This is a preprint version of: Herschbach, M., & Bechtel, W. (2014). Mental mechanisms and psychological construction. In Lisa Feldman Barrett and James Russell (Eds.), *The Psychological Construction of Emotion* (pp. 21-44). New York: Guilford Press.

Mental Mechanisms and Psychological Construction

Mitchell Herschbach and William Bechtel

1. Introduction

Psychological construction represents an important new approach to psychological phenomena, one that has the promise to help us reconceptualize the mind both as a behavioral and as a biological system. It has so far been developed in the greatest detail for emotion, but it has important implications for how researchers approach other mental phenomena such as reasoning, memory, and language use. Its key contention is that phenomena that are characterized in (folk) psychological vocabulary are not themselves basic features of the mind, but are constructed from more basic psychological operations. The framework of mechanistic explanation, currently under development in philosophy of science, can provide a useful perspective on the psychological construction approach. A central insight of the mechanistic account of explanation is that biological and psychological phenomena result from mechanisms in which component parts and operations do not individually exhibit the phenomena of interest but function together in an orchestrated and sometimes in a complex dynamical manner to generate it.

While at times acknowledging the compatibility of the mechanist approach with constructionist approach (Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012), proponents of the constructionist approach have at other times pitched their approach as anti-mechanist. For example, Barrett (2009) claims that the psychological constructionist approach rejects machines as the primary metaphor for understanding the mind, instead favoring a recipe metaphor; constructionism also purportedly rejects the "mechanistic" picture of causation, which it portrays as linear or sequential in nature (see also Barrett, Wilson-Mendenhall, & Barsalou, in press). While some mechanistic accounts do fit this description, we will see that the mechanisms generating phenomena can be complex and dynamic, producing phenomena far less stereotypic and more adaptive than people often associate with machines. Our goal, however, is not just to render constructionism and mechanism compatible. Philosophers of science have been examining the nature of mechanistic explanation in biology with the goal of gaining new insights into the operation of science. We will identify some of the places where the mechanistic account can shed new light on the constructionist project.

The constructionist account of emotion is presented as an alternative to what Barrett (2006) calls the "natural kind" approach, which segregates emotion from other psychological kinds (e.g., perception, memory), differentiates and treats different kinds of emotion (fear, anger, happiness, sadness, etc.) as psychologically primitive, and treats these as causal processes in the mind that are ultimately to be localized in the brain either in distinct brain regions (e.g., fear in the amygdala) or in discrete neural networks. The categories emotion, perception, memory, etc., and the subdivision of emotion into fear, anger, etc., are parts of our folk accounts through which we describe our own and other agents' mental states and activities. In that context they have an important role to play—they enable us to assess the individuals with whom we interact and coordinate our activities appropriately. They become problematic, for the constructionist and for us, when they are viewed, as on the natural kind approach, as identifying basic activities performed in the mind. These mental activities are constructed from operations that are different in character, and which thus require a different vocabulary to be described. On our view, it is the orchestrated functioning of a mechanism operating in a particular physical and social environment (and shaped by its history in that environment) that exhibits these mental activities.

The points of difference between the natural kind and constructionist approaches map onto fundamental issues highlighted in accounts of mechanism. On the mechanist approach, mechanisms are characterized in terms of the phenomena for which they are responsible. This makes the project of delineating the phenomenon critical since different types of phenomena require different explanations. The constructionist's contention is that when we appreciate that emotions are constructed, we will be attuned to such things as the wide variability found in the facial expressions, behaviors, and physiology associated, for example, with fear. We concur with this, but emphasize in section 3 the need to develop nonetheless an account of phenomena that identifies regularities, for it is such regularities that become the touchstones of explanatory endeavors. In section 4 we will focus on what is the key element in mechanistic science—the appeal to component parts and operations of a mechanism. This requires decomposing the mechanism. It is out of the component parts and operations that the mechanist and constructionist both view mental phenomena such as emotions as arising; but here the mechanist can both give some perspective on why it is so difficult to identify component parts and operations of a mechanism and the historical process by which they have been identified in mechanistic sciences that are further along than psychology is at present. As a result of the emphasis on decomposition in mechanistic science, mechanistic accounts are often viewed as reductionistic. But, as we develop in section 5, accounts of mechanism also require attention to how the whole mechanism is organized and situated in its environment—a mechanistic account requires both a reductionist and a holist perspective. This becomes particularly important when the organization of the mechanism is nonsequential and the operations are nonlinear, as these conditions can give rise to complex dynamics. In section 6 we turn to one of the intriguing consequences of mechanistic research—it sometimes leads to a reconception, or what we call a *reconstitution*, of the initial phenomenon for which an explanation is sought. We consider whether psychological phenomena themselves need to be reconceptualized in light of the understanding of the mechanisms that generate them. Before turning to these issues, however, we begin with a brief primer on the new mechanistic philosophy of science.

