Discovering Complexity: Further Perspectives

"At least in biology, most scientists see their work as explaining types of phenomena by discovering mechanisms, rather than explaining theories by deriving them from or reducing them to other theories and this is seen by them as reduction, or as integrally tied to it." (Wimsatt, 1976).

1. The appearance of a new mechanistic philosophy of science

William Wimsatt clearly perceived the centrality of seeking mechanisms to explain phenomena in biology, but it was a perception not shared by most philosophers who persevered in the attempt to fit the life sciences generally, or biology in particular, into the model of nomological explanation that had been advanced by the logical positivists. The biological sciences fit poorly within this framework, a conclusion which we began to appreciate as students of Wimsatt's, and came to appreciate even more as we conducted the research that culminated in *Discovering Complexity*.

When we initiated the research that led to the book, we were still in the grips of the thenprevailing philosophical accounts of scientific reduction that proposed that reduction took the form of deriving one theory from another, and various revisions of that idea that have since been proposed—by Kenneth Schaffner (1993), Clifford A. Hooker (1981), Paul and Patricia Churchland (1990), John Bickle (1998), and others—some with specific focus on the biological sciences. We soon found these to be utterly inapplicable to the cases we had in hand. Not least of the reasons was that explanations did not take the form of subsumption of phenomena to be explained under a theory/law construed as a logically ordered set of propositions. Instead, there was a much more confusing picture. Theories seemed to play little role, while models were ubiquitous. Laws were sometimes mentioned - as in "Mendel's laws" or "laws of growth" - but they either didn't play a significant role in explanations, or they didn't fit the philosophical portrayal of laws as unrestricted nomic generalizations. In large domains of biologyphysiology, cell and molecular biology, biochemistry, genetics, evolutionary biology, and neuroscience-the word scientists invoked in explanation was mechanism. Despite recognition of the conflicts in biology between vitalists and mechanists and the perception that the mechanists had prevailed, philosophers of science had had little to say about mechanisms. Among the exceptions were Wimsatt and Stuart Kauffman (1971).

Our central problems in the book were to understand the elaboration and articulation of models in the life sciences, how they featured in explanations, how the explanatory problem was conceived, and how these tasks affected one another. In thinking about these problems, we wanted a flexible expression. Appeals to *mechanisms* were prevalent, and they didn't seem to carry much philosophical baggage. One of our challenges was how to characterize mechanism and mechanistic explanation, without becoming entangled in the conceptual machinery. We characterized a machine and mechanistic explanation as follows: "A machine is a composite of interrelated parts, each performing its own functions, that are combined in such a way that each contributes to producing a behavior of the system. A mechanistic explanation identifies these parts and their organization, showing how the behavior of the machine is a consequence of the parts and their organization" (p. 17).

In the ensuing seventeen years that situation has changed dramatically. Unbeknownst to us as we were finishing the book. Stuart Glennan was developing his own analysis in his dissertation at the University of Chicago with William Wimsatt. His analysis retained the centrality of laws: "A mechanism underlying a behavior is a complex system which produces that behavior by of the interaction of a number of parts according to direct causal laws" (Glennan, 1996, p. 66). As he filled out the account, mechanisms are invoked to explain laws, at least those that are non-basic, but appeal is made to other laws to explain the interaction of parts. More recently Glennan (2002) has replaced the language of laws with that of "invariant changerelating generalizations," a characterization he borrows from Woodward (2000). He also provides an account of interaction as a process in which "a change in a property of one part brings about a change in a property of another part." Machamer, Darden, and Craver (2000) offered an analysis in which "Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions." They rejected Glennan's emphasis on interactions (which was also prominent in our account) and were keen to emphasize the dualism of entities and activities. This dualism, although differently described, was already implicit in our discussion of task decompositions and localization as "the identification of the different activities proposed in a task decomposition with the behavior or capacities of specific components" (p. 24). More recently, Bechtel and Abrahamsen (2005) have embraced this dualism in distinguishing parts and operations as well as organization and emphasizing structural decomposition into parts and functional decomposition into operations. But they have rejected the linearity implicit in Machamer, Darden, and Craver's reference to "regular changes from start or set-up to finish or termination conditions" and in their most recent account of what they call dynamic mechanistic explanation have picked up on Glennan's language of properties, appealing to the "orchestrated functioning of the mechanism, manifested in patterns of change over time in properties of its parts and operations" (Bechtel & Abrahamsen, in press-b). We return to this below.¹

Reconceptualizing explanation as often involving the identification of mechanisms, rather than drawing inferences from laws, has a variety of additional consequences for characterizing the reasoning of scientists. Laws are typically expressed linguistically, and inference is commonly represented as logical derivation. We do sometimes see functional dependencies in the articulation of models, but these bear little resemblance to classical laws. Moreover, scientists often diagram mechanisms and add text, whether in the form of figure captions or the text of an article, to provide commentary and explication of the diagram. Some reasons for this are obvious. The parts of a mechanism are organized in space and even the reduction to two dimensions in a diagram can more readily capture such organization than text. Moreover, the interest is not just in the spatial organization of the mechanism but the processes of change going

¹ Philosophical work on mechanism has exploded to such a degree that we cannot hope to provide a comprehensive list of major contributions. A few of the more significant contributions that have focused on the biological sciences are: Bechtel (2006), Craver (2007), Darden (2006), and Tabery (2004). Thagard (2003) has explored mechanisms in medicine. Bechtel (2008), Thagard (2006) and Waskan (2006) have focused on the role of mechanistic explanation in psychology and the cognitive sciences while Ramsey (2008) has explored its applicability to chemistry. One question on which there has been significant controversy is whether natural selection should be construed as a mechanism or a description of patterns (Skipper & Millstein, 2005). Likewise, Craver (2008) has argued that the Hodgkin-Huxley model is not a mechanism so much as a description of the phenomena a mechanism must display. This has been contested, both by Kenneth Schaffner (2008) and Marcel Weber (2005, 2008).

on in it. Change is more difficult to represent in static diagram, but arrows often serve to indicate which part of the mechanism is affecting other parts. With developments in computer graphics and computer animation, mechanisms are often represented in dynamical displays. The visual portrayal of change within the mechanism plays the role of logical inference in traditional accounts of explanation for these portrayals show what phenomena can be expected to result from the operation of the mechanism. It is a plausible assumption that when scientists attempt to understand mechanisms without reliance on scale models or computer animation, they often rely on visual imagery to simulate the working of a mechanism (for further discussion of representation and simulation, see Bechtel & Abrahamsen, 2005).

Although animated graphic displays are a powerful tool for enabling scientists to understand mechanisms (and reveal much more complex interactions than the mental simulations scientists traditionally have performed in their heads), they are not magic. Rather, they typically rest upon mathematical representations of change in mechanisms, often employing differential equations, and calculations from these equations. Such appeal to calculations using equations to characterize mechanisms may seem to bring us back to the framework of laws and logical derivation. But what the equations represent are typically not laws, but empirically discovered or hypothesized regular relations between parts of a mechanism. These mathematical descriptions do capture dependencies, often quantitative dependencies, but they are often limited in scope and riddled with exceptions. Importantly, the solutions are often not generated by constructing proofs from these equations, but through numerical simulations.

The change in focus from characterizing explanation in terms of derivations from laws to understanding the role of mechanisms in generating phenomena provides a very different perspective on a wide range of topics in philosophy of science. We touch on some of these in the sections that follow, but given the increasing attention to mechanisms by philosophers, many other implications are sure to emerge in future years. We must, however, raise a note of caution. The account of mechanism which we advanced was motivated by particular examples of research in the biological sciences. While attention to cases has also been characteristic of much other work on mechanism (including that identified in the references in note 1), we perceive a risk that philosophical analysis may loose its foundations in actual science and have much of the a priori character of law-based accounts (whose origin was similarly in scientific examples). We think this would be regrettable. There is much that explanations in which scientists appeal to mechanisms yet has to teach philosophy of science, and this could be lost if the notion of mechanism were to be reified. Particularly exciting to us are cases in which the mechanisms invoked by scientists differ in interesting ways from those on which we focused and the accounts of mechanism offered to date in the philosophical literature. Attention to real science has the potential to surprise in ways far more illuminating than philosophical thought experiments.

In this context, it is worth noticing that history of science can be used in two ways—as a reservoir of examples to illustrate a prior conception of science or as a source of insight into features of science that an adequate philosophical account of science needs to incorporate. Our use of the history of science is in latter tradition. We focus on historical examples in parts of biology that at the time we originally wrote the book had attracted little philosophical interest; in particular, we included work on cell biology and cognitive neuroscience together with some more well-worn areas such as the history of genetics. We then built the account around the cases,

rather than selecting cases to illustrate the account. As we have said, our focus was on what counted as 'reductionistic' research in these fields. We soon discovered that, whatever else, these fields could not be naturally fitted to existing philosophical models of reduction without doing serious damage to the science, and the actual historical development of the sciences. As Wimsatt had observed, these reductionistically inclined researchers sought mechanisms to explain the phenomena they observed.

2. Discovery Heuristics: Conceptual and Experimental

One topic on which a focus on mechanism is likely to be particularly illuminating is scientific discovery. In the heyday of logical analysis of science, philosophers typically eschewed discovery, treating it as a topic about which little could be said philosophically. So, for example, in works such as Carnap's *Aufbau* (1928), Nagel's *The Structure of Science* (1961), Hempel's *Aspects of Scientific Explanation* (1965), or Popper's *The Logic of Scientific Discovery* (1959), the emphasis was on the 'logical structure' of scientific explanation, or scientific justification, or scientific theories. Logical analysis was conceived of as essentially a normative enterprise, whereas discovery was treated as a psychological matter. As the very titles to the volumes above suggest, the emphasis is on the structure of theories. (Popper's might seem the exception, but the German title *Logik der Forschung* (1934) is perhaps better the logic of investigation than discovery; and in any case Popper explicitly disavows the existence of a logic of discovery.) Theories were seen as axiomatically organized structures, and explanations were deductions within these structures. Insofar as discovery could be understood, that would be a scientific matter rather than a philosophical project.

The view of science was not without its detractors. Thomas Kuhn, along with historically sympathetic philosophical works such as Feyerabend's *Against Method* (1975) and Hanson's *Patterns of Discovery* (1958), recognized that the approach embraced by logical positivists fit poorly with the actual history of science. In reflecting on the implications of an historical turn in philosophy of science, Kuhn observed in the second edition of the *Structure of Scientific Revolutions* that "History, if viewed as a repository for more than anecdote or chronology, could produce a decisive transformation of the image of science by which we are now possessed" (1970, p. 1). Nonetheless, these astute critics shared much of the emphasis on structure, and shared not only the conception of theories, but the accounts of confirmation and explanation, that reigned. Through the logical lens, the historical turn appeared to yield 'incommensurable' theories and revolutionary change. Discovery was still not part of the territory within philosophy of science.

