Reducing Psychology while Maintaining its Autonomy via Mechanistic Explanations[1]

William Bechtel
Department of Philosophy and Science Studies Program
University of California, San Diego

Abstract

Arguments for the autonomy of psychology or other higher-level sciences have often taken the form of denying the possibility of reduction. The form of reduction most proponents and critics of the autonomy of psychology have in mind is theory reduction. Mechanistic explanations provide a different perspective. Mechanistic explanations are reductionist insofar as they appeal to lower-level entities—the component parts of a mechanism and their operations—to explain a phenomenon. However, unlike theory reductions, mechanistic explanations also recognize the fundamental role of organization in enabling mechanisms to engage their environments as units (as well as the role of yet higher-level structures in constraining such engagement). Especially when organization is non-linear, it can enable mechanisms to generate phenomena that are quite surprising given the operations of the components taken in isolation. Such organization must be discovered—it cannot simply be derived from knowledge of lower-level parts and their operations. Moreover, the organized environments in which mechanisms operate must also be discovered. It is typically the higher-level disciplines that have the tools for discovering the organization within and between mechanisms. Although these inquiries are constrained by the knowledge of the parts and operations constituting the mechanism, they make their own autonomous contribution to understanding how a mechanism actually behaves. Thus, mechanistic explanations provide a strong sense of autonomy for higher levels of organization and the inquiries addressing them even while recognizing the distinctive contributions of reductionistic research investigating the operations of the lower level components.

Introduction

Two related legacies of mid-20th century philosophy of science continue to be entrenched in the philosophy of the cognitive sciences—the deductive-nomological model of explanation and the account of theory reduction. Philosophers talk as if the way to explain mental activities is by subsuming them under laws. Although psychologists sometimes advert to laws (as Cummins notes, usually designating them *effects*), these are seldom appealed to in order to explain their

---

[1] I thank Cory Wright for productive discussion of points in this paper and for presenting an earlier version of this paper for me at the 2004 meeting of the Society for Philosophy and Psychology in Barcelona, Spain. I also thank the participants and audience at a symposium of the Boston University Center for Philosophy and History of Science in May, 2005, during which a version of this paper was also presented, for their helpful questions and discussion. Finally, I thank Huib Looren de Jong for his helpful comments.

instances but rather to identify the phenomena (the regularities) requiring explanation (Cummins, 2000; Bechtel & Abrahamsen, in press). When they offer explanations, psychologists like biologists propose accounts of the mechanism responsible for the phenomenon. I will say more about mechanistic explanations below, but first I introduce the second legacy of mid-20th century philosophy of science, the model of theory reduction. On this traditional philosophical account of interlevel relations in science (Oppenheim & Putnam, 1958; Nagel, 1961) the laws of a higher-level science (e.g., psychology) are reduced by being derived from the laws of the lower-level science (e.g., neuroscience) together with bridge principles and boundary conditions (for a discussion of the traditional theory reduction model, variants on it, and criticisms of it, see Bechtel & Hamilton, in press; McCauley, in press). According to this view, reduction is antithetical to any claims for autonomy of higher levels of organization and the sciences that investigate them. In order to derive one set of statements from another, the former cannot assert anything not already asserted by the latter and hence is redundant to it. Hence, it makes no autonomous contribution.

Given the theory reduction framework, theorists defending the autonomy of psychology and other higher level sciences have argued against the possibility of reduction. They have targeted their criticisms on the possibility of bridge principles linking the vocabulary of the lower- and higher-level sciences, maintaining that higher-level phenomena are multiply realized by an extremely broad and not well delineated set of lower-level realizers (Fodor, 1974; Putnam, 1967). For over 20 years philosophers took multiple realizability as an obvious truth, but recently a number of philosophers have called into question its significance for science (Bechtel & Mundale, 1999; Bickle, 2003; Polger, 2004; Shapiro, 2004). Putnam appealed to examples such as hunger and pain, psychological phenomena which he claimed were realized in a huge diversity of animals despite radical differences in their brains. But it is important to note that hunger and pain behaviors vary radically across species, so if we view psychological processes at even a moderately fine grain, the psychological processes that are realized in different species are only similar, not identical. On the other hand, neuroscientists operate comparatively in identifying brain processes, treating as comparable brain activity in different species. If one uses a fine-grained account of both mental and neural processes, there is no evidence of the same mental state being realized in different ways. If, on the other hand, one adopts a coarse-grained account of both mental and neural processes, then, given the highly conserved nature of biological mechanisms, the different realizations of higher-level phenomena will themselves tend to be grouped in a common type. Accordingly, on neither a fine- or a coarse-grained account do appeals to multiple realizability provide a compelling argument against reduction and for the autonomy of higher-level explanations.

The theory reduction model, however, is much stronger than what scientists generally have in mind when they speak of reduction. For many scientists, research is reductionistic if it appeals to lower-level components of a system to explain why it behaves as it does under specified conditions. This sense of reduction is captured in the accounts of mechanistic explanation presented in the next section. I will argue in subsequent sections that the reductions achieved through mechanistic explanations are in fact compatible with a robust sense of autonomy for psychology and other special sciences, albeit a sense of autonomy no reductionist except one seeking hegemony for the lower level (Bickle, 1998; 2003) should have any desire to deny. This autonomy maintains that psychology and other special sciences study phenomena that are outside the scope of more basic sciences but which determine the conditions under which lower-level components interact. In contrast, the lower-level inquiries focus on how the components of mechanisms operate when in those conditions. Importantly, this defense of autonomy does not require appeal to multiple realizability, but only to the fact that investigations at higher-levels of organization provide information additional to that provided by the account of how the parts of a mechanism operate.

Just as this notion of reduction is much weaker than that offered by the theory reduction model, so also is the notion of autonomy less than that defended by the model's opponents. In making room for the autonomy of psychology or other special sciences, I am not arguing that these inquiries should be pursued in ignorance of lower- (or higher-) level inquiries. Sometimes knowledge about the components of a mechanism can guide inquiry into how the mechanism engages its environment and when such knowledge is available, ignoring it is foolhardy. The same, though, applies in the opposite direction—knowing how a mechanism behaves under different conditions can guide the attempt to understand its internal operation. Inquiries at different levels complement each other both in the sense of providing information that cannot be procured at other levels and also in the sense of providing information that can limit the range of possibilities at other levels.