2. A Primer on the New Mechanistic Philosophy of Science

If one consults traditional accounts of explanation in philosophy of science, one learns that explanation requires showing how the phenomenon to be explained follows from one or more laws together with some initial conditions (Hempel, 1965, 1966). While this nomological account of explanation may fit some cases in physics, it fares poorly in psychology, where there are few examples of laws and those laws which are to be found serve not to explain but to characterize phenomena in need of explanation (see Cummins, 2000, who notes that such laws in psychology are often called "effects"). It also fares poorly in biology (including neuroscience). When scientists advance explanations in both biology and psychology, they often speak of identifying the mechanism responsible for a given phenomenon. Responding to this, a number of philosophers of science have offered accounts of what counts as a mechanism in these disciplines. While the accounts differ in vocabulary and in some of their claims, they agree that a mechanism consists of an organized set of parts performing different operations whose orchestrated functioning results in the phenomenon of interest[1] (Bechtel & Richardson, 1993/2010; Bechtel & Abrahamsen, 2005; Craver, 2007; Darden, 2006; Machamer, Darden, & Craver, 2000). A mechanistic explanation both describes the mechanism and shows how it gives rise to the phenomenon. For example, beginning with William Harvey in the 17th century, explaining the phenomenon of the circulation of blood in animals involved identifying the heart as the responsible mechanism and then showing how the periodic contraction and relaxation of the muscles in the various chambers and the opening and closing of the valves (operations involving specific parts) resulted in the phenomenon. Over time the phenomenon was more fully characterized, capturing for example the precise timing of the muscle contraction, and new components such as neurons were included in the account. Although eventually researchers found it useful to characterize the phenomenon and the operations involved mathematically, the parts and operations are central to the account. Researchers do not simply construct the law of heart behavior and derive the behavior of a heart given initial conditions from such a law.

We have introduced the mechanistic account by contrasting it with more traditional accounts of explanation that present explanation as deriving a statement of the phenomenon from laws and initial conditions. The mechanistic approach also recasts a number of important issues in philosophy of science such as the format in which explanations may be represented and how scientists reason about mechanisms. Whereas it is natural to represent laws in propositions or equations, the challenge in understanding a mechanism is to grasp the types of parts and operations it employs and how these are organized. A diagram is often most effective in conveying this information. Logical deduction provides a way of showing that an instance (e.g., the movement of a given pendulum) follows from a law. But to see how a variety of operations generate a phenomenon, one must simulate the operation of the mechanism, either in a physical model, in one's head, or in a computational simulation.[2] Two issues on which mechanistic philosophy of science offers a different perspective than more traditional nomological ones will be of particular concern here.

One involves reduction. In one sense familiar to scientists, mechanistic research is inherently reductionistic insofar as it appeals to entities and events at a lower level of organization—parts and operations—to explain the behavior of the whole mechanism. But in another sense, it is inherently holistic in that mechanisms must be organized and appropriately

---

[1] The relationship between phenomena and mechanisms allows for multiple perspectives. Typically mechanistic philosophers of science have emphasized the delineation of the phenomenon as the reference point and construed the mechanism as consisting of whatever parts and operations are responsible for it. The same parts and operations may, on this view, be constituents of multiple mechanisms depending on what phenomenon one is explaining. However, sometimes researchers focus on the parts and operations constituting a mechanism, and then identify additional phenomena for which they are responsible. While there are contexts in which scientists operate in the latter manner, one reason to prefer the former approach is that mechanisms often do not come well delineated in nature (as they are more likely to in a human-made machine). Especially in biological systems, different components can be recruited into a functioning system when one activity is required and into another functioning system on another occasion. A similar problem arises with phenomena as well (biological phenomena are often integrated with each other) but in the context of a given research endeavor, the approach of the investigators provides a basis for characterizing the phenomenon for which explanation is sought.

[2] Recognizing that scientists sometimes model mechanisms via computational simulations is compatible with but does not entail that the modeled activities of the mechanism and/or their parts are themselves computational processes. A mechanistic approach in psychology is thus compatible with, but not necessarily committed to, a computational theory of mind which metaphysically characterizes psychological processes as computational or information processing activities.

situated in their environment in order to function. Accordingly, mechanistic accounts are multi-level and do not privilege the lowest level.

The second issue involves discovery. Advocates of the traditional nomological accounts of science generally eschewed discovery, focusing instead on justification. The reason was that while there was reason to hope that one could articulate processes of logical inference through which theories were justified (or at least could be falsified), there seemed to be little prospect of a logical specification of how one discovers laws (but see Langley, Simon, Bradshaw, & Zytkow, 1987; Thagard, 1988). Philosophers focused on mechanistic explanation have, however, identified some important aspects of how mechanisms are discovered. Of special relevance to us will be how reasoning can progress from initial proposals for localizing operations in the parts of a mechanism, to much richer accounts involving many parts that are organized in complex ways.

3. The Challenge of Delineating Emotional Phenomena

In referring to what is to be explained as a *phenomenon*, and considering examples such as biological respiration or encoding of long-term memories, it may seem as though phenomena are immediately apparent and obvious to anyone who looks for them. But in fact a great deal of scientific, often experimental, work is typically required to delineate phenomena in a manner that renders them appropriate as the target of explanation. Even something that we take to be basic to the life of mammals, circulation of blood, was not apparent even to physicians prior to the work of Harvey. Until then, physicians held that there were two distinct types of blood, arterial and venous, both of which flowed out from the heart to the peripheral tissues. Although Harvey himself was not able to show how the circuit from the arteries to the veins was completed, since capillaries were too small to be identified with the tools he had available, he presented what now seems like overwhelming evidence that the blood must circulate. First, he showed that the valves where the veins connect to the heart are oriented in the wrong direction for blood to flow out through the veins. Second, he calculated that the amount of blood that would have to be created from food on a regular basis if it were not recirculated vastly exceeded dietary intake. Finally, he showed that if veins were restricted in a limb, the limb would soon bloat from the amount of arterial blood reaching it. But none of this was compelling to many of his contemporaries, who insisted on the traditional account of the phenomenon in terms of two types of blood both traversing outward from the heart.