By the 1980s many philosophers of science took an historical turn, turning to the history of science for illumination as to how science works, including the process of discovery. A variety of different people took up serious 'case studies' in the attempt to understand discovery, several of which appeared in Nickles (1980b, 1980a). Lindley Darden, in *Theory Change in Science: Strategies from Mendelian Genetics* (1991) generated a distinctive account of discovery by focusing on strategies for theory revision, some of which lent themselves to formalization in computational models. Others, such as Paul Thagard in *Computational Philosophy of Science* (1988) and *Conceptual Revolutions* (1992), used computational methods to deal with historical cases. There were also psychologists who took up case studies, with computational models, such

as P. Langley, H. Simon, G. Bradshaw and J. Zytkow in *Scientific Discovery* (1987). By the early 1990's there were a great number of historians, psychologists, and philosophers living at the intersection of history, psychology and philosophy of science, many taking serious interest in discovery.

The distinctive focus we adopt in *Discovering Complexity* is on the reasoning strategies of the scientists engaged in mechanistic research, and especially on how these figured in the discovery of mechanisms. Our particular approach draws heavily on the work of cognitive scientists, especially the late Herbert Simon. Simon shifted our attention to a recognition of bounded rationality as a fundamental feature of cognition. In decision contexts generally, he argued, human beings make choices between alternatives in light of our goals, relying on incomplete information and limited resources. As a consequence, problem-solving cannot be exhaustive: we cannot explore all the possibilities which confront us, and search must be constrained in ways that facilitate search efficiency even at the expense of search effectiveness. If we think of problem solving as search through the space of possibilities (see Newell & Simon, 1972), limitations on search entail that for all but the simplest problems we can investigate at most a small proportion of the possibilities. Moreover, our evaluation of these possibilities will be uncertain and incomplete. We must rely on heuristic methods for pruning the tree of possibilities and also for evaluating the alternatives. In medical diagnosis, for example, the problem is one of inferring what disease is present from a given set of symptoms, when no one symptom is invariably present with a given disease and the same symptom can have various causes. (Even if a diagnostic test were decisive, so that the presence of some feature were invariably associated with a disease, there would be residual error in detecting the presence of the diagnostic feature.) Moreover, patients can be suffering from more than one disease, which compounds the difficulties of accurate diagnosis. Given that symptoms cannot generally be assumed to be independent, the size of the search space turns out to be large, and inferences concerning underlying diseases is inevitably uncertain. One solution is to limit the range of diseases considered to a narrow range of those antecedently likely. Good medical diagnosticians evidently narrow the range to no more than a handful of candidates. Search is limited in order to be efficient. This is bounded rationality.

Bounded rationality contrasts sharply with what Simon (1983) calls "Olympian rationality," embodied most clearly in the model of subjective expected utility, and in Bayesian models of human judgment. This includes Bayesian models of confirmation that still prevail in much philosophy of science. These models are often conceived of as normative models of decision making, by contrast with "merely descriptive" models garnered from empirical studies of human behavior. The difference between normative and descriptive enterprises was not always as sharp as it now appears (see Gigerenzer et al., 1989). Classical probabilists of the 17th and 18th centuries, for example, recognized no significant distinction between objective and subjective probabilities, much less between normative and descriptive problems. These classical probabilists sought a theory of "reasonable judgment," with all the ambiguity that suggests. We now are accustomed, living in Hume's shadow, to a sharper distinction between questions concerning how humans *do* reason and how they *should*. Economists, along with many philosophers, typically line up defending a theory of rational choice with a normative orientation, while psychologists often adopt a descriptive project. We are skeptical of any sharp distinction between to between questions about how humans *do* come to decisions, and how humans *should* come to

decisions. We clearly cannot assume that *however* people *do* reason is therefore appropriate; on the other hand, it would be nothing short of crazy to demand of us what we cannot do (cf. Stich, 1983; Flanagan, 1991). The relationship of descriptive and normative theories is at this point unresolved; in any case, when we are concerned with a theory of human choice, the limitations on human rationality cannot be neglected. Neither can the strategies we use to circumvent those limitations.

It was reflection on the factors which make Olympian rationality unrealistic as descriptions of the procedures of human rationality that led Simon, beginning in the mid-1950's, to focus on the roles of heuristics in human decision-making (Simon, 1979, 1983). Heuristic procedures, as Simon thought of them, are "rules of thumb" guiding our decisions and choices. Simon recognized that, in actual decision-making, humans rarely are simultaneously presented with all the available alternative modes of action. Even if we were thus informed, or if we simply restricted our decision procedure to those options known to be available, humans are not generally sufficiently aware of the set of likely outcomes attendant to those actions, much less the probabilistic array associated with each choice. Even given such comprehensive knowledge, we do not generally consider detailed scenarios for the future, with exhaustive and exclusive alternatives in place and conditional probabilities neatly assigned to them. Lastly, even granting us this wealth of information and the will to use it, our utility functions are notoriously ill-defined, in being intransitive, nonlinear, and encompassing incommensurable values. As Simon said,

The classical theory is a theory of a man choosing among fixed and known alternatives, to each of which is attached known consequences. But when perception and cognition intervene between the decision-maker and his objective environment, this model no longer proves adequate. We need a description of the choice process that recognizes that alternatives are not given but must be sought; and a description that takes into account the arduous task of determining what consequences will follow on each alternative (1979, p. 272).

Heuristics are just the sort of decision rules which recognize cognitive limitations and their impact on choices in complex circumstances.

Arthur L. Samuel's (1959) landmark studies in machine learning exhibit bounded rationality—this time artificial—as clearly as does human cognition. Checkers is a reasonably simple game, but complicated enough to tax time and search resources even for large computers. It is a well defined game in the sense that what constitutes a win, lose, or draw is sharply defined, as is the space of board positions. Samuel observes that there is no known algorithm which will guarantee a win or a draw in checkers. Moreover, since an exhaustive search would require exploring something on the order of 10⁴⁰ possibilities, there is no real alternative but to engage in selective search. That is, the only tractable procedures are ones which do not guarantee ideal solutions, and may not guarantee solutions at all. In the case of checkers, it is possible to enumerate the possible moves from a given board position, and to evaluate outcomes. Samuel's program evaluated board positions in terms of the ability of the players to move and the relative number of pieces. The evaluation of moves from an evaluation of the associated board positions followed a minimax procedure, choosing the move which maximizes the machine score

assuming that the opponent will respond in order to minimize it. In other words, a move was selected provided that among the possible moves, its worst outcome was better than the worst outcome associated with alternatives to it. Samuel programmed the machine to search for appropriate heuristic procedures for evaluating moves—in effect, to learn checkers. Yet the machine was inevitably limited in resources and its search reflected this fact. Samuel's machine, for example, limited how many moves it looked ahead depending on a number of variables, including whether a jump is available at various depths. In the limit, the program stopped at twenty moves ahead. That is more than humans generally pursue, but it is nonetheless a real limitation on search. Samuel's machine, no less than its human competitors, suffers from bounded rationality.

This perspective will be familiar to psychologists, and to those working in behavioral economics. It forms the heart of work by Amos Tversky and Daniel Kahneman (see, for example, classic studies included in Kahneman, Slovic, & Tversky, 1982), and a host of others, exploring the specific heuristics underlying human judgment. The comparison of heuristics and normative standards (that is, algorithms in a narrow sense) allows us to examine the "discrepancies" between actual and optimal behavior. The point is not, though, to press the (wholly unremarkable) conclusion that we are all too frequently not rational animals. That point hardly needs a research programme to make it evident. The point instead it to determine what the psychological mechanisms are which underlie our behavior. Whether Tversky and Kahneman are right or not in their description of the cognitive mechanisms, they are surely right that we do not rely on the Olympian standards. There is agreement on this point even from their most astute critics. For example, Gerd Gigerenzer and the ABC group recognize the sources of Olympian Rationality in the account of probabilistic reasoning that emerged since the 17th century. Gigerenzer and Todd say this:

The probabilistic revolution has shaped our picture of the mind ... Mental functions are assumed to be computations performed on probabilities and utilities (Gigerenzer and Murray 1987). In this view, the laws of probability describe or prescribe sound reasoning, judgment, and decision making (Gigerenzer, Todd, & the ABC Research group, 1999, p. 6).

Notice that this description too is not free of a normative cast, any more than is the work of Kahneman and Tversky. Yet Gigerenzer's own views are free of any Olympian tone. Gigerenzer and Todd are careful to situate their views under bounded rationality. They say this:

Fast and frugal heuristics employ a minimum of time, knowledge, and computation to make adaptive choices in real environments. They can be used to solve problems of sequential search through objects or options, as in satisficing. They can also be used to make choices between simultaneously available objects, where the search for information ... about the possible options must the limited, rather than the search for the options themselves. Fast and frugal heuristics limit their search of objects or information using easily computable stopping rules and they make their choices with easily computable decision rules. We thus see fast and frugal heuristics as two overlapping but different categories of bounded rationality (Gigerenzer et. al., 1999, p. 24).

Here, though, the similarities between the two approaches to understanding heuristics fade considerably.

Philosophers have not been wholly oblivious to the role heuristics play in human reasoning. One of the most influential philosophical advocates for heuristic models is Wimsatt. The importance of heuristic procedures is integral to his recent book, *Re-Engineering Philosophy* for Limited Beings: Piecewise Approximations to Reality (2007). Wimsatt identifies a wide variety of properties characteristic of heuristics (2007, pp. 39, 68, 76-77, appendix A). The key contrast is with algorithms. Whereas algorithms are procedures which are guaranteed to yield a correct solution to problems for which they are designed, heuristic procedures are not. The advantage heuristics offer to us lies not in accuracy, but in economy. Heuristics are less demanding in terms of time, memory, and computational requirements. Still, if they are useful they are accurate enough in the key cases to which they are applied. Any useful reasoning strategy must presumably be reliable, if not infallible. Michael Bishop and J.D. Trout similarly promote a set of "reasoning strategies" which they argue are "robustly reliable" (2005, p. 71 ff.). These are heuristic rules. These are reasoning strategies that are *resilient* across a wide range of environments. Bishop and Trout do not contend that these strategies are unerring, and they're not. And they notice that though a strategy may be reliable in one domain, it may utterly fail in another. That is, errors introduced by relying on heuristics will tend to be systematic. They will fail in characteristic classes of cases, and the errors are predictable. This was integral to much of Simon's experimental explorations of human reasoning, and it is central to what we do in Discovering Complexity. We are not, Wimsatt explains, "LaPlacean demons" for whom computational requirements are incidental. The "Re-Engineering" Wimsatt recommends derives from this simple observation together with the thought that much philosophy proceeds as if we were just such computationally unencumbered decision makers. Wimsatt writes:

"A more realistic model of the scientist as problem solver and decision maker includes the existence of such limitations and is capable of providing real guidance and a better fit with the actual practice in all of the sciences. In this model, the scientist must consider the size of the computations, the cost of data collection, and must regard both processes as 'noisy' or error prone" (2007, p. 78).