Mechanisms and mechanistic explanations

Within the life sciences, explanation frequently takes the form of identifying the mechanism responsible for a phenomenon of interest—circulation of the blood, photosynthesis, protein synthesis, reproduction, etc. Although there are differences among the various accounts of mechanism and mechanistic explanation that have been advanced (Bechtel & Richardson, 1993; Bechtel & Abrahamsen, 2005; Glennan, 1996; 2002; Machamer, Darden, & Craver, 2000), these are inconsequential for the question of the relation of reduction and autonomy. My preferred characterization of a mechanism is that

> a mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena (Bechtel & Abrahamsen, 2005).

Explanation then consists in representing (sometimes verbally, but often in diagrams or in a computational model) the mechanism responsible for the activity and showing how it accounts for the phenomenon.

As in biology, most explanations in psychology involve the identification and characterization of a mechanism responsible for the phenomenon of interest—decision making, memory encoding and retrieval, language comprehension and production, etc. (Wright & Bechtel, in press). Until recently, psychologists have lacked the resources to identify the parts of the brain responsible for the various operations invoked in a mechanistic account, and have settled for identifying operations and modeling their interactions in generating the phenomenon. The information processing models they advanced proposed sequences of operations on informational structures (representations) that would account for the phenomenon of interest (e.g., problem solving). These models were generally tested using behavioral measures such as reaction times (Posner, 1978). With the advent of cognitive neuroscience, mechanistic explanations of mental phenomena have increasingly included identification of the brain parts responsible for the component operations.[2] Techniques such as neuroimaging enable researchers to identify the brain regions involved in executing a cognitive task. The goal of such research is not just to learn where operations occur, but to use such knowledge to further constrain and revise proposed accounts of mechanisms (e.g., by discovering that what were taken to be two entirely separate cognitive activities invoke the same neural process and asking how the same cognitive operation figures in both activities).

To give some substance to the idea of mechanistic explanation, I will develop a few central features of how mechanistic explanations commonly develop. First, a researcher begins by delineating the phenomenon of interest. One product of the cognitive revolution of the 1950s and 1960s was the differentiation of types of memory. Miller, Norman, and Sperling distinguished echoic, short-term, and long-term memory in terms of their behavioral characteristics (Neisser, 1967) while Tulving and his collaborators differentiated various forms of long-term memory (semantic, episodic, and procedural) (Tulving, 1983). In addition, investigators distinguished different time stages in memory—encoding, storage, and retrieval. Although some investigators were interested in characterizing the mechanisms involved (Atkinson & Shiffrin, 1968), this inquiry focused on differentiating and characterizing various memory phenomena. The fact that the ability to encode new episodic memories was selectively destroyed in Scoville's patient HM both secured the delineation of episodic memory encoding as a distinct phenomenon and implicated the hippocampus in it (Scoville & Milner, 1957).

All of these steps, however, are preliminary to advancing mechanistic explanation. This requires decomposing the mechanism into component parts and operations and localizing each operation in the appropriate part. Although psychologists have developed tools to demonstrate when different tasks involve different operations (Kolers & Roediger, 1984; Roediger, Buckner, & McDermott, 1999), they have made far less progress in specifying what the operations are (Bechtel, 2005). In some cases the determination of the relevant brain structure can provide clues as to the nature of the operations. For example, the determination that the hippocampus plays a role in the encoding of long-term memories inspired researchers to investigate whether the neuroarchitecture of the hippocampus might provide clues to how it realized this phenomenon. The hippocampus is comprised of several different regions, each of which has a distinctive architecture (the relations between these areas are shown schematically in Figure 1). The dentate gyrus, for example, has ten times the number of neurons as the entorhinal cortex from which it receives inputs. This, together with the fact that only a few cells in the dentate gyrus respond to a given input stimulus, suggested that the dentate gyrus might serve to maintain separation between similar inputs and facilitate remembering events as distinct. The dentate gyrus projects

---

[2] Although some philosophers speak of reducing psychological processes to neural ones, the relation between the components identified in psychology and those identified in neuroscience is actually identity between the component performing a psychological operation and a neurally identified brain part. It is important to recognize that these identity claims are often advanced at the outset of research as hypotheses to guide further inquiry. Accordingly, McCauley and I speak of the *heuristic identity theory* (Bechtel & McCauley, 1999; McCauley, in press). This addresses an objection raised to traditional identity theories (Smart, 1959) that evidence could never support psycho-physical identity claims as opposed to psycho-physical correlations. Accepting an identity hypothesis commits a researcher to a host of further consequences. If these consequences are born out, the research program will have incorporated the identity claim at its foundation and researchers will not be asking whether a mere correlational claim could serve as well.

to the CA3 fields, whose neurons are highly connected to each other via recurrent loops, which suggests that they may compute similarity between patterns (Redish, 1999).

Figure 1. Schematic diagram of the hippocampal system. Information from widespread areas of neocortex converge on the parahippocampal region (parahippocampal gyrus, perirhinal cortex, and entorhinal cortex, EC) to be funneled into the processing loops of the hippocampal formation. The tightest loop runs from EC into the core areas of the hippocampus (CA1 and CA3) and back. An alternative route to CA3 goes through the dentate gyrus and an alternative route back to EC from CA1 goes through the subiculum, which is not part of the hippocampus proper. Not shown are a number of subcortical inputs and details of pathways and their synapses.

