact to obtain.

I am no proponent of Searle’s collective acceptance theory, nor of Hart’s or any other theory following the Standard Model. Nonetheless, I doubt that the views of Hart, Searle, and others are so easily self-defeating.\(^ {11}\) I suggest we take Searle at his word: what we collectively accept is the rule for being a dollar, or a boundary, or a President. But that means that the anchors are not among the grounds: the acceptance of the rule is not itself among the conditions for something to be a member of its kind.

\(^{10}\) One might take a self-referential approach, along the lines of Peacocke 2005, but then it becomes implausible that we collectively accept the complex rules that would require.

\(^{11}\) The same argument holds for other leading theories taking a Searle-like approach. For instance, Tuomela’s and Hindriks’ approaches are each consistent only if the collective acceptance facts are not among the grounding conditions.
Conjunctivism gets the grounding conditions wrong

The argument in the previous section applies specifically to Searle’s and others’ “Standard Model” approaches. Given that argument, it is always an option to abandon the idea that constitutive rules do what the Standard Model claims: that we collectively accept some proposition that gives the grounding conditions for social facts of some kind. If we abandon that idea of constitutive rules, then conjunctivism could still be viable. The real test, however, is which approach gets the grounding conditions for social facts right. And this is the fundamental reason for rejecting conjunctivism. To include anchors as grounds would get the grounding conditions for many social facts wrong.

As I discussed in Chapter 5, social kinds and facts are a sort of “universal tool”: even though we may anchor their frame principles idiosyncratically, they can be grounded in any situation whatever. We can look back at ancient societies, and evaluate whether there are classes or castes, aristocrats or serfs. We might look for baristas in the Ottoman Empire or in 17th century England, and variable annuities among the ancient Egyptians. We might find that the Egyptians do not have variable annuities, but only proto-annuities. Or we might find that there is, in their context, an entity satisfying the conditions for being a variable annuity. Social kinds and social facts are applicable across a universe of different situations. In assessing the grounding of social facts across possibilities, we take the grounding conditions to be fixed.

To decide whether conjunctivism is right or wrong, we need to assess its implications across other possibilities. That is, to think about whether a given social fact can obtain in possibilities where the anchors for its frame principle do not obtain. Complex cases like money are hard to resolve. But for many examples, it should be relatively uncontroversial that conjunctivism is too rigid. Certain social kinds, for instance, we explicitly anchor to apply across all situations, even retroactively.

Consider, for instance, the conditions for being a war criminal. One is a war criminal if one has committed or conspired to commit any of a long list of crimes in association with armed conflict. We can sensibly ask whether Caligula was a war criminal, or whether Genghis Khan was, having killed over a million inhabitants of a single city. We can also consider a possibility
in which some virtuous person instead committed terrible crimes, and sensibly ask whether that person would be a war criminal. It does not matter whether, in that possibility, there is an International Criminal Court. What matters is only whether the person satisfies the conditions we have anchored.

To collapse anchors into grounds would improperly restrict the social facts to ones having two-part grounding conditions. It would restrict the social facts to ones whose grounding conditions not only include the ones we want them to have, but also all the anchors involved in putting the conditions in place. This is simply not how we use social facts. They can have simple grounding conditions. And when we assess them across other times and possibilities, we do not deny that they obtain merely because the anchoring facts do not obtain at those times and possibilities.

*Making ontological individualism sensible*

In the last chapter’s examination of ontological individualism, I developed it as a thesis about full dependence, taken in a generalized way:

(OI1) For any possible world w, and any social fact f at w, there is some subset X of N (the set of possible individualistic facts), such that X grounds f at w.

In that analysis, however, I avoided discussion of frames and anchoring. In particular, I said nothing about whether “all worlds” are the ones within a frame or across frames, nor did I separate this from a thesis about anchoring.

There is, however, a clear way to understand this, which is both consistent with and charitable towards the original ontological individualism tradition. The idea is to take the thesis to be about just one frame, usually our current frame. Anchor individualism is a thesis about how anchoring in general can possibly work. Ontological individualism is a thesis about how social facts in our frame can possibly be grounded, given that the anchors are what they are. I will briefly explain why this is the best alternative.