It is not just obstinacy that renders phenomena challenging to identify. Rather, the activities that get characterized as phenomena and then explained typically depend on relations between entities that have to be identified. For example, Darwin had to notice how traits of organisms (such as the beaks of the finches he observed in the Galapagos) made them fit for the conditions in their environment. Previous theorists, such as Paley, whom he much admired, focused on the complexity of the characteristics of organisms, but only once Darwin identified traits as adaptive did it make sense to seek an explanation in a process such as natural selection. Moreover, sometimes quantification is required to delineate phenomena. It was widely known by anyone who thought about it that objects left unsupported near the surface of the earth drop. Yet it took Galileo to show that objects near the surface of the earth always fall 32 feet/sec^2, a phenomenon that would later be explained by Newton's law of universal gravitation.

With this as background, let us turn to the phenomena we expect psychology to explain. Here several challenges arise. First, the English word "emotion," like many other words in our mental vocabulary, likely refers to a host of different psychological phenomena that may share a family resemblance but also exhibit great variability. This extends not just to our folk talk of emotions, but also to the use of the term "emotion" by scientists. One finds emotions presented as conscious feelings, cognitive states (e.g., appraisals), peripheral physiological responses (e.g., heart rate, respiration), and behavioral tendencies or overt behavior (e.g., facial, bodily, verbal); sometimes they are treated as collections of these mental and bodily phenomena (for reviews, see e.g., Gendron & Barrett, 2009; Gross & Barrett, 2011; Prinz, 2004, ch. 1). Such differences are of fundamental importance from the perspective of mechanistic explanation since mechanisms are identified according to the phenomenon they explain.

The challenge is apparent if one considers the question of where the mechanism responsible for emotion is located. For mechanists, a phenomenon and its mechanism are not separate entities, with the mechanism *causing* the phenomenon. Rather, the activity of the mechanism is said to *constitute* or *realize* the phenomenon of interest. So if bodily events are literal parts of emotions, then clearly the mechanisms of emotion cannot be purely in the brain. But if emotions are disembodied mental states that can have bodily effects (which are not themselves literal parts of the emotions), and if the mechanisms of the mind are indeed to be found in the brain, then mechanisms of emotion will be neural mechanisms.

Second, as the constructivists emphasize (e.g., Barrett, 2006), emotional phenomena are highly variable. Episodes of fear, even when elicited in the same person by the same stimulus, show considerable variability in behavioral responses, autonomic activity, and subjective experience. The goal of a mechanistic account is not to construct a separate mechanism for each episode, but to identify a mechanism that accounts for regular patterns. The concept of a phenomenon, as presented by Bogen and Woodward (1988) as the appropriate target of explanation, involves regularities that can be exhibited repeatedly. This is not inconsistent with recognizing that each instance differs on a variety of dimensions, but it does require finding regularity even in variable instances.

The relation between phenomenon, regularity, and mechanism is more complex than these remarks suggest. Not every well-established quantitative regularity constitutes a phenomenon to be explained by a separate mechanism. Rather, there is often a reciprocal relation between identifying a phenomenon and identifying a mechanism—as an account of the mechanism develops, it may enable us to account for a variety of regularities, which may then be grouped together as aspects of the same phenomenon. It is at this point that the constructionist contention that emotion is constructed from many of the same operations that figure in other cognitive activities becomes a serious issue for scientists.

If the connections between components lead scientists to treat them as components of one mechanism, and we view all the activities to which this mechanism gives rise as aspects of the same phenomenon, then we have a recipe for an extremely problematic holism at least across biology and psychology.[3] One of the consequences of identifying the molecular basis of many biological activities has been the discovery that many are shared and that there are important pathways by which the activities affect one another. For example, there is increasingly compelling evidence that NAD, a central component of metabolic activities, affects circadian rhythms, which have traditionally be assigned to different mechanisms. The consequence would

[3] Russell (2009) seems to support the idea that psychological constructionism will involve a holism across psychology: "…the traditional assumption is that a 'theory of emotion' will differ from a theory of cognition, or behaviour, or conation. My claim, in contrast, is that any theory that explains all cases called emotion will be close to the whole of psychology, a theory that of course will not be limited to emotion but will extend to all psychological processes" (p. 1268).

seem to be that we should identify one phenomenon involving basic metabolism and circadian timekeeping. This same line of reasoning will quickly lead us to identifying only one biological phenomenon and one mechanism (the whole organism, plus perhaps its environment).

If the goal is explanation, we must resist such thoroughgoing holism. In practice mechanistic scientists differentiate phenomena and avoid holism while acknowledging points of interaction between the responsible mechanisms. Thus they differentiate circadian time keeping from the myriad of related phenomena to which it and the mechanisms responsible for it are linked. Scientists in part make the decision as to what to count as one phenomenon and hence one mechanism pragmatically in terms of what they think they can productively integrate into one account. But if the decisions as to how to delineate phenomena are to generate scientific understanding, these decisions must also aspire to track actual distinctions in nature. The mechanisms identified with different phenomena will not be totally separable from each other, but each must have sufficient organizational coherence to make it appropriate to treat it as one mechanism responsible for one phenomenon (see the discussion of clusters in small-world networks below in section 5—different clusters may constitute mechanisms for specific phenomena but their behavior may be modified in real-time through the long-range connections to other clusters).