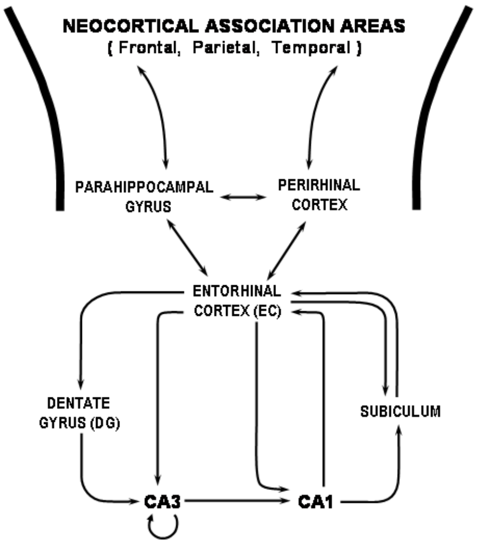

Figure 1. Schematic diagram of the hippocampal system. Information from widespread areas of neocortex converge on the parahippocampal region (parahippocampal gyrus, perirhinal cortex, and entorhinal cortex, EC) to be funneled into the processing loops of the hippocampal formation. The tightest loop runs from EC into the core areas of the hippocampus (CA1 and CA3) and back. An alternative route to CA3 goes through the dentate gyrus and an alternative route back to EC from CA1 goes through the subiculum, which is not part of the hippocampus proper. Not shown are a number of subcortical inputs and details of pathways and their synapses.

Beyond distinguishing phenomena and decomposing the responsible system functionally and structurally, mechanistic explanation requires researchers to determine how the various component parts are organized such that the operations are coordinated appropriately to realize the overall phenomenon. In the case of the hippocampus, it is important to know which regions send neural signals to which other regions. In the case of the hippocampus, the components are organized in a complex loop structure (see Figure 1). Furthermore, insofar as the hippocampus does not operate in isolation, investigators must relate the operations of the components of the hippocampus to other neural structures. Often computational modeling is required to evaluate whether a particular organization of parts carrying out proposed operations would be sufficient to realize the phenomenon (Rolls & Treves, 1998; McClelland, McNaughton, & O'Reilly, 1995).

When mechanistic explanations appeal to the components of mechanisms to explain their behavior, they are clearly reductionistic. Moreover, the process of decomposition is iterative—the operation of a component part can itself be explained by another round of decomposition and localization. In fact, however, few mechanistic explanations involve more than two iterations of such decomposition. A major reason for this is that each decomposition addresses a different phenomenon (the operation of a component part as opposed to the operation of the whole mechanism). Once researchers have identified the parts of a mechanism and determined what operations they perform and how these operations are coordinated so as to enable the mechanism to realize the target phenomenon when in a given environment, the question that drove the inquiry has been answered.

Levels of organization

Reduction is often characterized as appealing to lower levels. Despite frequent references to levels in discussions of reduction, what constitutes a level is often unspecified. In accounts of theory reduction, levels are often associated with broad scientific disciplines, so that one finds references to the level of physics, the level of psychology, etc. and to disciplines being reduced to disciplines at lower levels. Such accounts present a variety of problems. Physics deals with a broad range of entities, from the very small (sub-atomic particles) to the extremely large (galaxies). There doesn't seem to be a clear sense in which these reside at the same level. Nor is there an obvious sense in which the all the phenomena of physics lie at a lower level than those of say biology, which also deals with entities ranging from very small (viruses) to very large (ecosystems).

The division of inquiry into broad disciplines such as physics, biology, and psychology, has much more to do with what humans are interested in studying—the behavior of ordinary physical objects (physical sciences), living organisms (biological sciences), behaving systems (behavioral sciences), and social activity (social sciences)—than with a hierarchy of levels of entities (Abrahamsen, 1987). As Abrahamsen notes, the phenomena studied by parts of physics also figure in the phenomena associated with life; accordingly, there are bridges between the physical and biological sciences (as well as between the other main divisions). Within these main divisions, differences in discipline may more closely correspond to a levels hierarchy. For example, molecular biology seems to be focused on a lower level than physiology, which deals with phenomena at a lower level than ecology. But even here relating disciplines to levels runs into difficulty. Microbiology and bacteriology seem to deal with different phenomena at the same level.

Attempts to sort out levels in terms of disciplines are fraught with problems (for further discussion, see Craver, forthcoming, chapter 5). A very different approach is to start not with the categorization provided by disciplines, but with phenomena in nature. An initially plausible view is to demarcate levels in terms of the size of the entities involved—small things are at a lower level than big things. This is the picture Churchland and Sejnowski (1988) adopt when they appeal to size scales to delineate levels of organization in the nervous system: molecules (1Å), synapses (1µm), neurons (100µm), networks (1mm), maps (1cm), and systems (10cm). Wimsatt likewise proposes size as a way to differentiate levels, though he further elaborates the view by proposing that entities of the same size tend to "interact most strongly and frequently" (Wimsatt, 1976). Thus, levels are "*local maxima of regularity and predictability in the phase space of alternative modes of organization of matter*" (Wimsatt, 1994). Accordingly, he develops a stratified account according to which entities tend to fall into discrete clusters based on size:

levels "*are constituted by families of entities usually of comparable size and dynamical properties, which characteristically interact primarily with one another, and which, taken together, give an apparent rough closure over a range of phenomena and regularities*" (Wimsatt, 1994).

If it were true that entities of a given size range tended to interact most frequently with other things of similar size, then this would be a principled approach. There are, however, plenty of examples of things of different sizes interacting causally. There are gravitational forces between very large objects (the earth) and very small objects (a molecule of hydrogen in the atmosphere). Storms can sweep seeds from one local to another. Likewise, small things can causally affect big things—a bullet or a virus can kill an elephant. Absent a quantitative analysis, it is not clear that the claim that things tend to interact primarily with things of their own size is true.

Wimsatt often combines his analysis of levels by sizes with a compositional or mereological treatment of levels:

> By level of organization, I will mean here compositional levels—hierarchical divisions of stuff (paradigmatically but not necessarily material stuff) organized by part-whole relations, in which wholes at one level function as parts at the next (and at all higher) levels (Wimsatt, 1994).

The first thing to note is that, although a consequence of this compositional analysis is that parts are smaller than the whole they constitute, this account is very different. It does not require that all parts of an entity must be of the same size—parts may vary radically as long as each is smaller than the whole. Second, the analysis only applies locally. Parts will be at a lower level than the whole to which they belong, but the compositional analysis does not tell us how to relate the parts to things outside the whole.

