1 Laws and the Mind

Since the seventeenth century, our understanding of the natural world has been one of phenomena that behave in accordance with natural laws. While other elements of the early modern scientific worldview (i.e., that of the seventeenth and eighteenth centuries) may be rejected or at least held in question—including the metaphor of the world as a great machine, the narrowly mechanist assumption that all physical interactions must be contact interactions, and the idea that matter might actually be *obeying* rules laid down by its Divine Author—the notion of natural law has continued to play a pivotal role in actual scientific practice, in our philosophical interpretations of science, and in our metaphysics.

The history of philosophy since early modernity has been, in no small measure, an attempt to understand the things that are most central to our self-image as human beings against the backdrop of our best understanding of the natural sciences. This project, of course, has its fingers in many philosophical pies, including ethics, the metaphysics of freedom, the mind-body problem, and naturalized epistemology. In philosophy of psychology, one way it presents itself is as a set of problems about psychological and psychophysical laws. On the one hand, when one looks at the sciences of the mind, one might well come to doubt that they really possess anything as exact as natural laws at all, and to think that they possess only rough generalizations. On the other hand, to the extent that we are committed to the truth of physical *or* psychological laws, we would seem to be committed to denying the reality of free will. If the mind and the world are entirely governed by natural laws, there seems to be no room left for free will to operate.

This book seeks to address, and to work a form of philosophical therapy on, the apparent dissonance between the picture of the natural world arising from the sciences and our understanding of ourselves as agents who think and act. It is crucial here to distinguish between the sciences

themselves and their philosophical interpretations in the form of metanarratives in philosophy of science. The intuition that there is a principled difference between the generalizations one finds in the sciences of the mind and those one finds in physics and other natural sciences is based in part on considerations of the sciences themselves, such as the nature of the generalizations one finds in disciplines such as psychology and psychophysics. But it is also driven in large measure by a philosophical thesis about the nature of laws: that laws make universal claims about how objects actually behave. This interpretation of laws was the mainstream view of laws among Positivist and Logical Empiricist philosophers of science through much of the twentieth century. But it is not without alternatives, and indeed the Empiricist orthodoxy has suffered significant setbacks in recent philosophy of science. It is therefore incumbent on us to see whether familiar problems in philosophy of psychology can endure the great sea changes that have transpired in philosophy of science generally, or whether they are artifacts of an outmoded Empiricist view of science. The apparent conflict between laws and freedom is driven almost entirely by a particular philosophical understanding of laws, and is largely independent of the details of particular laws in psychology or physics. Here it is largely the philosophical issues that will prove telling, and again the crucial question is whether, on the best interpretation of the nature of laws, such problems actually arise.

This chapter presents a general overview of the problems and of how I intend to dissolve them. Let us begin by considering these problems in order.

1.1 Strict Laws and *Ceteris Paribus* Laws—Philosophical Problems

The first fundamental question would seem to be whether there are any psychological laws. In physics, we find an amazing variety of phenomena explained by a few simple equations. The laws can be verified with considerable rigor, the values of the physical constants can be determined with an almost arbitrary degree of accuracy, and the laws and models can be used to predict real-world behavior to such an extent that one can, say, send a spaceship to Mars. Our understanding of the mind presents a very different picture. There is a very great deal about the mind that we do not understand at all. What we do think we understand tends to come in the form of isolated insights and local models, and these models are often *informal* ones that lack the mathematical rigor of physical laws (and hence the preconditions for accurate prediction and confirmation).

Moreover, there are important philosophical views about the mind that are, to varying degrees, in tension with the idea that there are strict psychological laws. If there is libertarian free will, then at least some aspects of our psychology are radically anomic (that is, not governed by laws). If the soul is a non-material substance, it is not clear that we should expect its internal operations to be governed by laws (especially if we assume that it is a simple substance, one for which there are no component parts to interact in a lawlike way), and its interactions with matter have to be of an altogether different kind than interactions between two bits of matter. But one need not be a dualist to believe in free will. The assumption that we at least sometimes act freely is bound up in a much broader humanistic understanding of ourselves, one that is relevant to questions of ethics and responsibility as well as metaphysics. How to reconcile the humanistic assumption that we are free agents with our best understanding of the world of nature is, thus, an urgent question, regardless of one's assumptions about the metaphysical nature of the self.