*Ontological individualism is not anchor individualism*

Some advocates of the Standard Model are tempted to regard ontological individualism as a thesis about what I am calling “anchors.” That, however, is not a viable option.

The reason is this: a social fact’s grounds have at least *something* to
do with making it obtain. Regardless of how we anchor the conditions for being the Senate, or for being the freshman class at Tufts, ontological individualism holds that the social facts depend on the people in those groups. A fact like The freshman class at Tufts is studious is grounded — at least in part, if not fully — by facts about the individual freshmen. That is what ontological individualism is a thesis about — about grounding, not about what anchors the conditions for being the freshman class at Tufts. Or maybe it is about both the grounds and the anchors. But it is not about the anchors alone.

If we take ontological individualism to be just a thesis about anchoring, and to say nothing about grounding, then supervenience just about always fails. It is too easy to set up situations in which two social domains are exactly alike in terms of the individuals and the individual relations composing them, and yet do not share the same social properties at all. Ontological individualism is at the very least a thesis about the grounds of social facts, i.e., about what sorts of other facts can possibly ground social facts. It is at least that.

No individualist should prefer a thesis about both grounds and anchors

Is ontological individualism then a thesis about both grounds and anchors? That is, about how frame principles can possibly be anchored and about how social facts can possibly be grounded in any frame? This is the position the conjunctivist must take. I have argued that it is an error to collapse anchors into grounds. I want to point out, however, that the proponent of individualism has an even greater stake than I do in keeping them separate.

Consider the anchors and the grounds for a social fact like Billy is a dollar. A theory like Searle’s takes the anchors to be individualistic —

---

12 Here is a more technical illustration of this. Take some property z that is neither individualistic nor social. Anchor a new frame principle as Searle might: we collectively accept that if an object has z, it has status Y. Now consider two different worlds. In the first, there are people, but no objects having property z. The second is individualistically indiscernible, but that there is also one additional object that has property z. In the second domain, there is an object having status Y. In the first domain, there is not. So supervenience of the social on the individualistic fails.
individual “we-attitudes” toward the constitutive rule.\(^{13}\) But Searle has no stake in taking the grounding conditions to be individualistic. The “X-term” of his constitutive rule can involve facts about pieces of paper and lines of stones — things which are neither social nor individualistic. The very idea of his theory is that our attitudes project features onto the brute world of non-persons.

In general, the anchor individualist should want to keep the grounds of properties like being a dollar separate from the individualistic anchors. The anchor individualist should have no interest in being an individualist about grounds.

Conversely, the individualist about grounds will take issue with anchor individualism. There are various strategies the individualist about grounds could pursue, in arguing for her claim. She might tightly circumscribe which facts count as the social ones. She might, for instance, argue for individualism about the grounding of certain macroeconomic facts, but not other facts. Or alternatively, she might add epicycles to the set of facts that count as individualistic. These are the sorts of strategies that theorists working in the tradition of Watkins, Lukes, Kincaid, Pettit, and others pursue.

But whichever moves are made to preserve the thesis that the grounds of social facts are individualistic, these are entirely different from those made by the anchor individualist. The individualist about grounds gains nothing by taking a position on how anchoring occurs. The collapse of anchors into grounds only makes her task harder.

To many people, both ontological individualism and anchor individualism are appealing for the same reason. Both seem to deflate worries about dualism with regard to the social world. The social world is just us, both theses hold. However, the two theses deflate these worries in conflicting ways. For an ontological individualist, the prototypical example of a social fact is one about a group, like a court or legislature, which is composed of individual people. The ontological individualist typically regards social facts as emerging from interactions among individual people, in combination with

\(^{13}\) Recall that in Searle’s view, a “we-attitude” is not a group-dependent attitude, but simply a different sort of intentional state that an individual can be in, different from that person’s “I-attitudes.” (Searle 2010, pp. 42-50)
one another. For an anchor individualist, in contrast, the prototypical example is a fact about dollar bills, unkosher animals, or boundaries made of lines of stones. It is not just that the strategies are distinct. It is that they are at odds with one another.

*An olive branch*

Having argued for the model, and for the sharp separation of anchors from grounds, I want to extend an olive branch to skeptics, agnostics, and dissenters.

