Preface

The philosopher Thomas Kuhn (1962) has argued that science often progresses by fits and starts. At critical times in the history of any discipline, an individual or group of individuals defines a theoretical framework. That framework then serves as a foundation for inquiry.

Many of the greatest brain scientists of the last four centuries have believed that René Descartes played that critical role in defining how we study the connection between the biological brain and the behavior of humans and animals. Pavlov, for instance, argued that for physiologists "Our starting point has been Descartes's idea of the nervous reflex," and other neuroscientists from Marshall Hall in the early 1800s to Charles Sherrington in the 1900s have expressed similar sentiments. Descartes's framework, or paradigm, described how the sensory energies that impinge on an organism give rise to appropriate behavioral responses.

In his book *The Structure of Scientific Revolutions*, Kuhn suggested that paradigms like this succumb to alternative frameworks in a two-stage process. First, he proposed, scientists begin to accumulate data that fit poorly into the existing paradigm. Second, the paradigm begins to be challenged by alternative frameworks that attempt to reconcile all available data into a single alternative conceptual approach. It is a central thesis of this book that about 50 years ago psychologists and biologists interested in brain function began to gather data that fit poorly into the existing Cartesian framework, and that at the present moment in neurobiological history a number of alternative frameworks are being developed and tested.

One of these alternative frameworks has been of particular interest to my research group, and since the early 1990s we have joined a rapidly expanding coterie of social, behavioral, and physiological scientists exploring this alternative paradigm. We and others believe that a mathematically rigorous and conceptually complete description of the neural processes which connect sensation and action is possible, and that such a description will ultimately have its roots in economic theory. In the 1970s, social scientists and economists largely completed the development of what is now called classical microeconomics. This theoretical framework was intended as a tool for describing, at a mathematical level, the computations that would be required if an organism were to use incoming sensory data and a stored representation of the structure of the world to select and execute an optimal course of action. This seemed a powerful tool for describing computations that the brain might perform, and it quickly became very influential, perhaps even foundational, in cognitive science. By the late 1970s, however, evidence began to accumulate that humans often failed to select and execute optimal courses of action in their day-to-day lives, and this called into question the utility of economic approaches as tools for social and cognitive scientists.

More recently, however, biologists have returned to economic theory, using it as a tool for studying the decisions animals make about what to eat or with whom to mate. These biologists have returned to economics on the assumption that in the environment for which an animal has evolved, the decisions the animal makes may more nearly approximate optimal courses of action than do the decisions of humans operating in our modern society. Although this assumption has been controversial, there is no doubt that in many cases economic theory has allowed us to predict and define the behavior of animals with tremendous precision.

Economic theory offers physiologists a second advantage, one that might be even more important than its often debated predictive power. Economic theory allows us to define both the optimal course of action that an animal could select and a mathematical route by which that optimal solution can be derived. Without a doubt, the nervous systems of animals cannot produce perfectly optimal courses of action, but it is equally true that they cannot ever produce courses of action that are better than optimal. Economic theory thus provides us with one critical tool for understanding the nervous system: It places a clear boundary on what is possible and allows us to ask what nervous systems do in that light. If that were all that economics offered physiologists, that might be enough, but it offers another critical advantage. It provides us with a language for describing the computational architecture within which *all* possible solutions can be computed. In this volume I argue, as a member of a group of social, behavioral, and physiological scientists, that economic theory may well provide an alternative to the classical Cartesian paradigm.

Descartes believed that all of behavior could be divided into two categories, the simple and the complex. Simple behaviors were those in which a given sensory event gave rise deterministically to an appropriate motor response. We call these behaviors reflexes after Descartes's use of the verb *réfléchir* in his book *Passions de l'Âme* (1649). The second class of behaviors Descartes identified were those in which the relationship between stimulus and response was unpredictable, or chaotic. These behaviors, Descartes proposed, were the product of a more complicated process he called the soul, but which a modern scientist might label cognition or volition. Since the early twentieth century, physiologists and philosophers have returned again and again to debate this dualist notion. Many have questioned whether there is any real need for Descartes's second (cognitive) mechanism. Could all behavior be explained by reflexlike mechanisms? In fact, many have quite reasonably wondered whether cognitive mechanisms can even be considered scientific notions.

In the subsequent pages I will make two arguments about these issues. First, I will argue (as many others have before me) that if cognitive mechanisms are defined using tools like those developed for economic analysis, then they are in fact quite scientific. Second, I will argue that reflexes are not scientific. To be quite explicit, I will argue that reflexes are a framework for thinking about the connection between sensation and action that is outdated and mechanistically inadequate; that at a physiological level there is no such thing as a reflex. At first that may seem a shocking claim, but I am actually not the first to make it. Many distinguished physiologists working during the twentieth century have also made that claim in one form or another.

In summary, then, like many others I will argue that the Cartesian dualism which has frustrated neurobiologists for at least a hundred years operates from a false premise. But I will argue that the false premise is that the reflex is a useful model for describing anything. Instead, I will argue that the reflex is a primitive model which works well only in overly simple "toy worlds," not in the real world that animals inhabit. A mathematically rich cognitive theory, however, would face no such limitations. It could, by definition, solve the most difficult problems that any environment could present. It would, almost by definition, eliminate the need for dualism by eliminating the need for a reflex theory.

This book, then, has two sets of closely intertwined goals. At a neurobiological level it champions a conceptual approach to understanding the nervous system that is being developed by a growing number of researchers. This conceptual approach begins by arguing that there are two main models of how the nervous system connects sensation and action: the reflex model and the cognitive model. I then challenge, and I hope disprove, the utility of the reflex model for understanding the nervous system. Without the reflex model, I go on to outline what a mathematically complete cognitive model might look like and how one would begin to test that model empirically.

At a philosophical level this book attacks dualism in a slightly unusual way. It begins by arguing that the central error of dualism is the belief in reflex-type mechanisms. Reflex-type mechanisms are attractive for many reasons, but an appropriately developed cognitive theory does not call for them, even in principle. Of course, the existence of a mathematically complete cognitive theory raises important questions about determinism, free will, and the stochastic nature of complex behavior, issues that are dealt with at the end of the book.