Wimsatt doesn't actually explore the human limitations which would define this more realistic model of the scientist directly. He does explore an array of methods deployed within science, and problems that bear the earmarks of heuristic procedures. Models commonly deployed within the biological sciences, for example, embody systematic idealizations that are justified primarily because they simplify problem solving. The ubiquity of such models has a number of implications that are important.

First, the fact that there are systematic failures facilitates the idea that the elaboration of models is a self-correcting process. "False models" are, in Wimsatt's terms, the "Means to Truer Theories" (2007, ch. 6). Linkage mapping in early 20th century genetics is a case in point. A linear model of chromosomes, with the chances of recombination between pairs of genes dependent on their separation, violates the naïve Mendelian ratios, but does so in a systematic way. These linked traits exhibit non-random patterns of inheritance. If three genes are arrayed on a single chromosome, it is possible to determine their relative locations on the chromosome

using the degree to which they depart from the Mendelian expectations. Suppose that there are three genes A, B, and C, arrayed on a chromosome in that order. They should all exhibit some linkage. If so, then the chances of recombination between A and C would be expected to be the sum of the chance of recombination between A and B and that between B and C. This was the assumption Morgan and Sturtevant (see Sturtevant, 1913, 1965) began with. The deviation from the naïve Mendelian ratio is a measure of their linkage. However, the simple additive model also fails, even though it is approximately true. The experimental data showed that the recombination frequency of the more distant genes is sometimes *less* than would be expected based on the other recombination frequencies. This in turn can be explained as a consequence of multiple instances of cross over, which implies that the observed recombination frequency systematically underestimates the actual distance between them; as a result, the simple additive rule will break down. The simpler models in both cases break down. The pattern to that failure in turn became the lever for constructing a more adequate model. As we explore cases in both biochemistry and cognitive neuroscience, it is the failures again that serve as a corrective to the heuristics that initially guide research.

Second, Wimsatt turns the fallibility of the procedures into an epistemic advantage. The key concept is *robustness*, which is allied to Donald Campbell's idea of *triangulation*. Wimsatt says "Things are robust if they are accessible (detectable, measurable, derivable, definable, producible, or the like) in a variety of independent ways" (2007, p. 196). If we have available a number of independent, but uncertain, pieces of evidence which all point toward the same conclusion then that conclusion may in fact be extremely probable even if each piece of evidence is uncertain. By contrast, even a highly reliable procedure will have some likelihood of failure. It is a simple exercise in probabilities to see that the former may be superior to the latter. After all, the likelihood of getting an error in the latter case is the sum of the likelihood of errors in each step. Think of adding a long series of numbers. At each step there is a small chance of error, but the likelihood of some one error can be quite high. By contrast, with independent but unreliable pieces of evidence, the likelihood of error is the likelihood that all the evidence is wrong. If there are many witnesses to a crime, and they agree on the details, the chances of error are reduced.

Of course, this depends critically on the various measures being independent of one another. When the measures are not genuinely independent, we have a case of "pseudorobustness." These are artifacts. The biases in models of group selection are an interesting case that Wimsatt has discussed in many places. Commonly, these models assumed that all the migrants from groups are pooled, and then new groups are formed by randomly drawing from this migrant pool. The assumption is one that makes the models analytically tractable. However, it virtually guarantees that there will be no group selection. The fact that these models all indicate there is no group selection does not really give independent reasons for the conclusion (2007, pp. 84 ff.).

While our focus is on the cognitive strategies, and on their role as constraints facilitating the elaboration of mechanisms, there are engaging problems we noted but do not explore at length. Importantly, we did not focus on the experimental strategies which reveal the structure within complex systems. There are various strategies—and these too are heuristic procedures—for exploring the structure of a complex system. We noted that the evidence for mechanisms often resulted from experimentally lesioning or stimulating parts of mechanisms. In other cases,

evidence for mechanisms is the result of accidental damage. Inhibitory or deficit studies allow us minimally to determine which components within a system matter to the systemic behavior. The most dramatic recent cases in this category are genetic "knockouts" which reveal a great deal about normal functioning. Excitatory studies offer a complementary sort of evidence. In the study of developmental mechanisms, it is possible to implant regulatory genes and see how that affects development. The most salient cases, perhaps, are the regulatory genes controlling segment development in Drosophila. Though we identified the strategies, we did not focus in much detail on the procedures by which this is accomplished. These are only beginning to receive attention in philosophy of science, often inspired by social studies of science (Carl Craver, 2002). This raises a number of important philosophical issues. Experimental investigations are distinguished from observational studies in that they involve active manipulation. Often the manipulations are extremely violent to the system being investigated. For example, biochemical studies often begin by breaking apart the cell and homogenating its contents. In other cases, neurological studies often begin by disabling biochemical or neurological pathways. The risk is that the results of the experiment may be due more to the manipulation and unexpected consequences of the manipulation than to the operations in the mechanism. Scientists label such results artifacts and often the most contentious arguments between scientists are over whether a particular result is an artifact (Bechtel, 2000). The point we want to emphasize here is that these broad experimental techniques are themselves heuristic procedures. The problem is that we do not know beforehand exactly what the patterns of failure will be, and so we cannot readily disentangle fact and artifact.

3. Decomposition and Localization

Having characterized scientific discovery as relying on heuristics, we turn to the heuristics that figured in our account of the discovery of mechanisms. Assume for the moment that nature is hierarchically organized. Simon captures that idea under the title of "the architecture of complexity" (1969, p. ch. 7). Wimsatt (1974, 1976, 2007) has a characteristically more reticulated treatment of the 'levels' of nature. Craver (2007) explores perhaps the leanest account of levels that will suffice for mechanistic models in which levels are only defined locally in terms of the working parts of a mechanism. Hierarchical systems are organized around interrelated subsystems, with discriminable capacities. Simon thought that, as a matter of fact, not all systems exhibit interesting forms of hierarchy, but many-perhaps most-he thought are hierarchically organized. Among those that do exhibit some degree of hierarchical order, the key feature concerns the relative strength of interaction *among* as opposed to that *within* subsystems. The clear and straightforward thought is that as interaction among subsystems increases, the significance of interaction within subsystems decreases, and conversely.² So in relatively simple hierarchies, there will be a relatively high strength of interaction within subsystems as compared with the interaction among subsystems. These are systems that are at least nearly decomposable (Simon 1969). Here is what Simon says:

² Discussions concerning "modularity" are affected by this issue. Some claim there is "massive modularity," intending by this that the mind/brain is organized in a way that achieves relative independence for some parts, or that most functions are realized in discrete structures; others argue that there is more "plasticity," evidently intending that there is more modifiability in brain structures. We think the controversy is spurious. The real question, which few address, is the extent of local control as opposed to global control over systemic behavior.

At least some kinds of hierarchic systems can be approximated successfully as nearly decomposable systems: ... (1) in a nearly decomposable system the short-run behavior of each of the component subsystems is approximately independent of the short-run behavior of the other components; (2) in the long run the behavior of any of the components depends in only an aggregative way on the behavior of the other components (1969, p. 210)

One clear thought is that nature is organized in terms of wholes and parts, mereologically. Another is that there is a kind of order to this whole-part ordering. Organization sometimes is hierarchical without being decomposable; but when it is decomposable, or nearly so, there is a very natural kind of order. In the decomposable cases, there is a natural projection from component behavior to systemic behavior, both in the short and long term. If we have a failure of strict decomposability, but still retain *near* decomposability, then a projection from component behavior will yield at least a first approximation to actual systemic behavior. The behaviors of parts are then relatively independent of the behaviors of other parts, on both longer and shorter time-scales. In other cases this is not so; in these non-decomposable cases, the behavior of one component depends critically on the behavior of others. We'll return to this kind of case, since we regard it as extremely common in the life sciences. Yet even in these non-decomposable cases, research that assumed the system was decomposable has often been highly informative.

Decomposition and localization are the twin strategies—and we mean to say they are heuristics—we discuss for discovering mechanisms, and for articulating their structure. One kind of decomposition is essentially structural. So, for example, in the late 18th century, the great chemist Antoine-Laurent Lavoisier (1743-94) ingeniously managed to resolve water into its constituents, oxygen and hydrogen, by passing steam through an iron gun barrel. He was able to measure the weight of the hydrogen expelled at the far end of the barrel, as well as the increase in the weight of the barrel due to the oxygen that reacted with the red hot iron. It matched the weight of the original water. While this is important as a strategy, the kind of decomposition we focus on is not structural, but functional. Essentially, we begin with the assumption that a system has the simplest sort of hierarchical structure. Researchers assume that some phenomenon they seek to explain is due, at least in part, to different parts within the mechanism performing their own operations. The research challenge is then to decompose the phenomenon into the component operations and localize them within the parts of the mechanism.

In calling decomposition and localization heuristics, we are emphasizing that they are fallible research strategies—a given phenomenon may not be due to component operations discretely localized in component parts of a mechanism. Indeed, this is often not so. In those cases, mechanistic models elaborated by assuming decomposition and localization will fail to capture systemic behavior. There are many reasons this might be so. We might, for example, have failed to identify the actual component functions. In other cases that we regard as both common and significant, the very error is in assuming the system is decomposable or even nearly decomposable. In such cases, the failure of the mechanistic model may reveal that there is a more complex organization than was initially assumed. This echoes Wimsatt's (1987) suggestion that false models are sometimes the means to truer theories. In either case, the elaboration of the mechanistic models can be a means of discovery. That is the point of focusing on the discovery of complexity.

Critics of mechanism—vitalists and holists in the 19th and early 20th century, and many contemporary dynamicists—have rejected the quest for mechanistic explanation through decomposition and localization. In many cases, the rejections of particular mechanisms are well founded. The wisdom in their tout court repudiation of mechanistic research is less clear. An example may help here. In exploring the complexities of development, Wilhelm Roux (1831-1924) initiated an experimental version of *Entwickelungsmechanik*—literally, developmental mechanics-in support of the idea that there are internal determinants of development, and simultaneously minimizing the impact of external conditions. He had been led to a "mosaic" theory of development, according to which the hereditary determinants are distributed in a qualitatively uneven way within the fertilized egg. As the cell divides, the daughter cells are genetically differentiated and these differences explain the differentiation of cells within organisms. In a classic paper (1888), Roux describes experiments designed to test the idea of embryonic self-differentiation. At the first cleavage in the development of a frog, he destroyed one blastomere with a hot needle. The critical question was what happened to the remaining blastomere. In about 20% of the cases, the remaining blastomere continued to develop. Roux observed that the remaining blastomere did not compensate for the loss of its twin, but developed into half an embryo. He concluded that the blastomeres develop independently, depending primarily on their internal constitution. This supported Weismann's and Roux's view that development was controlled by material that was successively divided among the cells of the organism. This material, he thought, determined the growth of the organism in a fully mechanistic form.