The reason I argue for the sharp separation of anchors from grounds is that I think it makes for a better model, and that it illuminates the project of social metaphysics. It also makes it easier to see the flaws with anthropocentrism about the social world. One reason these flaws have been so hard to see is that the two different kinds of individualism have been confused with one another. And as I have said, however, separating the two theses does a favor for the individualist. It keeps individualism at least a little plausible, and makes us do some work to show why it fails.

The discussion of groups in coming chapters does not at all depend on the sharp distinction. I have found the grounding/anchoring model to be more powerful, and to produce more insights, so I do not see the point in hedging my bets. Still, all the rest of the points of the book remain intact, even for the conjunctivist. Conjunctivism does not compromise the force of the arguments before or after this chapter. If it turns out to be correct, that only means that the grounding inquiry I pursue is not an account of the full grounds, but only of partial grounds. And it means that the anchoring inquiry must be completed as well in order to get an account of the full grounds of social facts. Thus for the conjunctivist, the following chapters on the grounding of facts about groups will be part of the story, but not the whole one. Nonetheless, the same failures of individualism and consequences for modeling still apply.

*Grounding, anchoring, and the social sciences*

How is the grounding and anchoring of social facts pertinent to social science? How social facts are grounded matters for how we build models of
AGAINST CONJUNCTIVISM

them. And in a different way, how frame principles are anchored also matters for modeling. Of the two, grounding is the more directly relevant. As a result, the truth or falsity of ontological individualism is a more urgent matter for social science than is the truth or falsity of anchor individualism.

Suppose, for instance, there are too many first degree murders in our community, and we want to figure out how to reduce their numbers. Imagine a politician coming to the stump and proposing, “Under my administration, each of us will accept different conditions for being a first degree murderer in our community. Here is what I propose: we collectively accept that only people over 80 years old are first degree murderers. Then the rate would plummet!” Good idea, but not exactly the point. If we are interested in policy changes to reduce the incidence of first degree murder, we leave the grounding conditions fixed. If we want to model this incidence, we could model the factors that affect the grounds for different incidences of first degree murder. We might see what factors have a causal effect on killings, such as accessibility of guns. We might see what factors have an effect on premeditation, such as crime shows on TV. To reduce the incidence of first degree murder, however, we would not model changes in the anchors of the frame principle recorded in M.G.L. c. 265 §1. In trying to change the incidence of first degree murder, we are not normally interested in modeling what would change the conditions for having the status of a first degree murderer. Instead, we take the anchors to be fixed, and model what affects the grounds.

That is not to say that anchoring is always irrelevant to modeling in the social sciences. In some contexts, the anchors change rapidly, so we may be interested in assessing what affects them. On Wall Street, for instance, there is rapid innovation in financial instruments. New types of derivative instruments are continually created, not just simple options and swaps, but things like quantos, basket options, diff swaps, and so on. Sometimes they are created by explicit contracts, and sometimes by the practices of financial traders. Effectively, these people are anchoring new social kinds, setting up entities with different grounding conditions than there were before. Once such a kind is set up — that is, once various frame principles for facts about such kinds are anchored — then facts about particular instances of these kinds can be grounded. If we are interested in modeling financial markets, we
may just want to take the set of financial kinds fixed, anchored as they are, and see how changes in the world affect facts about them. But instead, we might want to model the dynamics of anchoring in these markets, to see how innovation can be affected.

Still, when we talk about the supervenience of social facts on facts about individuals, we are talking about grounding, not anchoring. When we talk about the emergence of social facts from facts about individuals, we are mostly talking about grounding, not anchoring. Most importantly, when we build models in the social sciences, we are mostly interested in the grounds of a set of possible social facts, and the causal factors that can affect those grounds. Typically, we take the anchors, and hence the frame principles, and hence the grounding conditions for social facts, to be fixed. And we model how changes in the world affect the grounds of social facts, thereby changing which social facts obtain.
Bibliography


BIBLIOGRAPHY


BIBLIOGRAPHY

BIBLIOGRAPHY


Analysis 51:192-94.


BIBLIOGRAPHY