The mereological account of levels on its own, though, allows for arbitrary differentiation of the parts of a whole (see Craver, forthcoming, for a discussion of this and many other problems with formal treatments of mereology). Wedding the mereological account directly to mechanisms solves this problem. The component parts of a mechanism are the entities that perform the operations which together realize the phenomenon of interest. A structure within the mechanism may be well-delineated (it has boundaries, continues to exist over time, is differentiated from the things around it, etc.). However, if it does not perform an operation that contributes to the realization of the phenomenon, it is not a working part of that mechanism. For example, while the gyri and sulci of the brain are well-delineated, they are not working parts of the brain but by-products of the way brains fold to conserve the length of axons (van Essen, 1997). The different working parts that constitute a level may be of different sizes—large parts, such as cell membranes, may interact with small parts such as individual sodium ions which are maintained in different concentrations on different sides of it.

To identify levels in terms of mechanisms, one starts with the mechanism identified in terms of the phenomenon for which it is responsible. The mechanism's working parts constitute the next lower level (see Figure 2). A consequence of this view is that levels are identified only with respect to a given mechanism; this approach does not support a conception of levels that extend across the natural world. Thus, it cannot address the question of whether glaciers, for example, are at the same level or at a higher or lower level than elephants, since they are not working parts of a common mechanism or related compositionally.

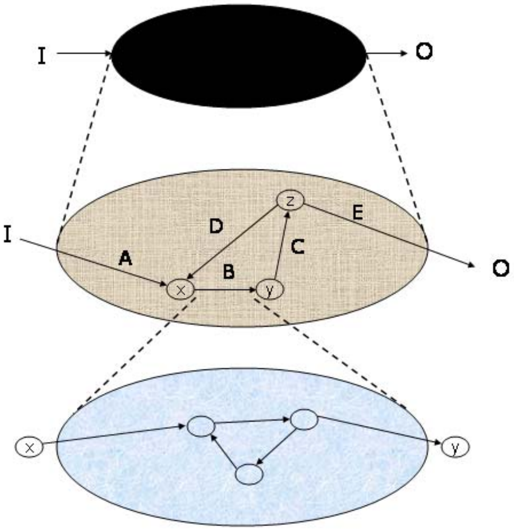

Figure 2. Moving down levels within a mechanism. The whole mechanism, shown in the top panel, is responsible for processing input I into output O. To explain how the mechanism is able to do this, investigators decompose it into its parts performing operations (indicated by uppercase letters) and determine how they produce changes in substrates (indicated by lowercase letters) in the system (middle panel). If one wants to explain how one component, B, performs its operation, a further round of decomposition into its parts and operations and determination of how they relate to one another is required (bottom panel).

The account of levels within a mechanism can be generalized to multiple levels of organization once we recognize that a working part of a mechanism may itself be a mechanism. To explain how it performs its operation, investigators decompose the part into its own working parts. These parts are at a lower level than the component of the initial mechanism, and hence two levels below the mechanism as a whole. This process can clearly be iterated. The local nature of levels, however, emerges again as soon as we move down two levels. Because of the lack of a compositional relation between the sub-parts of one component of the mechanism and those of another component of that mechanism, the question of whether the sub-parts of two components are at the same level is not well defined.

While this approach cannot provide a global account of levels, it is sufficient for understanding the respects in which mechanistic explanation is reductionistic. The local character of the treatment of levels also has a rather surprising consequence that distinguishes mechanistic reduction from traditional views of reduction. Traditional views tend to assume that one can reduce higher-level explanations level by level until one reaches a fundamental level. On a theory reduction account, the theories at this level provide the foundation on which all higher-level theories are grounded. Even those who forego a theory reduction perspective find it plausible that at some fundamental level we can identify the parts and operations out of which all higher-level mechanism are built. Theorists such as Kim (1998) then maintain that if we had a complete account of causal processes at this level, we would be able to determine all that happens in the universe. We would simply supply the initial conditions and make deductions from the laws governing the most basic level. Higher-level causal relations would over-determine outcomes since these would already be determined at the lower level. But if the notion of levels is defined only locally, then on the mechanistic account we are not confronted by the prospect of a comprehensive lower-level that is causally complete and closed. Such a picture of complete causal determination at a lower level is further brought into question when we consider why mechanistic explanations require relating levels of organization.

Relating levels: reduction and autonomy

Although mechanistic explanation is reductive insofar as it appeals to the component parts and operations within a mechanism to explain the behavior of the mechanism, the reductive aspect alone is insufficient to explain the behavior of the mechanism. The parts of a mechanism behave in a particular way because of how they are organized in the mechanism. Information about how the parts are organized goes beyond the account of the parts and their operations. Moreover, the mechanism interacts causally with other entities. These interactions provide the input and set the conditions for the operation of the mechanism and information about them is not part of the reductive account characterizing the parts and operations within the mechanism. Securing information about both the organization within the mechanism and the relations between the mechanism and its environment requires going beyond the reductive aspect of mechanistic explanation and incorporating the results of other, autonomous inquiries.

This need to incorporate both a reductive component and information secured from autonomous inquires at higher levels can be illustrated in a simple biological example. To produce your favorite ale, a brewer brings together the ingredients necessary for fermentation and creates the proper condition for fermentation to occur. At a lower level, individual yeast cells take in glucose and generate alcohol along with carbon dioxide. At a yet lower level, enzymes such as glucose-6-phosphatase in the yeast's cytoplasm catalyze specific reactions in the biochemical pathway from glucose to alcohol. The operation of the enzymes in catalyzing a specific reaction depends upon the context in the yeast cell, and that cell's situation in the brewer's vats. The enzymes in yeast remain able to catalyze reactions, but if no glucose is available or the brewer does not create the proper conditions, fermentation will not occur. These conditions are additional to providing the enzymes and the inquiry into which conditions are best suited for fermentation is distinct from the inquiry into the operation of the enzymes. Indeed, knowledge of how to set up conditions for fermentation was acquired by brewers long before the development of biochemistry in the early 20th century,[3] and that knowledge was not supplanted by the investigations of biochemists. On the contrary, biochemists employed such knowledge in setting up the experimental conditions in which they could study the operation of the enzymes.