In 1891, Hans Driesch (1867-1941) performed what seemed at first to be a very similar experiment, but with dramatically different results. Using sea urchins, he separated the blastomeres at the two cell stage. Instead of Roux's result, Driesch found that each blastomere developed into a smaller, but complete organism. He saw this result as inconsistent with Roux's mechanical account, and in particular as inconsistent with the idea that division of the cell involved a division of the "germ" controlling development. Since the blastomeres has the ability to develop into complete organisms, there could not be the kind of differentiation Roux had observed at this stage of development. Each had to have the full genetic complement. Driesch initially sought external, epigenetic, factors to explain development. Unlike Roux, Driesch saw development as the response of a whole living organism rather than a mechanically predetermined process, or, as he put it, as a "harmonious equipotential system." Eventually, he turned to a straightforward vitalism.

Driesch was of course right to reject Roux and Weismann's mechanistic model of development, even if his vitalism did not prove to be a viable alternative. Roux's initial mechanistic model was flawed. However, it's crucial to see that without elaborating and testing the mechanistic model—and finding, essentially, that the original decomposition was incorrect—scientists would not have turned to more satisfying alternatives. As a vehicle for discovery, the elaboration of the mechanistic model was fruitful even though the resulting model was incorrect. Once again, false models may be the means to better theories. Many of these, as in the case of developmental biology, will be accounts of mechanisms with greater ability to account for the phenomena that the critics saw as inimical to any mechanistic account. Contemporary research in developmental systems theory, for example, often identifies mechanisms, albeit ones that are far

more complex and integrated than those advanced by early pioneers such as Roux (see, for example, Oyama, Griffiths, & Gray, 2001; Stotz, 2008)

The task of decomposing a mechanism is often extremely challenging for those dealing with natural systems. Unlike many engineered mechanisms, natural mechanisms generally do not reveal their parts. Sometimes, as in the case of organ systems, with appropriate interventions we can determine the functions those parts serve. Sometimes, as in the case of cells and organelles within them, where membranes serve to demarcate boundaries, we can, with appropriate techniques, observe the parts. But in other cases what appear to be distinct parts of a system are not the relevant parts for understanding its functioning. The gyri and sulci turned out not to constitute the operational parts of the brain (alternative decompositions based on cytoarchitectonic criteria have fared far better). Operations within a mechanism are even harder to differentiate.

The notion of localization has often been the focus of criticism. Critics of neuroimaging, for example, often criticize it for assuming that mental activities are localized in discrete parts of the brain (for other recent critics, see Coltheart, 2004; Uttal, 2001; van Orden & Paap, 1997). We are not at all convinced that this is the best way to interpret most recent work on neuroimaging (see, for example, Logothetis, 2008); but in any case, the conception of localization we invoke is rather different than that assumed in such critiques. First, although sometimes research begins by assuming that a whole activity of a mechanism is due to a component within it-what we called direct or simple localization-this is often but a step in a research endeavor. Once a component is identified and its behavior explored, it turns out not to generate the phenomenon on its own, but to perform an operation that, together with a constellation of other operations performed by other components, generates the phenomenon. As research proceeds, it is not the whole phenomenon that is localized in a part of the system, but individual operations each of which contributes in some way or other to the phenomenon of interest. Second, although the word localization suggests a single discrete spatial location, that is not necessary and is often not correct. The functional component may be distributed in space, with the intervening space containing parts performing other operations or even entities that are not part of the mechanism responsible for the given phenomenon.

A further criticism of localization is that identifying the part where an operation is performed is of no intrinsic interest. This points to the third, and probably most important, difference in our construal of localization: we treat localization not an end of inquiry but a heuristic. The goal of localizing is not to find out where something occurs, but to acquire information about the part that is engaged in that activity that can inform further investigation. For example, once the part is identified, we can investigate whether it is also involved in other phenomena. If it is, that suggests investigating whether common operations figure in those different phenomena.

When direct or simple localization proves inadequate for understanding a given phenomenon, and the phenomenon is decomposed into multiple operations localized in different parts, the issue of how these parts are organized becomes important. Typically researchers begin with simpler mechanisms and hypothesize that a mechanism consists of components whose operations are performed sequentially. The simpler hypothesis may not be the best, but it is after all the simplest. That makes it easier to assess. This is what we characterized as *indirect* or *complex* localization and it is envisaged in Machamer, Darden, and Craver's (2000) characterization of a mechanism as functioning from "from start or set-up to finish or termination conditions." This is also compatible with Simon's appeal to near decomposability. (At least, it is in the spirit of Simon's characterization. He says the long range behavior of such components depend only in an aggregative way on the behavior of others. That is not true here, but at least it would depend only *linearly* on the behavior of other components.) If one imagines, as some seem to have done, that language processing depends on two regions-one specialized for semantic processing and another for syntactic processing, or one for speech comprehension and one for speech production-then it is natural to assume that one provides the input for the other, and that each performs its function in relative independence of the other. We now know that such a simple model is not defensible; but it is at least a natural first approximation. Once again, in determining how it breaks down, we at the same time uncover crucial information about the mechanism. We now know that language processing involved substantial portions of both hemispheres, and though we do not yet have a detailed understanding of the contributions we have a much more developed understanding of the process (cf. Richardson, 2009). The attempt to understand mechanisms in terms of linear execution of operations can be extremely productive, both as a heuristic that simplifies the search space and generates an initial testable model and as a fallible assumption whose ultimate failing may reveal far more complex modes of organization.

We took a small step beyond decomposable and nearly decomposable modes of organization in characterizing *integrated* systems as those whose behavior results from interactions between component parts and operations within the system. One of the simplest, and most pervasive, forms of interaction is negative feedback, which has become widely familiar and a staple of control theory. Yet even it was difficult for humans to understand initially, so common is our attempt to understand mechanisms as involving a sequence of operations. The first recognized use of negative feedback by human designers was in the 3rd century BCE when Ktesibios employed a float as a feedback device to regulate water flow into a supply vessel to insure a constant volume and hence a constant rate of water entering into his water clock (Mayr, 1970). Subsequently negative feedback was reinvented numerous times as new technologies required insuring stable conditions. James Watts' centrifugal governor for the steam engine was a celebrated exemplar that inspired Maxwell's (1868) mathematical analysis of governors, but the idea of negative feedback as a means of enabling engineered devices to obtain goals was only recognized as a general design principle with the emergence of the cybernetics movement in the mid-20th century (Rosenblueth, Wiener, & Bigelow, 1943). The cyberneticists also recognized the pervasiveness of negative feedback in biological systems (Wiener, 1948), an idea that had been adumbrated by Claude Bernard (1865) in his conception of organisms maintaining the constancy of their internal environments, and made more concrete in the various examples Walter Cannon (1929) identified in articulating the notion of homeostasis.

Although biologists in the late 20th century identified many instances of negative feedback at different levels of biological organization, only a few attended to the dynamical behavior of systems with negative feedback, typically assuming that negative feedback achieves the goal of smoothly regulated behavior or that negative feedback maintains some equilibrium state. A few, however, recognized what engineers often encountered as a problem: if there is

delay or if there are non-linear interactions in feedback, the system can enter into oscillations. Theorizing about the lac-operon model of negative feedback control in gene expression, discovered and elaborated by Jacob and Monod in the 1950's (cf. Jacob & Monod, 1961), Brian Goodwin (1965) developed a mathematical model in which feedback generated oscillations as a result of non-linear interactions of molecules in the feedback process. At nearly the same time Amal Ghosh, employing new techniques for rapid measurement of NADH levels developed in Britton Chance's laboratory, determined that NADH levels in yeast preparations performing glycolysis oscillated with a period of two to three minutes (Ghosh & Chance, 1964). Although there has been considerable uncertainty as to whether glycolytic oscillations occur under physiological conditions (see Richard et al., 1994, for claims that they do), glycolytic oscillations became a prototype for studying and modeling biological oscillations, in part because the responsible mechanism was readily identified (Hess & Boiteux, 1968; see Goldbeter, 1995, for a detailed review). The foundational ideas of how feedback loops in thermodynamically driven processes generate oscillations have received extensive uptake in research on circadian rhythms, the daily oscillations that affect nearly all physiological and behavioral processes in living organisms (see Bechtel & Abrahamsen, in press-a, for discussion).

Even as negative feedback was coming to be recognized as widespread and its potential for effecting control and generating complex dynamics came to be recognized, positive feedback was rejected as unlikely to occur as it was thought to generate run-away systems. In the 1950s the Soviet chemist and biophysicist Boris Belousov attempted to develop an inorganic analogue of the citric acid cycle, the biochemical cycle discovered by Hans Krebs (Krebs & Johnson, 1937) through which the product of a series of oxidative reactions starting with citric acid, oxaloacetic acid, combines with a molecule of acetyl-CoA to generate a new molecule of citric acid. When Belousov combined citric acid, acidified bromate (BrO₃⁻), and ceric salt, the resulting solution oscillated between yellow and clear. His results were rejected by mainstream chemists who, influenced by equilibrium thermodynamics, assumed all chemical reactions would proceed to a stable state, then stop, and he was only able to publish them in a conference abstract (Belousov, 1959). Anatol Zhabotinsky (1964) alone picked up on the results, and replacing citric acid with malonic acid, developed a preparation in which, when a thin layer of reactants is left undisturbed, produces varying geometric patterns such as concentric circles and spirals that propagate across the medium. Only when Field, Koros, and Noyes (1972) developed a detailed model of the kinetics of such reactions, showing them to involve coupled autocatalytic reactions in a system open to energy, did they become exemplars for the complex dynamics that can arise in naturally occurring open systems.

What is important about mechanisms with positive and negative feedback for our purposes is the tension it places on the assumption of decomposability or near decomposability. The more the various operations within the mechanism affect each other, the less successful is a sequential account of the mechanism that treats each operation as independent of the others. At an extreme there will be systems, or at least models of systems, in which the components are all uniform (e.g., nodes in a connectionist network or Boolean switches in Kauffman's gene regulatory networks) and the explanation of the resulting behavior appeals totally to the organization that is realized in the system. At such an extreme, as we discuss in chapter 9, the mechanistic heuristics of decomposition and localization cease to be productive, except insofar as their failure discloses functional integration. We suggested that such behavior could be seen as

'emergent' at least insofar as the organization of the system, rather than the contributions of its constituent components, determines systemic function (cf. Bechtel and Richardson 1992, and chapter 9).