The challenge this poses to delineating psychological phenomena and mechanisms is not simply resolved by adopting the mechanistic perspective, but it does give some guidance. The question of whether emotional phenomena and cognitive phenomena should be distinguished or whether some other differentiation is needed will in part turn on how successful researchers are at developing mechanistic accounts that distinguish them. It may be that one can maintain a productive distinction between emotion and some other mental phenomena and identify mechanisms that can be distinguished even if they are always interconnected. Likewise, it may be possible to distinguish particular emotion types from each other in the same manner.

Here we do not take a stand on whether the constructionists are right in challenging the traditional distinction between emotion and other mental phenomena or between different emotions as distinct phenomena, but have only tried to show some of what is at stake from a mechanistic perspective. One thing that we hope to have made clear, though, is that the question is not whether there are totally distinct phenomena/mechanisms or totally integrated, but whether the phenomena and the mechanisms responsible for them are sufficiently differentiable that they can be treated separately or not. In the final section we return to the question of delineating the phenomena and whether the constructionists' description of psychological phenomena so transforms traditional conceptions that they *reconstitute* the phenomena themselves.

4. Decomposition and Localization

Central to the mechanist approach to explanation is the idea that components parts perform different operations than the mechanism as a whole, and are organized and orchestrated to produce the phenomenon of interest. Thus, once one has delineated a phenomenon and linked it with an appropriate mechanism, a key task is to decompose that mechanism. Two types of decomposition are relevant to mechanistic explanation: decomposition into structural parts (*structural decomposition*) and into functional operations (*functional decomposition*). Decomposing in either way can be an extremely challenging endeavor—the appropriate ways of dividing a mechanism into parts or component operations is not just obvious to all those who look. It requires developing both the appropriate conceptual framework and adequate experimental tools. On the conceptual side, researchers must develop the concepts for particular sorts of parts or operations and the vocabulary for identifying them.[4] For example, in distinguishing regions of the brain, Brodmann (1908) had to construct the concept of brain areas distinguished by their cytoarchitecture. Likewise, in decomposing fermentation into component operations, biochemists needed concepts for the relevant different types of chemical reactions (oxidation, phosphorylation, etc.). In addition, researchers needed the requisite experimental tools for determining when these concepts are satisfied. Moreover, a given group of researchers may only have the tools for one type of decomposition, and indeed at a given time only one set of tools may be available in the scientific community. Brodmann aspired to distinguish brain regions that would perform distinct mental operations, but he had no tools for picking out such operations; and biochemists in the 1930s sought to link reactions with specific enzymes, but they could not determine the chemical constitution of enzymes (or even determine whether they were macromolecules or colloids). The ultimate goal, though, is to localize component operations in component parts of the mechanism, but the understanding of how to do so may only develop at the end of a sustained inquiry.

The process of decomposing a mechanism takes place over time. The importance of functional decomposition is often ignored in the early stages of a mechanistic research project, where it is common to attempt to localize a whole phenomenon in a single part of the mechanism. For example, Broca (1861) localized the faculty of articulate speech in the left prefrontal region that bears his name and Buchner (1897) localized fermentation in a single enzyme he named *zymase*. Bechtel and Richardson (1993/2010) refer to this as "direct" or "simple" localization. Sometimes such localization is correct, but even when it is, it simply relocates the mechanism in a component of what was taken to be the mechanism. As such it doesn't directly constitute any explanatory advance for it offers no account of how the phenomenon is produced. Yet, it can play an important heuristic role in opening a line of empirical inquiry.

One can view what Barrett (2006) calls "natural kind" approaches to emotion as engaged in the project of directly localizing individual emotions in brain regions. She identifies basic emotion theories but also some appraisal theories as exemplars of the approach. What distinguishes these approaches is that: (a) emotions are a discrete, basic category of psychological phenomena distinct from perception, memory, reason, etc.; (b) fear, anger, happiness, sadness, etc., are discrete types of emotions, irreducible to more basic psychological phenomena; and (c) the discrete emotion types are localized in distinct areas of the brain, whether these are single brain regions or discrete networks. Lindquist et al. (2012) call this the "locationist approach." What is characteristic of this approach is that the decompositions proposed in steps (a) and (b) are in terms of phenomena, not the operations realizing them. The reliance on such *phenomenal decompositions* (Bechtel, 2008) is extremely common in psychology. It was the strategy of the faculty psychologists, and provided the basis for Gall's program of phrenology. Although faculty psychology and phrenology are widely denigrated today, the approach to decomposing psychological processes is evident in the division of psychological processes into categories such as memory, perception, and language, of memory into long-term versus short-term memory, and long-term memory itself into episodic and semantic memory. With the advent of fMRI, these have often provided the psychological activities to be localized in the brain.

---

[4] Mechanists agree with Barrett (2009) that "the categories at each level of the scientific ontology capture something different from what their component parts capture, and each must be described in its own terms and with its own vocabulary" (p. 332).

The constructionists' contention is that the delineation of natural kinds of emotion "has outlived its scientific value and now presents a major obstacle to understanding what emotions are and how they work" (Barrett, 2006, p. 29). We agree that this approach needs to be supplanted by one that engages in decomposition of emotional phenomena into their operative parts and operations, and that these will not be properly characterized in the language of emotion, memory, or perception, but in more basic vocabulary that identifies the general operations that constitute the mind. But an important question is how to move forward to the promised land. Here it is important to recognize that even when it turns out to be incorrect, direct localization can be heuristically productive in the development of a science. One way it is so is by promoting the discovery of evidence that the part in which the phenomenon was localized is not the only part involved. Then researchers start to ask the question of what the various parts are doing, and this prompts finally decomposition into the contributions of the different parts. It can also bring the realization that different phenomena rely on the same part, prompting the question of what operation might be involved in these various phenomena. Lindquist et al.'s (2012) meta-analysis of the neuroimaging literature on human emotion is one recent example of this kind of research framed in opposition to direct localization accounts of emotion (see also Oosterwijk et al., this volume).