There are, however, psychological and psychophysical generalizations of various sorts, and it seems reasonable that at least some of these should be regarded as laws. Some of these generalizations are, indeed, employed as laws in the models of various human sciences. Economics, for example, models economies as statistical generalizations over populations of idealized rational decision makers whose actions are (or tend to be) functions of their beliefs and desires. Psychophysics, a discipline that began to acquire experimental and mathematical rigor in the late nineteenth century, has produced robust laws describing the relations between the intensities of stimuli and percepts. And neuroscience is mapping the neural correlates of various types of psychological states, and finding relationships that are robust in ways that it is tempting to treat as lawlike.

If we are content to use the word 'laws' in these cases, there is still an issue lurking in the wings for the philosopher of psychology. Generalizations cast at the level of belief and desire, and even the psychophysical laws stemming from the work of Ernst Heinrich Weber and Gustav Theodor Fechner, seem to differ in important ways from, say, the inverse-square law or Coulomb's law. The most familiar physical laws seem to apply exactly, always and everywhere, and to be scale invariant. The psychological and psychophysical laws, however, seem fraught with conditions and exceptions. Even a classical economist, when pressed, will usually admit that real human beings are not ideally rational decision-theoretic agents. Generalizations about how people behave as a consequence of their beliefs and desires are not particularly good for predicting how an individual will

behave on a particular occasion, even if one assumes (and this is controversial) that the ways individual behavior diverges from the models will cancel out statistically over sufficiently large populations. And the psychophysical laws hold only against a host of background assumptions about the organism—for example, that we are dealing with a normally sighted subject who has not been staring at a light or a single saturated color, and who does not suffer from macular degeneration. And the list of qualifying conditions is long and potentially unbounded.

Philosophers of psychology have recognized a problem here for several decades. It is often framed as follows: Physical laws are "strict" or "universal and exceptionless," whereas human sciences (e.g. economics, psychology, even psychophysics) have only "*ceteris paribus*" laws (i.e. generalizations that are hedged by the caveat "other things being equal.") In itself this characterization may seem only an observation about a particular special science, or about the human sciences in general, and not a problem. It leads to a problem, however, when combined with several additional philosophical theses. The first of these is the claim that having strict laws is some sort of defining characteristic of "real" sciences. One might take such a view on *a priori* grounds, as the Logical Positivists did in their reasoning about "the logic of science," which they viewed as serving as a kind of norm for the actual practice of science. Or one might take it on the basis of observation of the status of the more developed sciences. One characteristic of mature sciences (one might think) is that they employ strict and exceptionless laws; thus, if psychology (or any other human science) is to become a mature science, it too must be framed in terms of strict and exceptionless laws, and the fact that it does not do so at present is at best an indication of its immature state and at worst an indication that it cannot be made scientific (at least in anything close to its present form).

Another worry is metaphysical. Some philosophers take the view that we may distinguish "natural" kinds from other *soi-disant* "kinds" (really artifacts of human interests) as follows: A kind is a *natural* kind just in case there are strict and exceptionless natural laws that apply to it. Artificial kinds may indeed submit to useful generalizations; otherwise they would have been of no use in the first place. But the difference between classifications that provide only a rough grip on the world and those that reveal how things of a certain sort always behave is itself an indication that in the latter we have hit on something real and fundamental, rather than an artifact of our own interpretation or interests. And as a consequence, one is inclined to view the real—the "really real," if you will, or the fundamental—as the nomic, the law-governed.

Of course, if you combine this view with the view discussed above—that psychology does not have strict laws—you are faced with a problem for psychology. It is perhaps most usefully and charitably framed as a dilemma: Unless strict laws can be found for psychology, psychological kinds will turn out to not be real in some privileged sense. And thus a lack of strict laws becomes not merely an interesting observation about psychology, but a call to the bucket brigade to help put out a dangerous philosophical fire and "vindicate" psychology, and likewise other sciences of the mind and indeed the human sciences generally.

Much philosophical discussion of these matters has accepted the general assumptions (a) that there is a difference between strict and exceptionless physical laws and *ceteris paribus* laws in psychology and other human sciences and (b) that the real is the nomic, yet has tried to dissipate the sense that there is a crisis for philosophy of psychology. Jerry Fodor (1974), for example, has argued that the messiness of psychological generalizations is not only explained but predicted by the fact that psychological kinds are functionally individuated kinds which are multiply realizable. The multiple-realization model of functional kinds, according to Fodor, underwrites both the ontological legitimacy of psychological kinds and the prediction that they should not have strict laws. I think this analysis yields some good insights for the philosopher of psychology, and indeed does so better when fleshed out with real examples than it does as pure philosophy of psychology. However, Fodor's response is only as good as the terms in which it is cast, and these terms have turned out to be problematic.