A similar relationship between reductive appeals to the parts of a mechanism and autonomous investigations of the organization of those components and of how the mechanism is situated in

---

[3] Indeed, physiological investigations of yeast were conducted while it was still unknown what happened inside them. Pasteur (1860), for example, conducted detailed studies of the conditions in which fermentation would occur and established the important result that yeast carry out fermentation when in an oxygen-free environment, an effect known as the Pasteur effect. But Pasteur himself rejected the prospects of a chemical account of the reactions occurring within yeast cells.

an environment is found in the case of the primate visual system. Reductive investigations in the second half of the 20th century identified the different brain areas that comprise the visual system and the operation of each in extracting information from visual inputs (van Essen & Gallant, 1994). Such inquiry determined that V1 detects edges (amongst other things), V4 computes color constancy and identifies shape, and MT computes perceived motion. None of these operations is itself the activity of seeing an object. Although one might be tempted to identify seeing objects with the firing of cells in inferotemporal cortex, where the process of computing the identity of the object is thought to occur, this temptation should be resisted. Inferotemporal cortex does not operate in isolation, but only in conjunction with other components of the visual system. It simply represents an output component of a complex mechanism. Each of the various components of the mechanism carry out operations that together realize the phenomenon of perceiving objects both as entities of a particular kind and also as having a given shape or a certain color. Moreover, seeing only occurs when all the parts are organized in the right way and the whole organism engages a visual world. A person recognizes a dog when her retinal cells are activated by patterns of light reaching them from a dog in the environment, the components of her visual system are properly connected, and other brain processes, including those required for attention, operate normally. These other processes, some of them involving mechanisms themselves, produce the conditions for the phenomenon in question.

Understanding how the visual system is organized, coordinated with other physiological systems, and responsive to external stimuli, requires knowledge beyond the specification of the parts of the visual system and their operation. Moreover, investigations of these relations can determine important regularities that are not provided by the reductive inquiry. For example, psychophysicists began to identify regularities about how our visual system responds to sensory stimuli almost 150 years ago (Fechner, 1860; Stevens, 1957), and these investigators did not know about the parts of the system or how those parts operated. James Gibson (1966; 1979) and subsequent ecological psychologists identified many sorts of information to which the visual system is responsive without engaging in any study of the components of the system. Moreover, these researchers have produced knowledge that could not have been acquired by such reductive inquiry. Psychophysicists and ecological psychologists complement the reductive inquiries of neurophysiologists and have not been rendered unnecessary by the neurophysiologists' success.[4]

Not only can one study the performance of a mechanism without knowing its component parts and their operations, but what the mechanism as a whole does is typically quite different than the operations performed by its parts. Neurons in different parts of the visual system generate action potentials in response to release of transmitters by cells on which they synapse, for example, while the (ventral) visual system as a whole identifies what object is presented to the person and makes that information available to other cognitive systems (those engaged, for example, in encoding memories or making decisions). As this illustration also makes clear, the mechanism as a whole may in fact constitute a component of a larger mechanism that does something still

---

[4] Adopting this perspective, we can recognize the mistake of eliminative materialists in maintaining that knowledge of the brain will eliminate the need for folk psychology. Folk psychology, like social psychology, characterizes regularities in the way cognitive agents respond to situations arising in their environment (Bechtel & Abrahamsen, 1993). This is not information that neuroscientists themselves are interested in or have the tools to procure (but see Bickle, this volume, for an opposing view).

different (enabling the organism to act). The information processing of neural ensembles is different from the production of action potentials in individual cells.

The fact that mechanisms perform different activities than do their parts manifests itself in the fact that the activities of whole mechanisms are typically described in different vocabulary than are component operations. Traditional accounts of theory reduction implicitly recognized this fact by requiring bridge principles to connect the different vocabularies used in different sciences, but little notice was given as to why different sciences employ different vocabularies. The vocabulary used in each science describes different types of entities and different operations—one describes the parts and what they do whereas another describes the whole system and what it does. Relating the vocabulary used in the different sciences requires consideration of the compositional relations between the entities and their operations. The substantive knowledge required to establish these relations is not derived from the lower-level laws but requires additional empirical investigation. Recognizing this reveals that even the theory-reduction account must incorporate higher-level knowledge, and so is not as reductive as it appears.

Throughout this discussion I have been making reference to the way the components of a mechanism are organized. It is the fact that mechanisms organize the operations performed by the parts that enables them to do things no part alone can do, which in turn requires higher-level inquiries into how the mechanism engages its environment. Organization itself is not something inherent in the parts (even self-organizing systems only organize themselves under appropriate conditions). Accordingly, investigators who already understand in detail how the parts behave are often surprised by what happens when they are organized in particular ways. To appreciate this, consider engineering design. The primary activity of a designer is to put components together in novel ways to produce new activities. A designer does not develop a new mechanism by creating it de novo, but by imposing novel organization on components that already exist. An indication of the critical contribution of organization is the fact that engineers who are successful in discovering organizational principles that enable mechanisms to perform activities are able to secure patents for their design are rewarded financially or with fame.

In virtue of being organized systems, mechanisms do things beyond what their components do. But beyond this, the organization of the components typically integrates them into an entity that has an identity of its own. As a result, organized mechanisms become the focus of relatively autonomous disciplines—disciplines which deploy their own tools of investigation and develop their own distinctive accounts of the phenomena associated with these mechanisms. Thus, to understand the autonomy achieved by higher levels within a mechanistic framework, we need to focus on the sorts of organization that figure in mechanisms.

Types of organization and the generation of higher levels

By introducing the notion of *aggregativity*, Wimsatt (1986) provided a baseline account of collectives that lack organization. In an aggregate, such as a pile of sand, the component parts are simply amassed together without any specific organization. The components each behave as they would outside of the aggregate; specifically, they do not interact in ways that result from specific dependencies of one part on another. Parts can be substituted for one another and, as new parts are added to the aggregate, the behavior of the whole depends simply on the number of parts present (at least until critical points are reached where interesting non-linear interactions such as landslides begin to occur). As a consequence, aggregates do not produce entities at higher levels or require new inquiries.

A first step away from aggregativity is for some parts to depend on the prior operation of others in order to perform their own operations. In a linear organization (of which human assembly lines provide exemplars), the product of the operation of one part is operated on by the next in a sequence until an output product is produced. A system organized in this way can accomplish more than can any given component. Whereas in an aggregate, any part can replace another, linearly organized systems depend on having the right order of parts so that the products of one operation are made available as inputs to a subsequent operation. Nonetheless, the organization is still essentially additive: the operation of one component is simply added to the operation of another component. Understanding linearly organized systems requires only modest additions to the knowledge of how the components work.