While it is easy to identify examples where psychologists and cognitive neuroscientists have settled for simple localizations, neuroimaging (as well as the other main tool of cognitive neuroscience, analysis of lesions) has itself provided the evidence of the need to decompose psychological phenomena and not just localize them in the brain. In one of the first PET studies of cognitive processes, Petersen, Fox, Posner, Mintun, and Raichle (1988) found that the verb-generate task elicited increased blood flow in left dorsolateral prefrontal cortex, the anterior cingulate, and the cerebellum. Although much of the interest focused on the left dorsolateral prefrontal cortex, the researchers themselves were puzzled by the activity in the anterior cingulate and interested in what it and the cerebellum contributed. Moreover, the fact that nearly the same region of left dorsolateral prefrontal cortex showed increased activation in a semantic memory task led Gabrieli, Poldrack, and Desmond (1998) to conclude that "operations may be the same whether they are considered in the context of language, working memory, episodic memory, or implicit memory. The left prefrontal cortex thus serves as a crossroads between meaning in language and memory" (p. 912). The current generation of neuroimaging studies that focus on networks of brain regions that exhibit synchronized behavior in the resting state and are then recruited in specific cognitive tasks (Mantini, Perrucci, Del Gratta, Romani, & Corbetta, 2007; Fox & Raichle, 2007; van den Heuvel, Mandl, Kahn, & Pol, 2009; Sporns, 2010; Moussa, Steen, Laurienti, & Hayasaka, 2012) is the outgrowth of the earlier attempts to localize cognitive activities in single brain regions. Barrett and Satpute (2013) describe this transition in neuroscientific research on emotion and social cognition (see also Lindquist & Barrett, 2012).

Attempts to directly localize phenomena in single components of an organism have often revealed the limitations of such an approach. In anticipating this, the vitalists (a diverse group of biologists from the 17th through the 19th century who rejected mechanist approaches to the activities of living organisms and often appealed to a vital force to explain these phenomena) were often correct in their objections to early proposals of mechanistic biologists. But whereas they lacked a positive research program of their own, an important virtue of the mechanist project was that it generated hypotheses for which falsifying evidence—e.g., that other parts were involved—could emerge, forcing researchers to address the question of what operations they performed and to develop a functional decomposition of the phenomenon into multiple operations.

As difficult as it often is to identify the parts of a mechanism, identifying operations can be even more challenging. The problem is that when the mechanism is functioning, the component operations interact smoothly so that the result of one operation feeds directly into others. Barrett (2009) captures this point when she comments “The contents of a psychological state reveal nothing about the processes that realize it, in much the same way that a loaf of bread does not reveal the ingredients that constitute it” (p. 330; however, we would prefer a better analogy since a loaf of bread is not an active mechanism). Often the clues as to the component operations must come from sources other than direct interventions on the mechanism such as lesioning or stimulating its parts. For biochemistry, for example, they were provided by organic chemists, who discovered that organic compounds are composed of groups such as hydroxyl or phosphate groups and that operations might consist of adding or removing whole groups or modifying them in predicable ways. With these operations in mind researchers could attempt to organize the different potential intermediates they identified when the reaction was interrupted at various points into comprehensible sequences of chemical reactions.

Only in a few areas has it been possible to propose psychological operations out of which mechanisms responsible for psychological phenomena can be constructed. In the case of sensory systems it has been possible in some cases to identify the types of stimuli to which particular neurons are responsive—center-surround contrasts in the retina and LGN, oriented and moving edges in V1, illusory contours in V2, shapes in V4—and to ask then what sorts of information processing operations would enable the downstream region to identify a feature such as an illusory contour from representations of edges. Likewise, the discovery of different cells that respond to information relevant to navigation in the medial temporal lobe (place cells, head-direction cells, grid cells, boundary-vector cells) has provided researchers a basis on which to hypothesize about the information processing operations that enable navigation by path integration. In many regions such as the hippocampus and surrounding areas the neural architecture provides both clues and evidence for hypotheses about the operations being performed.

It is noteworthy that the examples in the last paragraph appeal to neuroscience to identify the information processing operations underlying the cognitive activities of whole agents (e.g., seeing or navigating). In contrast constructionists often focus on *psychological* primitives and appeal to examples such as categorizing and core affect as the basic operations (Barrett, 2009; Russell, 2003). There are, however, virtues in appealing to neuroscience in identifying operations. The first is that the operations to which constructionists appeal are often surprisingly close to the phenomena to be explained. Categorizing is an activity agents (even single-celled organisms such as bacteria) perform. Moreover, it is an activity that may be performed in different ways, relying on different operations. We worry that despite constructionists’ concern to supplant folk psychological operations with ones from which they are generated, they may have stayed too close to the folk level. This is not surprising since, short of insight from outside, theorists have few places to go to identify the component activities. However, we are sympathetic to the view that in appealing directly to the operations of the brain is to look to too low a level of organization to develop explanations of psychological phenomena. Contrary to some philosophical reductionists (Bickle, 2003), the goal in explanation is not to descend to the lowest possible level, but to the level at which operations are found out of which a mechanism can generate the phenomenon of interest. For many higher cognitive phenomena we need more complex operations than those that might be performed by single-neurons. The virtue, though, of operations specified in terms of the neural architecture is that they are independently grounded and not just the projections of phenomena onto the mechanism that is to explain them. Even if, for many purposes, they are operations at too low a level of organization to figure directly in the explanation of cognitive phenomena, they do afford the prospect of identifying higher-level structures whose operations may be appropriate to the phenomena of interest. These might include neural columns or brain regions. Barrett (2009) is critical of the suggestion of appealing to columns on the grounds that they have many projections beyond themselves. This, however, is not a reason to reject them as the structures whose operations support cognition since we expect the parts of a mechanism to interact with each other to generate phenomena.