The whole distinction between strict laws and *ceteris paribus* laws is embedded in a set of deeper assumptions about the nature and the logical structure of laws. In particular, it assumes that laws are (true) universally quantified statements ranging over the real-world behavior of objects. This familiar interpretation of laws was made explicit by various Positivist and Logical Empiricist writers in the twentieth century. However, it came under withering attack in the last decades of that century. Nancy Cartwright, in particular, has pointed out that if we were to construe the laws of physics this way, they would turn out to be false. I would strengthen her case to say that most of the laws would have no true substitution instances—that is, that nothing ever behaves exactly as the gravitation law describes it. If this is the case, it is hard to see how psychological laws could fare worse. Cartwright seems to change her views, however, on the implications of this. Should we conclude that the laws are false? Or perhaps that they are true, but are not to be interpreted as universally quantified claims over objects and events? And if the latter, what characterization of

laws should we put into the place left vacant by abandoning the Positivist view of laws?

On this last question, Cartwright offers two suggestions, neither of which I ultimately consider satisfying. (See chapter 4.) However, I offer (in chapter 6) my own account of laws as idealized claims that pick out potential partial causal contributions to real-world behavior. This account shares with Cartwright's the virtue of dissolving the supposed problem for philosophy of psychology by rejecting the opposition between strict laws and *ceteris paribus* laws. It also yields more positive fruit: An investigation of the different types of idealization that may be at work in different laws and models both reveals the deep unity between physical and psychological laws and explains why the latter are, in intelligible and principled ways, messier than the former, in the sense of being less susceptible to integration with one another into a single "super-model" of the mind, and in terms of a greater rift between modeling and prediction.

In the case of physical laws, we are blessed to have a small number of fundamental forces, which are mutually independent. Given that we are dealing with more than one basic variable, this is a best-case scenario for getting a good fit between the theoretical goal of revealing deep invariants and the more practical goal of predicting real-world kinematics. Independent forces are factorable; they can be evaluated separately and then summed through vector algebra. Computational problems may arise from chaotic systems, but we are still in the best sort of scenario nature presents us with. With mind and brain, on the other hand, we are dealing with a complicated feedback system. When we model one part of it, we necessarily idealize away from facts about other parts that may matter crucially *in vivo* in modulating the behavior of the system we are studying. This kind of nonlinear, dynamic system is more complex than a system with only independent physical forces, and this kind of complexity makes the relationship between model and prediction much more tenuous.

My account of laws and modeling represents a view that I call Cognitive Pluralism. It is cognitivist in that it traces features of our scientific models to the cognitive process of modeling features of the world. Our models of the world, including our scientific models, are not simply reflections of how things are in their own right; they are idealized representations of particular features of the world, features taken in isolation, and represented in a particular representational system. It is pluralist in suggesting that the *de facto* plurality of models, and the apparent impossibility of integrating them all into a single "super-model" that allows us to explain everything at once, may be principled and abiding features of our science, rather than symptoms of science's immature state. The cognitivism and the pluralism

are related in the following way: I suggest that the reason for a principled plurality of models may lie in facts about *how* the mind models the world. If the mind necessarily understands the world piecemeal, through idealized models employing diverse representational systems, this may itself present significant barriers to the unification of the sciences. Models may be separate, inconsistent, or incommensurable, not because of anything about the world, but because of how the mind is constrained to understand it. At the very least, the alternative assumption that we *can* integrate our insights about the world into a single "God's-eye view" involves a significant empirical assumption about the nature of our minds—an assumption that seems improbable once one begins to consider it carefully.

1.2 Laws and Freedom

I contend in chapters 7–9 that my analysis of laws also resolves other problems for philosophy of psychology and its relation to other philosophical problematics, namely problems presented by claims for libertarian free will. Such claims might have been thought to conflict with the very notion that there *are* psychological *laws* on the ground that laws are universal and thus imply a kind of determinism incompatible with freedom.

My analysis of laws, however, shows that one can embrace the truth of individual laws, or indeed any set of such laws, without any implication of determinism, because the idealization conditions of each law are essentially open-ended. That is, no law includes a clause that says, in effect, "and this is the entire story about the universe." A gravitational law does not claim, for example, that dynamics is closed under gravitation. Nor does our commitment to gravitational laws plus strong, weak, and electromagnetic laws imply that the universe is closed under those forces. The truth of those laws is compatible both with the discovery of additional laws and with the possibility of genuinely anomic events, including voluntary spontaneity. Likewise, psychological laws, as idealized laws, do not claim to govern all possible behavior, but only to extract a partial list of real invariants in psychodynamics. In no way are further lawful invariants or voluntary anomic spontaneity excluded.