Engineers often begin by designing linear systems and scientists begin by proposing linear models for natural phenomena. One reason is that our conscious thinking is sequential—we think one thought and then another. So when we try to understand a system, we conceptualize one operation followed by another. (Even when processes are known not to be strictly linear—when they depend upon feedback loops that connect operations earlier in the sequence to those later in the sequence—scientists often represent them linearly. Thus, despite the fact that some of the reactions in fermentation are linked with ones earlier or later in the pathway, the fermentation pathway is commonly represented as a linear sequence of reactions.) But scientists, especially in the life sciences, have repeatedly discovered that natural systems are not organized linearly but exhibit various types of interactivity such as cycles. In cycles, the product of a sequence of operations feeds back into an earlier step in the process. We have already encountered one example: the cyclic pathways linking the various regions of the hippocampus (see Figure 1). The Krebs cycle in oxidative metabolism is another clear well-known exemplar of cyclic organization—the final product of a series of oxidation reactions, oxaloacetic acid, is combined with additional acetyl-CoA arriving from carbohydrate, fat, or protein metabolism, to produce citric acid. This is then oxidized in another round of the cycle. Biochemists, physiologists, and ecologists have discovered a plethora of such cycles.

Cycles often allow the results of an operation to feedback on earlier operations and regulate them. Figure 3 shows a negative feedback loop at the entry point to the Krebs cycle where accumulation of acetyl-CoA serves to inhibit the earlier formation of pyruvate from phosphoenolpyruvate, thereby preventing additional accumulation of acetyl-CoA when it is unneeded. The usefulness of negative feedback organization seems to have first been discovered in the third century BCE by Ktsbios in his design of the water clock. In order for water to flow at a regular rate into the vessel measuring time, a constant volume of water in the supply tank was required. Ktsbios achieved this by inserting a float where water entered the supply tank so that whenever the water exceeded that height, the float would block the entrance, preventing any more water from entering. The idea of negative feedback, however, was not easily generalized to other contexts but had to be continuously rediscovered (Mayr, 1970). Watt's introduction of the centrifugal governor for the steam engine, and Maxwell's (1868) mathematical analysis of it, led finally to recognition of negative feedback as a general technique for regulating behavior of mechanisms. In the mid 20th century the cyberneticists celebrated negative feedback and promoted cyclic design as a fundamental principle (Wiener, 1948) in the organization of biological and social systems. From the point of view of understanding a mechanism, when negative feedback is employed, the operation of some components falls under the influence of other components, and a theorist cannot account for the behavior of the whole system by just adding together the outputs of the component parts. Rather, the theorist must appreciate how the components are constrained by the conditions created by the operation of other components.

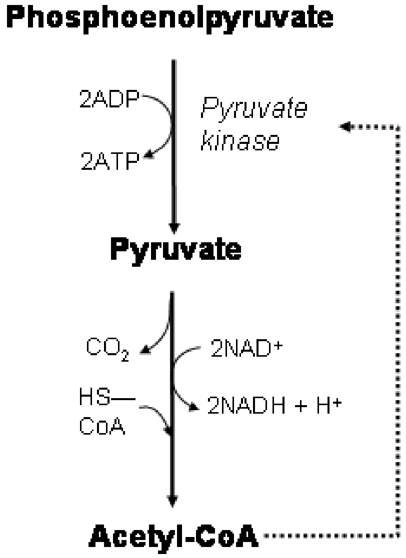

Figure 3. Feedback loop (dotted line) at the point of entry to the Krebs cycle. If acetyl-CoA builds up from not being metabolized by the Krebs cycle, it feeds back onto the dephosphorylation of phosphoenolpyruvate, causing that reaction to halt.

Already with a simple mechanism that performs a function different than the operations of its parts we had introduced a higher level of organization that exhibited a kind of autonomy from inquiries at the lower-level. The higher level investigated the engagement of the mechanism as a whole with its environment. But as more complex modes of organization are introduced, this autonomy grows. When designs such as negative feedback are introduced, the system develops a kind of insulation from certain perturbations in the environment. Understanding the behavior of the mechanism requires not just knowing its parts and operations but the capacities provided by the negative feedback.

Although negative feedback is now reasonably well known, positive feedback, in which the products of two operations each facilitate the other, often receives less attention. This is due to the fact that positive feedback in many situations leads to out of control, runaway behavior. But in some contexts such interactions are a powerful force for developing higher-level structure. Positive feedback can provide a basis for self-organization—the ability of a set of components to organize themselves into structures which perform operations beyond what the components themselves are capable (Kaufmann, 1993). Here I will limit my focus to positive feedback within networks of already organized components.

For the most part, human thinking about networks has started from one of two designs—regular lattices in which components are connected to those closest to them and randomly connected networks (investigated by Erdös and Rényi) in which each component has a random probability of being linked directly to any other unit. But two developments in thinking about network design in the late 20th century revealed forms of organization in networks that result in far more interesting sorts of system behavior. Inspired by Stanley Milgram's 1967 letter mailing experiment and John Guare's 1990 play "Six degrees of separation," Duncan Watts (Watts & Strogratz, 1998) explored what he termed *small worlds*. These are networks in which there is high local connectivity, as in regular lattices, supplemented by a few long-range connections. These long-range connections serve to bring the path-length between any two nodes into the same range as in a randomly connected network. Such networks show how one can have systems that are largely modular in design (clusters of units are primarily connected with each other and can operate as a higher-level unit) without the modules becomes completely isolated from one another, as in Fodor's (1983) account of mental modules. A few long range connections enable these modules to coordinate their activity, and thus function as integrated systems at yet higher-levels of organization.

The second new idea was that nodes in networks might have very different numbers of connections to other nodes, some being connected to just one or two, and a very few to a very large number. Barabási and Albert (1999) characterize networks in which the connectivity of nodes drops off according to a power-law as *scale-free*. In scale-free networks many components of a network can be destroyed while the rest of the network retains its integrity and continues to function. Many naturally occurring networks, such as metabolic systems in cells and human social networks, exhibit the properties of small worlds and scale-freeness. The ubiquity of such networks raises the question of what additional fruitful properties result from such designs that we have previously failed to appreciate.