A second concern is with the prospect of differentiating psychological and neural operations. Barrett's (2009) proposal of psychological primitives seems to impose an unnecessary and unhelpful boundary between psychology and neuroscience. In mechanistic research there is always a level of decomposition that particular researchers, with the tools and techniques available to them, typically do not go beyond. Often this presents no problem—to explain the phenomenon they seek to explain, it is enough to show the parts and operations out of which it is constructed. They or someone else may then become interested in the components themselves, treat them as phenomena and seek to explain them mechanistically. This does not mean denying the higher-level construction! In going further down, these researchers are asking different questions and appealing to different types of parts and operations to explain the component mechanism. For purposes of explanation one may establish a level of parts and operations one takes as primitive, but there is nothing magical or necessarily psychological about the operations in terms of which one decomposes a phenomenon—especially given the lack of consensus over the nature of the "mental" or "psychological." Researchers trained in and employing techniques of psychology may in fact identify and invoke the same parts and operations as someone trained in and employing the tools of neuroscience. One of the goals of inquiry in neuroscience is to identify the operations (information processing activities) various brain regions perform. As these are characterized, they may equally be employed by psychologists as the operations they invoke in psychological explanations.

Identifying the parts and operations into which to decompose psychological phenomena remains a major challenge for psychology. It, however, is a challenge all mechanistic sciences face. Moreover, even if the early attempts have stuck too close to folk categories and so have not identified the parts and operations required to explain psychological phenomena, there is no reason for despair. In particular, collaboration with neuroscience may help generate higher-level psychological operations in terms of which the desired explanations can be developed. Barrett's recent work (Barrett & Satpute, 2013; Lindquist & Barrett, 2012) appears to be taking these lessons to heart, by exploring intrinsic functional neural networks as the "core systems" realizing higher-level, domain-general psychological operations out of which emotions and other agent-level psychological phenomena are constructed. In addition to using neuroscience as a guide to discovering higher-level psychological operations, Barrett's newer terminology of "core systems" better suggests that these domain-general psychological ingredients are not essentially "primitive," and can themselves be structurally and functionally decomposed into parts and operations.

5. Decomposition and Recomposition

As we have stressed, the goal of decomposition is to identify parts and operations that contribute to the phenomenon but do not individually exhibit the phenomenon. The explanatory gain in explanation comes from understanding how components that don't themselves exhibit the phenomenon can nonetheless exhibit it when working collaboratively. The challenge is to understand how such collaboration can yield the phenomenon. Again, in pressing this point we concur with Barrett (2009) that "there must be an explicit accounting (i.e., a mapping) of how categories at each level relate to one another" (p. 332). To emphasize how this endeavor complements decomposition, we refer to it as *recomposition*. Even though, just as in the case of decomposition, researchers may not literally put the parts back together, they must at least do so conceptually.

Recomposition, however, turns out to be considerably harder than decomposition, as we have only the beginnings of an understanding of the consequences of different modes of organization. When humans think about combining operations, they tend to do so sequentially, as in an assembly line. The same holds true of scientists trying to recompose mechanisms. When operations are envisaged as sequentially ordered, researchers can recompose them in their minds, imagining the results of executing the first operation, then imagining the second applied to it, etc. Thus, one can imagine DNA being transcribed into various RNAs, them being transported to the cytoplasm, where tRNA binds with an amino acid and then with a locus on a mRNA, the amino acid forming a bond with the last amino acid to have been added to the polypeptide sequence, etc. Likewise, one can imagine visual perception resulting from processing first by center-surround cells in the retina and LGN, then by edge detection cells in V1, etc.

This sequential conception of a mechanism has been enshrined in Machamer, Darden, and Craver's (2000) characterization of a mechanism as operating "from start or set-up to finish or termination conditions." It, moreover, is a factor behind the frequent view that mechanisms are impoverished in what they can accomplish and are not up to the challenges of producing cognition. For example, Barrett and colleagues' (Barrett, 2009; Barrett, Wilson-Mendenhall, & Barsalou, in press) critique of the "machine" metaphor for the mind often targets this linear organization as one its main inadequacies.

This sequential conception of mechanism is, however, both inadequate to most biological mechanisms and vastly underestimates what can be accomplished when mechanisms are organized in a more complex manner. But the failure to look beyond sequentially organized mechanisms is not surprising, as it turns out to be very difficult for humans to understand how systems organized in a nonsequential manner will behave. One of the simplest departures from sequential order is negative feedback in which a process later in an imagined sequence affects the execution of an earlier one. Its first known use in human design was in Ktesibios' 3rd century BC design for a water clock, but it had to be continually reinvented over the next two thousand years when designers sought to maintain a system at a target level. In the early 20th century it became more widely used and was then heralded by the cyberneticists as a fundamental principle found in both biological and social systems.