1.3 Are There Really Psychological and Psychophysical *Laws*? Case Studies

Parts I and II of this book take up the philosophical issues about psychological and psychological laws that have just been discussed. The kind of general and principled philosophical case developed in their chapters may

or may not prove persuasive to the reader. However, philosophy of psychology would be ill served if it were discussed only at this highly abstract level, without looking at the particulars of real psychological and psychophysical generalizations. The final three chapters, which make up part III, look more concretely at three types of such generalizations: laws of what Fechner called "outer" psychophysics, relating stimuli to percepts (chapter 10), computational neuroscience's network models of the dynamics of cortical systems involved in vision (chapter 11), and generalizations cast at the level of the common-sense inventory of mental states(such as beliefs, desires, perceptions, decisions, and actions) and relating two mental states (chapter 12). Both the "entities" and the "laws" turn out to look very different in these three cases, and to involve distinct philosophical issues.

In outer psychophysics, one finds things that look very much like natural laws in their form and perhaps in their robustness, even though they are implicitly hedged by a much larger number of background assumptions than are the most familiar physical laws. Chapter 10 takes this as an opportunity to flesh out more details of the general idealization account, and to explain how outer psychophysics is related to projects of localization and formal modeling.

The models of mechanisms accounting for psychophysical transformations examined in chapter 10 share many features with familiar models from the physical sciences. They involve laws relating quantitative data, and they provide explanations in terms of straightforward circuit-like mechanisms involving only feedforward causation. Chapter 11 examines the explanation of further psychophysical effects through "later" neural processes in the lateral geniculate nucleus and in the visual cortex, using a family of models developed by Stephen Grossberg and his associates at Boston University over the past three decades. These indeed rely on formal modeling techniques, but ones in which algebraic equations of the sort involved in laws play a far smaller role. They also involve complicated feedback processes. And this, I argue, results in a kind of idealization not found in basic physics, in outer psychophysics, or in the modeling of early vision. Mechanisms standing in a feedforward chain can, like independent causal forces, be factored, because the operation and the output of the earlier mechanism are independent of the operation of the mechanism standing later in the causal chain. As a result, one may model each system separately and then recombine them, using the outputs of the earlier mechanism as inputs to the later mechanism. But when two systems are related by a feedback loop, this is not possible. To model either system in

isolation—to idealize away from the contributions of the other in modulating its behavior—is distorting. Separate models of interconnected areas of the cortex, therefore, are idealized in ways that separate models of, say, gravitation and electromagnetism are not, because the latter are independent forces, whereas the processes modeled in the former case are, in real life, radically interdependent. This difference in the types of idealization employed in different modeling contexts results in very different kinds of gulfs between abstract models and real-world behavior.

Chapter 12 explores both common-sense belief-desire psychology and various attempts to regiment such psychology into more exact models, such as decision theory, Freudian psychology, cognitive/computational psychology, and explorations of knowledge representation in artificial intelligence (semantic networks, frames, scripts). Common-sense belief-desire psychology, I argue, does employ models of the mind, but these models lack many of the benchmark features of scientific models, including methodological and formal exactitude. It is, I believe, dangerously misleading to use scientific theory as a paradigm for understanding such processes as acquiring grammar or understanding other minds. However, the shortcomings of common-sense psychology by no means prohibit the development of truly scientific theories that either invoke belief-desire explanations in a more rigorous way or postulate other inner representational structures to explain features of belief-desire reasoning. Models of the sort explored in cognitive psychology and artificial intelligence seldom have laws in the form of algebraic equations, and often lack any quantitative element at all. However, computer programs and data structures are themselves an alternative form of formally exact modeling.

1.4 Modularity and Cognitive Pluralism

The case studies in part II also develop themes that support the general Cognitive Pluralist account developed in part I. The case studies, particularly those in chapters 11 and 12, all proceed by supposing that mind and brain have a number of distinct modules that represent particular parts or features of the world. The modules explored in chapter 11 plausibly are products of natural selection and are innate or at least strongly biased toward particular functions. Those explored in chapter 12 are acquired through learning and must be viewed as "soft modules"—partially autonomous structures that employ proprietary representational systems for their problem domains yet are acquired and fine-tuned through learning. This

supports the claims, made in chapter 6, that scientific modeling is but a specially regimented case of a more ecumenical phenomenon of mental modeling and that the abiding plurality of scientific models is a consequence of a basic design principle of human cognitive architecture.