Since 1993 there has been extensive analysis of network architectures which suggest that complex dynamical systems often end up considerably short of the limit at which decomposition and localization fail. Classically, graph theorists in mathematics focused on two ways of organizing nodes into a network: as regular lattices in which each unit is connected to its neighbors and as random networks. In a regular lattice there is interaction between neighbors and there is a path, albeit often a long one, whereby each node can communicate with others. In random networks, when the number of connections exceeds half the number of nodes, a giant component whose nodes are all connected by a short path is likely to emerge (Erdös & Rényi, 1960). But, beginning with the work on Duncan Watts on communication between fireflies, theorists have focused on a different mode of organization known as small-world networks in which nearby units are highly interconnected but there are a few long-range connections between units. Even with only a very few long range connections, these networks exhibit the short path length associated with random networks as well as the high clustering between local components found in regular lattices (Strogatz, 2001).³ Such organization is found to be widespread in biological systems, especially complex dynamical systems in which coordinated behavior of independent oscillatory components is required, either to synchronize behavior or to create waves.

Independently, examination of network structures by Barabási showed that another common assumption, that the number of connections originating from a node in a network is randomly distributed, is often false. Rather, in many cases the distribution more closely resembles a power-law, although often one with a cut off tail (Albert & Barabási, 2002). The existence of a few highly connected components enhances communication through the network and also accounts for the robustness of such networks to the destruction of even a large number of individual connections. These ideas are receiving increased attention in neuroscience where they are seen as supporting complex integrated dynamical activity and local components that perform distinct processing tasks (Sporns, Heidi, & Timothy, 2009; Rubinov, Sporns, van Leeuwen, & Breakspear, 2009).

These recent developments, and others we cannot include here, support two critical points about the investigation and understanding of mechanisms. The first concerns the manner in which mechanisms are studied. Traditionally, mechanistic explanations could be presented with little reliance on mathematics. But when the components of a mechanism are highly integrated, so that the behavior of given components can be affected by activity in many others, mechanistic explanations increasingly rely on mathematical modeling, often invoking sets of differential

³ Stuart Kauffman argued in the *Origins of Order* that much natural order is what he calls "spontaneous order" in complex systems (cf., 1993, p. 173). As we've noted, Kaufman models complex systems there as Boolean networks, and explores the effects of varying connectivity patterns. With large numbers of nodes, connectivity values around 2 (so that each node outputs to 2 others), yield systems that are neither chaotic nor frozen. There are essentially small world networks. See Richardson (2001) and Burian & Richardson (1991) for discussion. One fascinating aspect of Kauffman's work, and of early connectionist modeling, is the range of dynamical behaviors that can be realized with components that are modest in their functional properties. This is why we put the emphasis on organization.

equations. These often cannot be solved analytically, but only in simulation. This means, among other things, that models must be quantitative and not only qualitative. Moreover, the visual representation techniques and the analytical concepts and tools that have been developed in dynamical systems theory (e.g., concepts such as limit cycle, chaotic attractor, bifurcations), are increasingly needed to characterize the behavior of mechanisms. Although for some theorists such invocation of dynamics entails a repudiation of mechanism, others recognize the need for integrating mechanism and dynamics in what Bechtel and Abrahamsen (in press-b) have characterized as dynamic mechanistic explanation. Similarly, Richardson and Stephan (2007b) as well as Boogerd, Bruggeman and Richardson (in preparation) argue that the sorts of explanations offered within systems biology are dynamic and should be understood as mechanistic.

The second point concerns a different sense of "emergence" from the one we noted above. There is an important tension between traditional strategies of decomposition and localization and recognition of dynamical relations in mechanisms. Approaches to decomposition and localization often assume that when a component is removed from the system—e.g., an enzyme is extracted from the cell milieu, and studied in vitro—the properties crucial to its behavior will not be affected. But insofar as interactions within the system alter its behavior, that assumption will prove false. Component behavior will be system sensitive. In fact, this is often the case. This gives rise to an important-and entirely naturalistic-sense of emergence which does not require going to the limit considered in chapter 9 and that has been explored by Boogerd, Bruggeman, Richardson, Stephan, and Westerhoff (2005). Assume we have a semi-open system with a definite organization, and well-defined components. To say it is semi-open means there is some exchange of mass or energy with the environment, though that flow is restricted by a boundary (perhaps a cell membrane) separating the system from the environment. The semi-permeable boundary will also dissipate free energy. With constant sources of energy to the system, there will be characteristic behaviors it exhibits. When there is sequential organization with linear interactions, the system will tend toward equilibrium states. When there is non-sequential organization with nonlinear organization, systems exhibit much more complex dynamics. This much characterizes a wide array of systems, from organisms to ecosystems or economies. These are nearly the thermodynamically open systems of Nicolis and Prigogine (1977). It may help to be a bit more concrete. The components of interest in cell physiology—cells are among these thermodynamically semi-open systems studied within systems biology⁴—characteristically include low molecular weight molecules (e.g., metabolites such as lactate and pyruvate), and macro-molecules (including enzymes, proteins, DNA), sometimes isolated into subsystems (e.g., in organelles). Proteins interact either directly or indirectly through the action of intermediaries. The interactions among the molecules, which are typically rather specific, define the organization of the network. These networks of interacting molecules, which are organized in such as manner as maintain or propagate themselves (Collier & Hooker, 1999; Ruiz-Mirazo, Peretó, & Moreno, 2004), are cells.

Even such simple systems do exhibit a form of emergence. Emergence is, as we know, an ambiguous notion (cf. Stephan, 1999). The emergence exhibited in systems biology is neither

⁴ "Systems biology" is a heterogeneous discipline, united by the focus on systems and the methods for exploring them. It encompasses some work in ecological modeling, organismic modeling, proteomics, genomic, and more. In all these cases, it is the focus on whole system modeling, which is dynamic, with semi-open systems that is common to the range. Systems biologists also typically deploy computational models, and are data driven.

weak nor epistemic. That is, though there are systemic behaviors exhibited which are unique, and differ from anything exhibited by components, that is not what constitutes this form of emergence; and the emergence is not merely a matter of our epistemic limitations. Let's see how this works within systems biology, and specifically within cell physiology. After all, if there is emergence within cell biology, and in particular in the behavior of 'simple' one celled organisms such as *E coli*, we should certainly expect there is emergence quite broadly. (Similar issues arise much more generally, including such places as neural networks, or in the modeling of ecosystem dynamics; but the point here is not to argue for the generality of the case.) Within cellular systems biology, what is called *kinetic modeling* begins with the kinetic parameters that characterize the interactions among metabolites and enzymes, and 'predicts' the dynamical behavior of the system from these parameters. Given that we know the constituents of the system, their organization (i.e., the network of interactions), together with the internal and external conditions, that is sufficient to determine not only the dynamic behavior of the system, but the range of dynamical behaviors within the repertoire of the system. We can predict the behavior of systems given these parameters (cf. Richardson & Stephan, 2007a). These are fully mechanistic models which are dynamic in the sense that Bechtel and Abrahamsen (in-press-a, inpress-b) articulate. Under some conditions these systems also exhibit emergence. Understanding how this is so is usefully approached in two stages. The first defines a class of dynamic systems that are both natural and mechanistic, but which do not fit comfortably within the category of decomposable or nearly decomposable systems. The second carves out a class of systems whose behavior is usefully treated as emergent from within the broader class of dynamic but nondecomposable systems.

So we begin with two preliminary conditions to be met under if these systems are to exhibit naturalistic emergence, though these are not sufficient for emergence since many systems meet these conditions but do not exhibit emergence. First, in such a natural system, it must be true that the component behaviors are systematically and nonlinearly affected by the contributions of other components. This can be the case when there is negative and/or positive feedback within the system; in the absence of such systemic interaction, there may be influences, but none that affect sequential operation or the linearity of interactions, and therefore none inhibit decomposition and localization. In a fairly straightforward case, the flux through a biochemical network sometimes depends nonlinearly on the concentrations of its enzymes; and the actual contributions any one enzyme makes depend quantitatively on the concentrations of other enzymes. Second, the characterization of the state of the system must be framed exclusively in terms of the state-independent properties. So, among other things, in order to predict the behavior of, say, a bacterium, we need to know the concentrations of nutrients, enzymes, and metabolites, as well as external conditions such as temperature and pressure that affect the organism. The component properties of parts are dependent on this total stateindependent characterization of the system. As we've said, this total characterization is sufficient to determine the state and the dynamics of the system as a whole.

Typically, in such systems, there is distributed control. That is, there is no single reaction which determines the state of the system, or the dynamics of the system; and changing, say, the concentration of an enzyme can have a wide array of effects, including some not under its immediate control. This raises a number of engaging problems in determining the component properties and capacities. There are sometimes technical barriers to accomplishing such a

complete characterization, of course. However, it is sometimes possible experimentally to determine the properties of components, including their functional properties of components within the system of interest. In some cases, such things as the kinetic parameters can be determined *in vivo*, using techniques such as knock-out genes. In other cases, by carefully matching experimental conditions to those experienced *in vivo*, they can be determined *in vitro*. The resulting models are enormously complex, with many components and elaborate networks of interactions. They are sufficiently complex even with real systems such as *E coli* that there is no simple solution. In general, simulations are necessary to yield any predictions concerning cellular dynamics. (See Ideker et al., 2001, for computational simulations of E. coli.; van Beek, 2006)

Though such systems are fully naturalistic, there is still space for emergence in a sense that is not merely epistemic. Again there are two conditions. If these are met as well as the first two, there will be significant naturalistic emergence. First, we have already noticed that in nonsequential systems, what we called a 'common assumption' of decomposition will fail. The behavior of components within the system will depend on the actual behavior and the capacities of other components within the system; and to the extent that feedback that is system wide, these dependencies will make the component behaviors-the realized properties of componentsspecific to the system. As a consequence, outside the systemic context, component behavior will be different than it is within the systemic context. As Boogerd et al say, the behavior of components within the system will not be 'predictable' from the behavior in simpler systems. (On the relevant sense of 'prediction,' see Richardson and Stephan 2007a. The point is essentially that the behavior of the system can be described, but not predicted *ab initio* from first principles.) Second, the nonlinearities affecting component behaviors must in turn affect systemic behaviors. For example, such organization can introduce oscillatory states, or instabilities that would not be present in simpler systems or linear systems. The two conditions insure that the behavior of components is system dependent, and that this affects system behavior as a whole. When these two conditions are met, the systemic behavior is reasonably counted as emergent, even though it is fully explicable mechanistically.

In the above paragraphs we have identified two conceptions of emergence. One represents a limit condition in which mechanistic explanation ceases to be relevant since there are no component parts and operations to be distinguished. The organization of the system is then the focus of explanations of systemic behavior. The second embraces mechanistic explanation and recognizes the importance of characterizing the component parts and operation of the relevant mechanism, but treats the phenomenon produced by the mechanism as emergent when the operation of the components cannot be predicted from their behavior in simpler systems and the nonlinearities in the behavior of the parts affects the behavior of the whole system. Our broader objective in this section has been to point to the range of organization beyond simple and complex localization on which we focused in the book and to emphasize the importance of considering such modes of organization in accounts of mechanism that will be adequate to developments within mechanistic science such as are currently being explored in systems biology and related fields.