Negative feedback is sufficiently simple that most people can grasp how it works in their imagination: as a furnace heats the air, it causes metal to expand, opening a switch, causing the furnace to stop. When the temperature drops sufficiently, the metal contracts, the switch is closed, and the furnace again generates heat. But while this use of negative feedback to control a mechanism so as to maintain or approach a target state can be intuitively understood, another important effect of negative feedback—that it can generate sustained oscillations—is even today not widely recognized except by engineers who view it as a nuisance they seek to minimize. The reason is that by mentally simulating a system with negative feedback one cannot ascertain whether it will continue to oscillate or will dampen. Whether it will do so depends on whether the operations are appropriately nonlinear, a feature we cannot capture in our mental rehearsal. Confronted with nonsequential mechanisms with nonlinear operations, researchers adopt a different strategy to recomposing the mechanism: they mathematically or computationally model it by representing the operations in terms of differential equations and employ them in a computational model to simulate the functioning of the mechanism.

The nonsequential, nonlinear nature of biological mechanisms not only serves to make them more challenging for scientists to understand, but also provides them with resources to behave in a manner very different from the way machines we have designed typically behave. Moreover, these behaviors are crucial to the ability of living organisms to maintain themselves as distinct entities, recruiting matter and energy from their environment and continually building and repairing themselves. Biological organisms are in this sense *autonomous* (Ruiz-Mirazo, Peretó, & Moreno, 2004). Autonomous systems are not just reactive systems; they are endogenously active. Endogenously activity is already manifested in the simplest life forms—single-celled organisms without neurons or brains, including prokaryotes that lack differentiated organelles. Even such "simple" organisms carry out complex processing of information about their environment and use it to modulate their endogenous activity including motor behavior, a task for which animals employ a nervous system. This endogenously active feature of organisms carries over to brains—a signature of this endogenous activity is the oscillatory behavior that can be identified at a wide range of frequencies and shows up in the synchronized activity across networks in the brain (Abrahamsen & Bechtel, 2011).

In rejecting the machine model for the mind, Barrett (2009) opts instead for the recipe model: mental phenomena are constructions in accord with recipes (see also Barrett et al., in press). But the entries in a recipe book are static entities, and only when utilized by a skilled chef do they figure in the generation of the rich variety of foods that we associate with them. It is the chef, not the recipes, that is critical to this process, and he or she does so as an endogenously active system engaged in ongoing, variable, interactions with an environment. We suggest that it is more appropriate to maintain the mechanistic perspective but recognize that the mechanisms involved in the brain involve nonlinear operations organized in complex ways that we are only beginning to identify and understand.[5] One feature of brain organization that researchers are beginning to understand is that at different levels of organization it exhibits a small-world architecture in which units (neurons, columns, brain regions) are primarily connected to their neighbors but with a few long-range connections (Watts & Strogratz, 1998). The high clustering of local units allows clusters to specialize in specific types of information processing while the short path length enables these clusters to coordinate their activities as appropriate. The local clusters in such an organization can be viewed as specialized information processing mechanisms whose operations cognitive science and neuroscience are trying to identify. The mental activities and behaviors of cognitive agents result from the coordinated operation of many of these clusters as a person as a whole engages his or her environment. This is the conception of

[5] Barrett et al. (in press) introduce two other supposedly "more apt" non-machine-based models for the relation between the mind and brain: "molecules that are constructed of atoms" and "chamber music emerging from the interplay of instruments." Unfortunately neither quite captures both the nonlinear organization and endogenous activity we emphasize in the dynamic mechanistic perspective.

the functioning of the brain that the metaphor of construction is advanced to capture, and we suggest that when the complex dynamic behavior that is possible in mechanisms is recognized, their account is best served by the framework of mechanistic explanation. Cunningham, Dunfield, and Stillman (this volume) and Thagard and Schröder (this volume) offer examples of mechanistic accounts of emotion emphasizing a complex, dynamic organization of neural parts and operations.

6. Reconstituting the Phenomena

Mechanistic research is directed at explaining phenomena that have already been delineated. But as we have already noted, delineating phenomena is an ongoing activity and it goes on even as explanations are being developed. Researchers might devote considerable effort to explaining a phenomenon only to realize that it has quite a different character than initially assumed. In one particularly dramatic example, biologists spent over a century trying to explain the phenomenon of animal heat. For much of that time, heat was presumed to be an energy resource. In the process of trying to explain animal heat, researchers came to understand many of the important operations that go into basic metabolism. But eventually heat came to be recognized as a waste product—once free energy was turned into heat, it was no longer available to do useful work in living organisms. Only that energy which was captured in chemical bonds, such as those of ATP, is available for work. Bechtel and Richardson (1993/2010) refer to such revisions in the characterization of the phenomenon as *reconstituting the phenomenon*.

Does the fact that the operations performed in the brain are not themselves appropriately characterized in terms of the folk categories of emotion, thought, memory, self-knowledge, etc., entail that mental phenomena themselves should be reconstituted? Eliminativists in the philosophy of mind have long contended that these and other categories of folk psychology should be eliminated from science when they are found incapable of being reduced to the best available theory of how the brain works (Churchland, 1981). But the assumption that these categories track internal operations in the mind or brain may misrepresent their role in a fundamental way. Whole persons use folk psychological concepts to characterize themselves and other people, and these characterizations can have significant psychological and behavioral consequences. For example, if we can conceptualize what our goals are, we may be better able to identify and avoid acting on desires which are incompatible with them. And when we attribute beliefs and desires to others, we may be able to anticipate their behavior and better coordinate our own actions with theirs. These uses of folk idioms have consequences for behavior and so make them an appropriate focus of psychological research even if they do not track the basic operations of our mental mechanisms.