1.5 Rhetorical Slant

This book has several rhetorical objectives. One of these is directed primarily at fellow specialists in philosophy of mind. There is a division among philosophers of mind on the question of whether philosophy of mind should be pursued as a largely autonomous discipline concerned primarily with the metaphysics of mind or should be pursued in close conjunction with recent work in the sciences of cognition and philosophy of science. I fall in the second camp, and I hope that this book will help to show the importance of taking both the sciences of the mind and recent philosophy of science into account when pursuing issues in philosophy of mind and philosophy of psychology. To the extent that philosophers of mind are engaged with philosophy of science, their discussions are often mired in philosophical views of science made popular by the Logical Positivists and the Empiricists—views that have been fairly decisively rejected within philosophy of science in the past several decades. For the purposes of this book, the most of important of these views is the Empiricist interpretation of the nature of laws. It is my claim that some problems in philosophy of mind and in philosophy of psychology are artifacts of an outmoded philosophy of science, and that they dissolve if one adopts more adequate and more up-to-date views.

The second rhetorical thrust is aimed both at fellow specialists and at the educated public. Since early modernity, people have worried that there is a kind of dissonance between the view of the world presented by modern science and the things about our own self-image that we hold most near and dear, such as the role of our mental states in determining behavior and the freedom of the will. On some philosophical views, the very reality of consciousness, beliefs, and desires seems to be threatened. Fodor (1990) puts the point eloquently: "If it isn't literally true that my wanting is causally responsible for my reaching . . . then practically everything I believe about anything is false and it's the end of the world." Likewise, we tend to assume that a commitment to the truth of laws involves a commitment to determinism, and hence to a denial of free will, and with it the abandonment of any moral evaluations that make sense only on the assumption that we at least sometimes act freely. Accepting modern

science is sometimes seen as implying that we must move "beyond freedom and dignity" (Skinner 1971) to a more mechanistic understanding of ourselves. This, too, is a view that many find quite threatening, as is evidenced by the number of works that attempt to argue that determinism ought not, on closer examination, to feel so threatening after all (e.g. Dennett 1984 and Flanagan 2002).

While I applaud attempts by convinced determinists to salvage human dignity and at least some types of (non-libertarian) freedom, I think they have bought too quickly into the assumption that a commitment to scientific laws implies a commitment to determinism. If my view of laws is correct, no such implication follows. There may be other reasons to be a determinist, and determinism may turn out to be true in the end, but we cannot get determinism out of scientific laws alone. And hence those who feel they have reason to believe in libertarian free will are free to embrace scientific laws without fear that doing so will compromise their commitment to freedom. This falls short of any sort of proof that we are, in fact, free. But it works a kind of philosophical therapy on a widespread but flawed way of coming to the conclusion that freedom is not compatible with what modern science tells us about ourselves and the world we live in.

1.6 How to Read this Book

The book is written for several audiences with different interests. Its principal theme is that questions of philosophy of mind are often best approached in close conjunction with explorations of the best that is offered by both the sciences of cognition and contemporary philosophy of science. Indeed, its main argument is that certain philosophical problems about the mind, such as the status of mental states and the possibility of human freedom in a world with natural laws, are attributable to unfortunate and outdated views in philosophy of science that can cause us to misunderstand what the sciences actually tell us about the mind and the world of nature. The primary goal of the book is to work therapy on these problems by applying a more adequate philosophical interpretation of natural laws.

Part I should be of general interest to readers concerned with issues in philosophy of mind. It also provides the background for parts II and III. Parts II and III, however, are largely independent of one another, and not every reader who is interested in the issues discussed in one of them will be equally interested in those discussed in the other. Part II deals with free

will, part III with a more detailed examination of how laws and other types of rigorous models are found in particular sciences of the mind. In part III, I occasionally make a brief return to issues of freedom; but the main argument in part II does not depend on those passages, and in other respects these two parts of the book are parallel and independent continuations of part I. Readers who are interested in free will but indifferent to case studies in philosophy of cognitive science will probably be more interested in the chapters of part II; readers who after finishing part I are still bothered by the question of whether there really are laws or rigorous models in the cognitive sciences may wish to read part III before returning to the material on free will in part II.