4. Recomposing and Situating Mechanisms

As we noted above, our project began as an attempt to understand reductionistic research in science but soon found that neither the account of reduction prevalent in philosophy (the theory reduction account presented in Nagel, 1961) nor the modifications philosophers had advanced (Kenneth Schaffner, 1967; Nickles, 1973; Hooker, 1981; Bickle, 1998) to be helpful in understanding reductionistic research in science. We ended up recharacterizing the problems we faced in terms of the pursuit of mechanisms that involve decomposing systems into their component operations and dropping the use of the word *reduction* altogether. Since then some philosophical treatments of reduction have abandoned the theory reduction framework and characterized reduction in terms that are apparently compatible with our account of decomposition and localization. Bickle's (2003) ruthless reductionism, for example, advocates explaining cognitive processes in terms of the cellular and molecular processes in the brain. Bickle here abandons the emphasis on theory reduction that characterized his earlier work, in favor of molecular mechanisms for behavior. But Bickle's account retains an important aspect of the theory reduction account-the view that the lower level processes are what provide genuine explanation. For him, higher level accounts are "merely heuristic." It also retains a place for laws or generalizations, neither of which plays a significant role in our accounts.

Implicit in our discussion of organization, complexity, and emergence in the previous section is the basis for rejecting ruthless reductionism and for distinguishing mechanistic reduction from ruthless reduction or any account that construes lower levels as the source of all explanation (e.g., Kim, 1998); we also reject accounts that place laws center stage in the explanation of complex systems, though we acknowledge that some laws feature in understanding the mechanisms. In a mechanism, parts and operations are organized, and in many cases the operations are orchestrated in time, so that the mechanism as a whole behaves in a specific way. It is not the operations of the parts alone, or even as simply added together, that explain what the mechanism does, but the specifics of their organization and their orchestrated execution. Moreover, the parts and operations are often modified by activities occurring elsewhere in the mechanism, and, in some cases, by activities external to the mechanism. Scientists must attend to the whole mechanism, in its characteristic environment, and not just its lower-level constituents, to understand what the mechanism does, or even how the components function.

This may be clearer if we relax our focus on the parts and operations, and look toward the system. The counterpart to decomposing a mechanism is recomposing it. We earlier mentioned Lavoisier's structural decomposition of water into constituent molecules. As Lavoisier's compounds were decomposable into their constituents, it was equally important to him that he could construct or recompose the compounds from the simple constituents. The simplest experiments of this sort involved burning metals, but one of Lavoisier's most famous experiments involved the combustion of hydrogen, yielding water. This is structural recomposition. As before, our focus is on functional matters, but the principle is the same. One must show that the postulated components, with an appropriate organization, are sufficient to yield the systemic behavior (of course, within the right boundary conditions). Recomposing a mechanism is often as difficult if not more difficult than decomposing it, but both are fraught with dangers. There are occasions when scientists literally try to put mechanisms back together. For example, after decomposing the parts of a biochemical system by centrifugation and studying what operations occur in the separated fractions, biochemists often try to reconstitute

the original system to show that indeed they had identified all the necessary components of the mechanism (Racker, 1976). Not infrequently, these efforts result in failure—the reconstituted system does not generate the original phenomenon, or only some aspects of it. One possibility is that the researchers have failed to identify some critical component. Another is that the organization was not adequately recovered in the reconstituted system and it contributed in essential ways to the production of the phenomenon. In either case, the failure to reconstitute the system behavior signals a failure to identify key constituents of the mechanism.

In many other cases, actual reconstitution is not a possibility and investigators appeal to mathematical or computational models to determine how a system of components that they have identified would behave when put together. The equations employed in elaborating models often involve a number of crucial parameters, and to ensure that the models are realistic, researchers must ascertain what are plausible values of these parameters. At the time we wrote this book, the sort of data that is needed was simply not available in a number of scientific fields. We are now beginning to acquire the kind of data rich domains in genetics and biochemistry, together with the computational power to manage it, so that it looks possible to construct such computational models. Especially when the organization is non-sequential and the interactions between components are nonlinear, models end up behaving in surprising ways. In some cases, these surprises turn out to reveal features of the phenomena that had not been noted previously. In other cases, however, they turn out to involve discrepancies between the model and the actual mechanism, and considerable research is required to determine where the model departs from the actual mechanism.

The first point to emphasize against the ruthless reductionist is that it is the whole organized mechanism that exhibits the phenomenon, not its individual parts. The glycolytic system as a whole, for example, converts glucose to ethanol and captures energy in ATP. Moreover, this phenomenon is characterized in terms of how the mechanism relates to its environment-concentrations of glucose are present in the environment of a yeast cell and it excretes quantities of alcohol and captures energy for itself. Similarly, in the case of the lac operon, the synthesis of β -galactosidase, which allows the use of complex sugars, depends on the prevalence of lactose in the extracellular environment. It also depends on the ability of the organism to import lactose into the cell, which depends on a permease that is also regulated by the lac operon. In both these cases, it is a mistake to focus on one component as opposed to an orchestrated set of responses. To use another example we discuss, whole organisms listen to or read words and speak or write them. These are people. These regularities involve the whole mechanism as it is situated in its environment and a crucial part of the challenge in developing an adequate mechanistic explanation is to identify the ways the mechanism engages its environment. In this case, the problem is understanding the development of the child in a linguistic environment. The point is that it is the child, in a family, and part of a culture, that learns language. Broca's area is a piece, but only a piece.

The second point is that mechanical models are not typically governed by a single level. Even in the simple cases portrayed here where we differentiate only the whole mechanism and the operations performed by its parts, the mechanistic account spans at least two levels. At each level causal processes are occurring and neither is sufficient to provide a complete account of the phenomenon in question. Researchers require appropriate tools both for characterizing the ways in which the mechanism interacts with its environment, and the ways its components interact with each other. In addition, they must figure out how, as a result of the parts performing the operations they do and interacting in an organized and orchestrated way, the organism exhibiting the phenomenon when impinged in particular ways by its environment. The challenge of mechanistic explanation is far greater than that of identifying the parts and operations at the lower level.

These lessons take on even greater significance when we recognize that often mechanistic explanations must span more than two levels of organization-the parts of a mechanism are themselves mechanisms and for some purposes it is necessary to decompose the parts to understand the original phenomenon. So the third point is that there are significant top-down constraints on the development of mechanistic models. For example, when the hippocampus was identified as the locus of spatial memory, a natural question was how the hippocampus performs this operation. Researchers were in part able to appeal to the distinctive neuroarchitecture of the hippocampus, with distinct fiber bundles linking regions each with specific neuronal composition. The discovery of place cells in the CA1 and CA3 regions suggested how these regions might contain maps of local environments. To understand how individual cells come to represent places, researchers needed to examine the internal processes within neurons that result in altered generation of action potentials. But to understand how the maps as a whole are organized, it is necessary to examine the interaction of cells. And to understand how the organism uses these maps to navigate requires considering the organization of the hippocampus as a whole and its relation to other neural structures. It is also very relevant that rats are oriented toward their environment spatially. Processes at each of these levels of organization, and the particular history of the organism in its environment, are all part of the mechanistic explanation. No level on its own gives a complete account and research at any given level does not supplant the need for research at other levels. Moreover, the characterization of higher level mechanisms places serious constraints on what constituent mechanisms must be able to do. It will not be enough, as ruthless reductionism would have it, to focus on the lowest level and let the phenomena fend for themselves.

Given the widespread use in mechanistic science of the term *reduction* for the decomposition of mechanisms into their parts and operation, we no longer need to avoid the term. Some mechanistic scientists have assumed that once the parts and operations were identified, the hard work was nearly done. But even they recognized that what they had discovered would only characterize the parts, not their organization or the environmental processes impinging on the mechanism. And, as scientists in many domains have come to recognize, as challenging as decomposition and localization are, recomposing and situating mechanisms is equally challenging. The calls for systems biology, even by those who have been highly accomplished in mechanistic reduction, is evidence of the magnitude of these challenge (Noble, 2006).

5. Model Systems, Conserved Mechanisms, and Generalization

A topic that was just beginning to receive much attention in the history and philosophy of science when we wrote the book was the role of model organisms and model systems in biological research. Biologists typically conduct their research on specific species that are

conducive to their research but intend their results to have much larger applicability. Disproportionate amounts of research are conducted on the bacterium Escherichia coli, bakers' yeast Saccharomyces cerevisiae, the nematode worm Caenorhabditis elegans, the fungus Neurospora crassa, the mustard plant Arabidopsis thaliana, the fruit fly Drosophila melanogaster, the bacteriophage lambda, and the mouse Mus musculus. Typically these species have been selected for research due to the fact that they leant themselves to experimental research because they were perceived to be relatively simple, were easy to handle and maintain, had short generation times, were readily available, etc. Over time large corpora of results from research on model organisms have been developed, enabling new research to draw on a significant body of relevant information. While researchers hoped to learn general lessons from work on these species, they were not viewed as typical. Model organisms are not even, generally, typical of their species. Research is often conducted on organisms that have been specially bred so as to minimize genetic variability, thereby increasing standardization of research results. So, for example, the laboratory rat—surely a prototypical model organism for behavioral biology—is an inbred and unusual rat by comparison with wild rats. Likewise, the increased use of genetic knock-outs as model organism certainly represents a shift to model organisms that are atypical.

The research we reported here was mostly obtained from model organisms—rats in the case of memory research, *Neurospora* in the research on gene regulation, and yeast in the case of glycolysis. The goal of the research, though, was not just to learn about physiological and cognitive processes in these species, but to procure a broader understanding. In particular, the goal is often to understand processes in humans. This raises the question of how researchers generalize from research on model organisms to other target organisms. If indeed explanation in biology involved derivations from general laws, the answer would be relatively straightforward—one discovered general laws in specific model systems and simply applied them (perhaps using different initial conditions) to the target systems of interest. But mechanistic research does not give rise to general laws. Rather, the result is typically a detailed account of the parts and operations of a mechanism and how they are organized and orchestrated in a specific model system.