One can gain further appreciation of the importance of non-linear modes of organization in networks of components by considering the recent history of research on artificial neural networks. The discovery of the backpropagation training algorithm for multi-layer networks in the mid-1980s (Rumelhart, Hinton, & Williams, 1986) focused a great deal of attention on feedforward networks in which activity in one layer of units contributed to activity of units in the next layer. These networks operate in a sequential manner.[5] But other researchers were already concentrating on far more interactive networks (Hinton & Sejnowski, 1986). One simple addition to a feedforward network is to employ the activity pattern on later units as additional inputs to earlier layers of units in subsequent epochs, resulting in what is called a *simple recurrent network*. Elman (1991; 1993) showed that such networks could learn to process complex grammatical forms involving the type of long-range dependencies found in natural languages without invoking internal representations of linguistic structure. Other network modelers have used positive feedback between units in networks to capture some of the rich dynamics of human behavior (van Leeuwen, Steyvers, & Nooter, 1997). Recognition of the fact that many real world networks employ small-world and scale-free features has led investigators to explore their use in neural network modeling (Gong & van Leeuwen, 2003; 2004). A general feature of such

---

[5] Backpropagation, as the name suggests, already uses backwards projections from output units to weights in the network, but only in the process of training the network, not when the network is solving a particular problem.

research is to reveal how networks of simple processing units organized in more complex ways could generate behavior that initially seemed quite beyond the capacity of such networks.

My interest in this section is in how organization enables systems of components to exhibit behavior different in character than that exhibited by the components. Such organized systems become the focus of their inquiry that is autonomous from inquiry into the behavior of the components and focuses on how these systems engage the world in their own way. Herbert Simon drew attention to this in his analysis of hierarchically organized complex systems (Simon, 1962). A century earlier Claude Bernard (1865) investigated the ability of living systems to maintain themselves through a wide-range of environment changes and introduced the important distinction between the internal environment of an organism and its external environment. He further articulated the idea that individual parts of organisms perform operations needed to maintain the constancy of the internal environment. This idea was partially explicated with Cannon's (1929) account of mechanisms of homeostasis and the cyberneticist's account of negative feedback. But a critical feature of complex systems to which Simon drew attention was the interface—the boundary between a system and its environment. In biological systems, membranes provide a means by which organisms (including single-cell organisms) can control admission of foodstuffs and other needed materials from their environments and expel toxins back into the environment. Maintenance of membranes does not come for free—organisms must build and maintain their own membranes, drawing on energy and resources from their environment. Accordingly two of the critical components that Tibor Gánti (1975; 2003) included in his conception of a chemoton, the simplest chemical system that he maintained would exhibit the properties of life, were a metabolic system and membranes, with the membrane segregating the metabolic system from the environment and controlling access to it. (The third was a control system, which I will not consider here.) At the center of Gánti's conception of the chemoton are cyclic processes which enable the chemoton to maintain itself and even build more of its metabolic components and its membrane (relying on physicochemical processes to divide into daughter chemotons).

The important feature of such systems as Gánti's chemoton is that they are able to maintain themselves in relative independence of their local environments. To capture this feature of living systems, Alvaro Moreno and his colleagues speak of them as *autonomous*,[6] and characterize an autonomous system as:

> a far-from-equilibrium system that constitutes and maintains itself establishing an organizational identity of its own, a functionally integrated (homeostatic and active) unit based on a set of endergonic-exergonic couplings between internal self-constructing processes, as well as with other processes of interaction with its environment" (Ruiz-Mirazo, Peretó, & Moreno, 2004).

Although the notion of an autonomous system is different from the autonomy of different inquiries, which has been my focus here, they are related. The fact that biological systems are autonomous is part of what motivates inquiries into their behavior and the organizational

---

[6] The notion of autonomy has also been developed by Varela (1979; see also Maturana & Varela, 1980). Varela, however, does not invoke thermodynamic considerations in his account of autonomy. It is the thermodynamic phenomenon of being far from equilibrium which imposes on living systems the need to recruit resources and energy from their environment and to use these to maintain themselves. It is this feature that makes autonomy so critical to understanding biological mechanisms.

conditions that give rise to it that is distinct from inquiries into their component parts and their operations. The relationship, though, is the same as in mechanisms generally: the parts, which are the focus of reductive inquiry, are organized and situated in an environment. Understanding how they are organized and situated in not redundant to what the reductive inquiry reveals. It rather provides additional information that is required to understand the mechanism. The recognition that biological systems are autonomous systems simply makes the need for these inquiries focused on higher levels all the more important.

Conclusion

Mechanistic explanation is reductionistic insofar as it appeals to the components of a mechanism to explain its activity. But insofar as the phenomenon generated by a mechanism depends upon the organization of the parts and the conditions impinging upon the mechanism from without, mechanistic explanation also recognizes the autonomy of higher-level investigations. Modes of organization are not determined by the components but are imposed on them. (This is true even when systems self-organize—once they have organized, the components are subject to the constraints imposed by the organization as a whole.) The contribution of organization in creating mechanisms that do things their parts cannot do undergirds the need for scientists to discover the particular forms of organization realized in a mechanism. This is what higher level disciplines, such as psychology, have the resources to provide. Their autonomous contribution is secure even if higher-level activities are not multiply realizable. At the same time, mechanistic explanation also recognizes the value of reductionistic investigations into how the components perform their operations. Higher level inquiries and reductionistic inquiries complement each other, and often provide heuristic guidance to each other. Neither on its own suffices and neither can do the work of the other.

References

Abrahamsen, A. A. (1987). Bridging boundaries versus breaking boundaries: Psycholinguistics in perspective. *Synthese, 72*(3), 355-388.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 2, pp. 89-195). New York: Academic.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*, 509-512.

Bechtel, W. (2005). The challenge of characterizing operations in the mechanisms underlying behavior. *Journal of the Experimental Analysis of Behavior, 84*, 313-325.

Bechtel, W., & Abrahamsen, A. (1993). Connectionism and the future of folk psychology. In R. Burton (Ed.), *Minds: Natural and artificial* (pp. 69-100). Albany, NY: SUNY University Press.

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). Phenomena and mechanisms: Putting the symbolic, connectionist, and dynamical systems debate in broader perspective. In R. Stainton (Ed.), *Contemporary debates in cognitive science*. Oxford: Basil Blackwell.

Bechtel, W., & Hamilton, A. (in press). Reduction, integration, and the unity of science: Natural, behavioral, and social sciences and the humanities. In T. Kuipers (Ed.), *Philosophy of science: Focal issues*. New York: Elsevier.

Bechtel, W., & McCauley, R. N. (1999). Heuristic identity theory (or back to the future): The mind-body problem against the background of research strategies in cognitive neuroscience. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the 21st Annual Meeting of the Cognitive Science Society* (pp. 67-72). Mahwah, NJ: Lawrence Erlbaum Associates.

Bechtel, W., & Mundale, J. (1999). Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science, 66*, 175-207.

Bechtel, W., & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton, NJ: Princeton University Press.

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.

Cannon, W. B. (1929). Organization of physiological homeostasis. *Physiological Reviews, 9*, 399-431.

Churchland, P. S., & Sejnowski, T. J. (1988). Perspectives on cognitive neuroscience. *Science, 242*, 741-745.

Craver, C. (forthcoming). *Explaining the brain: What a science of the mind-brain could be*. New York: Oxford University Press.

Cummins, R. (2000). "How Does It Work?" versus "What Are the Laws?": Two Conceptions of Psychological. In F. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 117-144). Cambridge, MA: MIT Press.

Elman, J. L. (1991). Finding structure in time. *Cognitive Science, 14*, 179-211.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71-99.

Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf und Härtel.

Fodor, J. A. (1974). Special sciences (or: the disunity of science as a working hypothesis). *Synthese, 28*, 97-115.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Gánti, T. (1975). Organization of chemical reactions into dividing and metabolizing units: The chemotons. *BioSystems, 7*, 15-21.

Gánti, T. (2003). *The principles of life*. New York: Oxford.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

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.

Gong, P., & van Leeuwen, C. (2003). Emergence of a scale-free network with chaotic units. *Physical A: Statistical Mechanics and its Applications, 321*, 679-688.

Gong, P., & van Leeuwen, C. (2004). Evolution to a small-world network with chaotic units. *Europhysics Letters, 67*, 328-333.

Hinton, G., & Sejnowski, T. J. (1986). Learning and relearning in Boltzmann machines. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1, pp. 282-317). Cambridge, MA: MIT Press.

Kaufmann, S. A. (1993). *The origins of order*. Oxford: Oxford University Press.

Kim, J. (1998). *Mind in a physical world*. Cambridge, MA: MIT Press.

Kolers, P. A., & Roediger, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior, 23*, 425-449.

Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1-25.

Maturana, H. R., & Varela, F. J. (1980). Autopoiesis: The organization of the living. In H. R. Maturana & F. J. Varela (Eds.), *Autopoiesis and Cognition: The Realization of the Living* (pp. 59-138). Dordrecht: D. Reidel.

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.

McCauley, R. N. (in press). Reduction. In P. N. Thagard (Ed.), *Philosophy of psychology and cognitive science*. New York: Elsevier.

McClelland, J. L., McNaughton, B., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419-457.

Nagel, E. (1961). *The structure of science*. New York: Harcourt, Brace.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Oppenheim, P., & Putnam, H. (1958). The unity of science as a working hypothesis. In H. Feigl & G. Maxwell (Eds.), *Concepts, theories, and the mind-body problem* (pp. 3-36). Minneapolis: University of Minnesota Press.

Pasteur, L. (1860). Mémoire sur la fermentation alcoolique. *Annales de Chimie, 3e Ser, 58*, 323-426.

Polger, T. (2004). *Natural minds*. Cambridge, MA: MIT Press.

Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Putnam, H. (1967). Psychological predicates. In W. H. Capitan & D. D. Merrill (Eds.), *Art, Mind and Religion* (pp. 37-48). Pittsburgh: University of Pittsburgh Press.

Redish, A. D. (1999). *Beyond the cognitive map*. Cambridge, MA: MIT Press.

Roediger, H. L., Buckner, R. L., & McDermott, K. B. (1999). Components of processing. In J. K. Foster & M. Jelicic (Eds.), *Memory: Systems, process, or function* (pp. 32-65). Oxford: Oxford University Press.

Rolls, E. T., & Treves, A. (1998). *Neural networks and brain function*. Oxford: Oxford University Press.

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.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning Internal Representations by Error Propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. Foundations*. Cambridge, MA: MIT Press.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20*, 11-21.

Shapiro, L. (2004). *The mind incarnate*. Cambridge, MA: MIT Press.

Simon, H. A. (1962). The architecture of complexity: hierarchic systems. *Proceedings of the American Philosophical Society, 106*, 467-482.

Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review, 68*, 141-156.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64*, 153-181.

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.

van Essen, D. C. (1997). A tension -based theory of morphogenesis and compact wiring in the central nervous system. *Nature, 385*, 313-318.

van Essen, D. C., & Gallant, J. L. (1994). Neural mechanisms of form and motion processing in the primate visual system. *Neuron, 13*, 1-10.

van Leeuwen, C., Steyvers, M., & Nooter, M. (1997). Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. *Journal of Mathematical Psychology, 41*, 319-344.

Varela, F. J. (1979). *Principles of biological autonomy*. New York: Elsevier.

Watts, D., & Strogratz, S. (1998). Collective dynamics of small worlds. *Nature, 393*(440-442).

Wiener, N. (1948). *Cybernetics: Or, control and communication in the animal machine*. New York: Wiley.

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. (1986). Forms of aggregativity. In A. Donagan, N. Perovich & M. Wedin (Eds.), *Human nature and natural knowledge* (pp. 259-293). Dordrecht: Reidel.

Wimsatt, W. C. (1994). The ontology of complex systems: Levels, perspectives, and causal thickets. *Canadian Journal of Philosophy, Supplemental Volume 20*, 207-274.

Wright, C., & Bechtel, W. (in press). Mechanisms and psychological explanation. In P. Thagard (Ed.), *Philosophy of psychology and cognitive science*: Elsevier.