Target systems are assumed to be similar, but commonly vary in specific, but unpredictable ways. One cannot simply assume that exactly the same mechanism responsible for a similar phenomenon in a different species (or even in different variants of the same species, or in different organs of the same organism) will contain identical parts and operations or that these will be organized in the same way. They may turn out to be quite similar, so that what was learned in the model species can be applied with relative ease to the target case, but typically some differences must be taken into account. So, for example, it is important, and significant, that so-called 'higher' organisms share a dipoloid mechanism of sex determination. It is equally important that only some share the xx/xy mechanism of mammals; it is equally important that some are xx/xo as are *Drosophila* and that some are haplodiploid, as are the social insects. There are many mechanisms of sex determination, as there are many ways for eyes to develop. The point is that there may be similar mechanisms, and similar phenomena. The interesting task of identifying a model organism concerns both its potential for generalization, and its experimental utility. The actual history of science once again promises to yield dividends to the philosophical problems in this area. The work on *Neurospora* in the explorations of biochemical genetics that was our focus in chapter 8 depended on assuming that *Neurospora* manifested something like sexual exchange. This was not obvious, to say the least, since *Neurospora* is not diploid. As we note, this was important for the research, since it meant there would be no genetic masking. This made it a problematic case for genetic research, of course. Just as the early decades the the 20th century genetics involves the development of *Drosophila* as a model organism, the history of genetics in the 1920's and 1930's is in part a history of the development of bacteria as model organisms for genetics, including the mechanics of bacterial 'sex.' Model organisms must after all be *model* organisms, revealing something more general about mechanisms of interest. This was not part of the story we told, but it is part of the story to be told (Richardson, 2008a, 2008b).

One way to characterize the relation between model and target systems is to assume that they are similar. Appeals to similarity, as Nelson Goodman noted, are often vacuous. But in biology they are often constrained by appeals to evolution as a process that results in descent with modification, giving rise to what are often characterized as conserved mechanisms. Sometimes conserved mechanisms are construed as frozen mechanisms, mechanisms in which no changes have occurred. This is seldom the case. Change is ubiquitous. Nonetheless, change is also not wholesale, enabling researchers to extrapolate from a mechanism in one organism to those in its descendants. In some cases the conservation is remarkable. The structure of the genetic regulatory genes that determine segment identity in *Drosophila* occur also in humans; and the Pax-6 genes that regulate eye development in the octopus are clearly homologues of the Pax-6 that regulate eye development in humans and in Drosophila. (They are both structurally similar and functionally intersubstitutable. So they are similar in sequence, and when transplanted they both induce eye development.) In generalizing explanations biologists often appeal to the claim that the mechanism of interest is conserved. This involves establishing the phylogenetic descent of the target mechanism from the one studied by identifying homologous parts and operations, often via identifying homologous genes.

Focusing on phylogenetic descent is a useful heuristic for developing accounts of mechanisms both when the target turns out to employ essentially the same mechanism and when it manifests significant changes. When variants are uncovered, they pose new research questions. Between *Drosophila* and mouse, for example, the homolog of one protein, cryptochrome, which serves to entrain the central oscillator to light in *Drosophila*, appears to have become a component of the central oscillator. That discovery raised questions both about how the protein was enabled to take on this new role, and what has taken over its old role since entraining the oscillator remains an important operation (Bechtel, in press).

Appeals to conservation in the study of biological mechanisms predate the attempts to map the genomes of humans and other species, but these have significantly affected how biologists appeal to conservation. Once the gene coding for an enzyme involved in a particular mechanism is identified, researchers often search gene databanks to find homologues in other species. Sometimes the homolog will be involved in the same or closely related mechanisms in the other species, but it may turn out to be involved in quite different processes. Both results can be highly informative. What has been learned about the same or similar mechanisms in the other species provides a valuable guide to investigating the mechanism in the current species. When a gene/protein is found to be involved in different mechanisms in different species, this too is informative in providing understanding into the operations the part can perform and how differences in organization and orchestration of operations can generate different results from the same components.

Generalization of mechanistic explanations beyond model systems via appeals to conserved mechanisms has yet to receive detailed analysis in philosophy of science (see, though, Kane, Richardson, & Fong, 1990). It represents a distinctly different mode of generalization, though, than is achieved by universally quantifying statements of laws and is one that is very important in biological practice. As we've suggested throughout, laws don't play a significant role. Moreover, generalization through conservation gives rise to a rather different conception of the unity of science than that fostered by the nomological approach. Unity is achieved by identifying common features among the diverse mechanisms that have been produced in the biological world. Sometimes these threads are topical—similar mechanisms are used to extract energy in all living systems. But sometimes they cross topics as when homologues of the same protein are found to be involved in mechanisms performing different functions. So, e.g., PAX-6 is involved not only in the induction of eye formation, but in pancreatic functions. The integration that is achieved is not deductive, but takes the form of a network of related mechanisms. Moreover, identifying these threads is often of great heuristic advantage to scientists.

Conclusion

The analysis we offer in *Discovering Complexity* is very much driven by the cases in the history of biology we chose to explore. These were, at the time, rather unconventional cases for philosophers to consider. They are also doubtless eccentric. Because they did not easily fit into extant frameworks in philosophy of science, we were forced to develop alternative analyses. Since its publication, other philosophers, exploring other cases, have developed similar accounts of mechanism. The emphasis on decomposition and localization as heuristics invoked in the discovery of mechanisms has, however, had less uptake. There are also important themes that were implicit in our account, but we had not made explicit, including the specific research techniques and strategies employed in decomposing and localizing operations, the importance of recomposing mechanisms, the multilevel nature of mechanistic explanations, and the role of evolutionary conservation in generalizing mechanistic accounts. Surely there are others that we have not yet noted, some of which will be revealed either by more detailed analyses of the cases we have considered, or the exploration of different cases in the life sciences. We do not expect that the same account will fit all instances in science. Rather, as scientists extend their analyses from model systems to others, fully expecting to find novelty as well as similarity, we expect that exploration of other parts of biology and other scientific fields will support similar analyses but also prompt new insights into the operation of science. We have at least indicated some directions we find enticing.

Perhaps the theme that excites us most in extending beyond the account we offered in 1993 is an appreciation of the complexity that lies beyond what we described. Mechanisms that depart significantly from sequential organization and involve nonlinear interactions between constituents are increasingly recognized to be common in the biological world. New tools of mathematical and computational modeling, of graph theoretical analysis, and of dynamical systems theory are being applied successfully to understanding the behavior of such mechanisms. These tools do not supplant the results of traditional mechanistic science, guided by the heuristics of decomposition and localization, but build upon it. Assuming the parts of a mechanism have a determinate character and perform operations that can be characterized independently of other operations occurring in the mechanism offers only a first approximation. We now see it is a crude approximation. As tools for studying the dynamics of such mechanisms are added to the traditional techniques of mechanistic science, researchers can move to better approximations of the actual mechanism. Likewise, as we incorporate these strategies scientists are cultivating into the account of mechanistic science, we can offer a more adequate account of the science itself. Our heuristic strategies of decomposition and localization were, we thought, the strategies of a reductionistic assault. We look forward to understanding the strategies of a reconstructionist science. That is what we've tried to describe here.

References

- Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74 (1), 47-97.
- Bechtel, W. (2000). From imaging to believing: Epistemic issues in generating biological data. In R. Creath & J. Maienschein (Eds.), *Biology and epistemology* (pp. 138-163). Cambridge, England: Cambridge University Press.
- Bechtel, W. (2006). *Discovering cell mechanisms: The creation of modern cell biology*. Cambridge: Cambridge University Press.
- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. London: Routledge.
- Bechtel, W. (in press). Generalization and discovery through conserved mechanisms: Cross species research on circadian oscillators. *Philosophy of Science*.
- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History* and *Philosophy of Biological and Biomedical Sciences*, *36*, 421-441.
- Bechtel, W., & Abrahamsen, A. (in press-a). Complex biological mechanisms: Cyclic, oscillatory, and autonomous. In C. A. Hooker (Ed.), *Philosophy of complex systems. Handbook of the philosophy of science, Volume 10.* New York: Elsevier.
- Bechtel, W., & Abrahamsen, A. (in press-b). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*.
- Belousov, B. P. (1959). A periodic reaction and its mechanism. *Compilation of Abstracts on Radiation Medicine*, 147, 145.
- Bernard, C. (1865). An introduction to the study of experimental medicine. New York: Dover.
- Bickle, J. (1998). Psychoneural reduction: The new wave. Cambridge, MA: MIT Press.
- Bickle, J. (2003). *Philosophy and neuroscience: A ruthlessly reductive account*. Dordrecht: Kluwer.
- Bishop, M., & Trout, J. D. (2005). *Epistemology and the psychology of human judgment*. New York: Oxford University Press.
- Boogerd, F. C., Bruggeman, F. J., Richardson, R. C., Stephan, A., & Westerhoff, H. V. (2005). Emergence and its place in nature: A case study of biochemical networks. *Synthese*, 145, 131-164.

- Burian, R. M., & Richardson, R. C. (1991). Form and order in evolutionary biology. In A. Fine, M. Forbes & L. Wessels (Eds.), *PSA 1990* (Vol. 2, pp. 267-287). East Lansing, MI: Philosophy of Science Association.
- Cannon, W. B. (1929). Organization of physiological homeostasis. *Physiological Reviews*, 9, 399-431.
- Carnap, R. (1928). Der logische Aufbau der Welt. Berlin: Weltkreis.
- Churchland, P. M., & Churchland, P. S. (1990). Intertheoretic reduction: A neuroscientist's field guide. *Seminars in the Neurosciences*, *2*, 249-256.
- Collier, J. D., & Hooker, C. A. (1999). Complexly organised dynamical systems. *Open Systems and Information Dynamics*, 6, 241-302.
- Coltheart, M. (2004). Brain imaging, connectionism, and cognitive neuropsychology. *Cognitive Neuropsychology*, 21, 21-25.
- Craver, C. (2002). Interlevel experiments and multilevel mechanisms in the neuroscience of memory. *Philosophy of Science*, *69*, S83-S97.
- Craver, C. (2007). *Explaining the brain: What a science of the mind-brain could be*. New York: Oxford University Press.
- Craver, C. (2008). Physical law and mechanistic explanation in the Hodgkin and Huxley model of the action potential. *Philosophy of Science*, *75* (5), 1022-1033.
- Darden, L. (1991). *Theory change in science: Strategies from Mendelian genetics*. New York: Oxford University Press.
- Darden, L. (2006). *Reasoning in biological discoveries: Essays on mechanisms, interfield relations, and anomaly resolution.* Cambridge: Cambridge University Press.
- Erdös, P., & Rényi, A. (1960). On the evolution of Random Graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, *5*, 17-61.
- Feyerabend, P. K. (1975). Against method. London: New Left Books.
- Field, R. J., Koros, E., & Noyes, R. M. (1972). Oscillations in chemical systems. II. Thorough analysis of temporal oscillation in the Bromate–Cerium–Malonic acid system. *Journal of the American Chemical Society*, 94, 8649–8664.
- Flanagan, O. J. (1991). Varieties of moral personality : ethics and psychological realism. Cambridge, Mass. :: Harvard University Press.
- Ghosh, A., & Chance, B. (1964). Oscillations of glycolytic intermediates in yeast cells. Biochemical and Biophysical Research Communications, 16 (2), 174-181.
- Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Kruger, L. (1989). *The Empire* of chance: How probability changed science and everyday life. Cambridge: Cambridge University Press.
- Gigerenzer, G., Todd, P. M., & the ABC Research group. (1999). Simple heuristics that make us smart. New York :: Oxford University Press.
- Glennan, S. (1996). Mechanisms and the nature of causation. Erkenntnis, 44, 50-71.
- Glennan, S. (2002). Rethinking mechanistic explanation. Philosophy of Science, 69, S342-S353.
- Goldbeter, A. (1995). *Biochemical oscillations and cellular rhythms: The molecular bases of periodic and chaotic behaviour*. Cambridge: Cambridge University Press.
- Goodwin, B. C. (1965). Oscillatory behavior in enzymatic control processes. *Advances in Enzyme Regulation, 3*, 425-428.
- Hanson, N. R. (1958). Patterns of discovery. Cambridge: Cambridge University Press.

- Hempel, C. G. (1965). Aspects of scientific explanation. In C. G. Hempel (Ed.), Aspects of scientific explanation and other essays in the philosophy of science (pp. 331-496). New York: Macmillan.
- Hess, B., & Boiteux, A. (1968). Mechanism of glycolytic oscillation in yeast. *Hoppe-Seyler's Zeitschrift für physiologische Chemie*, 349, 1567-1574.
- Hooker, C. A. (1981). Towards a general theory of reduction. *Dialogue*, 20, 38-59; 201-236; 496-529.
- Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., et al. (2001). Integrated Genomic and Proteomic Analyses of a Systematically Perturbed Metabolic Network. *Science*, 292 (5518), 929-934.
- Jacob, F., & Monod, J. (1961). Genetic regulatory systems in the synthesis of proteins. *Journal* of Molecular Biology, 3, 318-356.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press.
- Kane, T. C., Richardson, R. C., & Fong, D. W. (1990). The phenotype as the level of selection: Cave organisms as model systems. In A. Fine, M. Forbes & L. Wessels (Eds.), *PSA 1990* (Vol. 1, pp. 151-164). East Lansing, MI: Philosophy of Science Association.
- Kauffman, S. (1971). Articulation of parts explanations in biology and the rational serach for them. In R. C. Bluck & R. S. Cohen (Eds.), *PSA 1970* (pp. 257-272). Dordrecht: Reidel.
- Kauffman, S. A. (1993). *The origins of order: Self-organization and selection in evolution*. Oxford: Oxford University Press.
- Kim, J. (1998). Mind in a physical world. Cambridge, MA: MIT Press.
- Krebs, H. A., & Johnson, W. A. (1937). The role of citric acid in intermediate metabolism in animal tissues. *Enzymologia*, *4*, 148-156.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (Second ed.). Chicago: University of Chicago Press.
- Langley, P., Simon, H. A., Bradshaw, G. L., & Zytkow, J. M. (1987). *Scientific discovery: Computational explorations of the creative process*. Cambridge: MIT Press.
- Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, 453 (7197), 869-878.
- Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67, 1-25.
- Maxwell, J. C. (1868). On governors. Proceedings of the Royal Society of London, 16, 270-283.
- Mayr, O. (1970). The origins of feedback control. Cambridge, MA: MIT Press.
- Nagel, E. (1961). The structure of science. New York: Harcourt, Brace.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Nickles, T. (1973). Two concepts of intertheoretic reduction. *The Journal of Philosophy.*, 70, 181-201.
- Nickles, T. (Ed.). (1980a). Scientific discovery: Case studies. Dordrecht: Reidel.
- Nickles, T. (Ed.). (1980b). *Scientific discovery: Logic and rationality*. Dordrecht: D. Reidel Publishing Company.
- Nicolis, G., & Prigogine, I. (1977). Self-organization in nonequilibrium systems: From dissipative structures to order through fluctuations. New York: Wiley.
- Noble, D. (2006). *The music of life: biology beyond the genome*. Oxford: Oxford University Press.

Oyama, S., Griffiths, P. E., & Gray, R. D. (Eds.). (2001). *Cycles of contingency : developmental systems and evolution*. Cambridge, MA: MIT Press.

- Popper, K. (1934). Logik der Forschung Vienna: Springer.
- Popper, K. (1959). The logic of discovery. London: Hutchinson.
- Racker, E. (1976). A new look at mechanisms in bioenergetics. New York: Academic Press.
- Ramsey, Jeffry L. (2008). Mechanisms and Their Explanatory Challenges in Organic Chemistry. *Philosophy of SCience*, *75* (5), 970-982.
- Richard, P., Diderich, J. A., Bakker, B. M., Teusink, B., van Dam, K., & Westerhoff, H. V. (1994). Yeast cells with a specific cellular make-up and an environment that removes acetaldehyde are prone to sustained glycolytic oscillations. *FEBS Letters*, 341 (2-3), 223-226.
- Richardson, R. C. (2001). Complexity, self-organization, and selection. *Biology and Philosophy*, *16*, 655-683.
- Richardson, R. C. (2008a). Boris Ephrussi. In N. Koertge (Ed.), *New Dictionary of Scientific Biography* (Vol. 2, pp. 398-403). New York: Charles Scriber's Sons.
- Richardson, R. C. (2008b). Edward Lawrie Tatum. In N. Koertge (Ed.), *New Dictionary of Scientific Biography* (Vol. 7, pp. 12-15). New York: Charles Scriber's Sons.
- Richardson, R. C. (2009). Multiple realization and methodological pluralism. *Synthese*, *167*, 473-492.
- Richardson, R. C., & Stephan, A. (2007a). Emergence. Biological Theory, 2, 91-96.
- Richardson, R. C., & Stephan, A. (2007b). Mechanism and mechanical explanation in systems biology. In F. C. Boogerd, F. J. Bruggeman, J.-H. Hofmeyr & H. V. Westerhoff (Eds.), *Systems Biology: Philosophical Perspectives* (pp. 123-144). New York: Elsevier.
- Rosenblueth, A., Wiener, N., & Bigelow, J. (1943). Behavior, purpose, and teleology. *Philosophy of Science*, *10*, 18-24.
- Rubinov, M., Sporns, O., van Leeuwen, C., & Breakspear, M. (2009). Symbiotic relationship between brain structure and dynamics. *BMC Neuroscience*, 10 (1), 55.
- Ruiz-Mirazo, K., Peretó, J., & Moreno, A. (2004). A universal definition of life: Autonomy and open-ended evolution. *Origins of Life and Evolution of the Biosphere*, *34*, 323-346.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal*, *3*, 210-229.
- Schaffner, K. (1967). Approaches to reduction. *Philosophy of Science*, 34, 137-147.
- Schaffner, K. (1993). *Discovery and explanation in biology and medicine*. Chicago: University of Chicago Press.
- Schaffner, Kenneth F. (2008). Theories, Models, and Equations in Biology: The Heuristic Search for Emergent Simplifications in Neurobiology. *Philosophy of Science*, 75 (5), 1008-1021.
- Simon, H. A. (1969). The sciences of the artificial. Cambridge, MA: MIT Press.
- Simon, H. A. (1979). *Models of thought*. New Haven: Yale University Press.
- Simon, H. A. (1983). *Reason in human affairs* (Vol. 1982). Stanford, Calif. :: Stanford University Press.
- Skipper, R. A., & Millstein, R. L. (2005). Thinking about evolutionary mechanisms: Natural selection. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 327-347.

- Sporns, O., Heidi, J.-B., & Timothy, E. J. B. (2009). The Human Connectome: Linking Structure and Function in the Human Brain. In *Diffusion MRI* (pp. 309-332). San Diego: Academic Press.
- Stephan, A. (1999). *Emergenz: Von der Unforhersagbarkeit zum Selbstorganisation*. Dresden: Dresden University Press.
- Stich, S. P. (1983). From folk psychology to cognitive science: The case against belief. Cambridge, MA: MIT Press.
- Stotz, K. C. (2008). The ingredients of a postgenomic synthesis of nature and nurture. *Philosophical Psychology*, *21*, 359-381.
- Strogatz, S. H. (2001). Exploring complex networks. Nature, 410 (6825), 268-276.
- Sturtevant, A. H. (1913). The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of association. *The Journal of Experimental Zoology*, *14*, 43-59.
- Sturtevant, A. H. (1965). A history of genetics. New York: Harper and Row.
- Tabery, J. (2004). Synthesizing activities and interactions in the concept of a mechanism. *Philosophy of Science*, *71*, 1-15.
- Thagard, P. (1988). *Computational philosophy of science*. Cambridge, MA: MIT Press/Bradford Books.
- Thagard, P. (1992). Conceptual revolutions. Princeton, NJ: Princeton University Press.
- Thagard, P. (2003). Pathways to biomedical discovery. Philosophy of Science, 70, 235-254.
- Thagard, P. (2006). *Hot thought: Mechanisms and applications of emotional cognition*. Cambridge, MA: MIT Press.
- Uttal, W. R. (2001). *The new phrenology: The limits of localizing cognitive processes in the brain*. Cambridge, MA: MIT Press.
- van Beek, J. H. G. M. (2006). Channeling the Data Flood: Handling Large-Scale Biomolecular Measurements in Silico. *Proceedings of the IEEE*, 94 (4), 692-709.
- van Orden, G. C., & Paap, K. R. (1997). Functional neural images fail to discover the pieces of the mind in the parts of the brain. *Philosophy of Science*, 64 (4), S85-S94.
- Waskan, J. (2006). Models and cognition. Cambridge, MA: MIT Press.
- Weber, M. (2005). Philosophy of experimental biology. Cambridge: Cambridge University Press.
- Weber, M. (2008). Causes without Mechanisms: Experimental Regularities, Physical Laws, and Neuroscientific Explanation. *Philosophy of Science*, *75* (5), 995-1007.
- Wiener, N. (1948). *Cybernetics: Or, control and communication in the animal and the machine.* New York: Wiley.
- Wimsatt, W. C. (1974). Complexity and organization. In K. F. Schaffner & R. S. Cohen (Eds.), *PSA 1972* (pp. 67-86). Dordrecht: Reidel.
- Wimsatt, W. C. (1976). Reductionism, levels of organization, and the mind-body problem. In G. Globus, G. Maxwell & I. Savodnik (Eds.), *Consciousness and the brain: A scientific and philosophical inquiry* (pp. 202-267). New York: Plenum Press.
- Wimsatt, W. C. (1987). False models as means to truer theories. In M. H. Nitecki & A. Hoffman (Eds.), *Neural Models in Biology* (pp. 23-55). New York: Oxford Univesity Press.
- Wimsatt, W. C. (2007). *Re-engineering philosophy for limited beings: Piecewise approximations to reality*. Cambridge, MA: Harvard University Press.
- Woodward, J. (2000). Explanation and invariance in the special sciences. *British Journal for the Philosophy of Science*, *51*, 197-254.
- Zhabotinsky, A. M. (1964). Periodic processes of malonic acid oxidation in a liquid phase. *Biofizika*, 9, 306-311.