Barrett (2012; this volume) and other psychological constructionists take a position along these lines about the psychological phenomena identified by folk categories. They treat emotion, cognition, etc., as genuine psychological phenomena, rather than endorsing their complete elimination from the science of psychology. But they endorse an alternative characterization of these psychological phenomena. Rather than being primitive categories of the mind, these phenomena are constructed from more basic psychological ingredients, many of which are shared across psychological phenomena the folk distinguish as emotional or cognitive. In the case of the different emotion categories, constructionists deny that they are distinguished by unique psychological ingredients (and the mechanisms that realize them). Rather, emotional episodes of anger, fear, etc., are ones in which core affect and other interoceptive and exteroceptive sensory information are conceptualized by the agent as instantiating an emotion category. According to psychological constructionists, the psychological and behavioral components of emotional episodes do not exhibit the tight correlations assumed by natural kind theories of emotion. These psychological ingredients are involved in many phenomena that the folk distinguish as emotional and non-emotional. Emotional episodes in ourselves and others occur, for constructionists, when sets of these loosely correlated ingredients are conceptualized by an observer as an instance of an emotion category. This categorization process is a real phenomenon that can have psychological, behavior, and social consequences, which are to be studied by psychologists. It is just that these emotional phenomena do not exist independent of agents' conceptualization process. Accordingly, Barrett (2012, 2009) describes emotions as "observer-dependent" phenomena, distinct from "observer-independent" phenomena whose existence does not depend on being recognized or conceptualized by minded beings.

We accept this metaphysical distinction between entities which depend on minded beings for their existence and those which do not, but question its being given any special status. Many metaphysical categories, especially ones in biology, may be in this way relational—i.e., the metaphysical nature of the entities depend on their relations with other entities—without depending on the existence of *minds*. Niches are defined with respect to the traits of the organisms that inhabit them, and chemical attractants for bacteria are defined with respect to the organism's chemical receptors. Further, we reject that the distinction between ontologically-dependent and independent entities entails a sharp *epistemological* divide between folk and scientific concepts, or between psychological and neuroscientific concepts. Concepts across both divides are heavily dependent upon the theories and interests of human cognizers, and may not carve reality (whether it is mind-dependent or independent) at its joints.[6] As humans, we have as much scientific interest developing categories that facilitate understanding of ourselves and directing our actions as we do those that characterize the physical, chemical, and biological processes that constitute us.

Making these points, however, does not call into question the psychological constructionist view that some psychological phenomena are constructions rather than psychological primitives, and dependent upon the psychological concepts of the people under study. We agree that "Psychology must explain the existence of cognition and emotion because they are part of the world that we (in the Western hemisphere) live in" (Barrett, 2009, p. 330). This makes the study of emotion part of the larger interdisciplinary research program studying people's folk psychological or "mindreading" abilities. It is a debated issue in the literature whether attributing mental states to oneself involves a distinctive sort of introspective access to our own mental states, or engages the same non-introspective methods we use to attribute mental states to others (e.g., Carruthers, 2009). The latter position can easily be accommodated by constructivists, since on this view the process of attributing mental states to others and oneself can be seen as a construction based on a variety of available evidence. But so can a version of the introspectionist account if the objects of introspection are restricted to the sensory states constructionists include in their accounts of emotion categorization. While much of the mindreading literature assumes that Western folk psychological concepts do reflect the actual architecture of the mind, some emphasize that the importance of mindreading depends on its

[6] Although it may seem strange to say a metaphysical category that is constructed via the conceptualization of minded beings is not accurately conceptualized, this is not contradictory. It is possible for agents to use one set of concepts in unreflectively constructing the category, but use another set in their reflective thought about that metaphysical domain. See also Barrett (2012, p. 422).

usefulness in mediating social interaction and regulating one's own behavior, rather than on its accurately describing our inner workings (e.g., Godfrey-Smith, 2004, 2005). Such accounts have clear affinities to psychological constructionism.

We do not here take a stand on whether psychological constructionism is correct in its call to reconstitute psychological phenomena. From the perspective of mechanistic explanation, we emphasize that such revision in how the phenomena are described can be the result of research at the level of the whole—as when constructionists call into question the proposed correlations between the various behavioral measures of emotion—or the indirect result of research at lower levels investigating the mechanisms responsible for these phenomena. In the latter case findings such as that basic emotion categories fail to be localized distinct brain regions can motivate inquiry into whether these categories should or should not be retained at the level of characterizing persons. In the case of psychological constructionism, its central contention is that psychology has been restricted by its adherence to Western folk psychological concepts, especially in the case of emotion. While many folk concepts have been eliminated as science has progressed, folk psychological concepts have the unique status of being used by us humans to characterize ourselves, and thus partially constitute and influence our own minds. Folk concepts of emotion thus have an important place in psychological research, whether or not they accurately reflect the nature of our mental mechanisms.

7. Conclusion

In this chapter we have highlighted several ways that the new mechanistic philosophy of science can shed light on psychological constructionism. The mechanistic approach offers a multilevel framework for conducting research and explaining psychological phenomena in terms of more basic components. Research into the psychological construction of mental phenomena can benefit from mechanism's emphasis on both decomposition of a mechanism into its basic parts and operations, and recomposing those parts into an organized whole that is environmentally situated. In this context, we can understand the move from natural kind theories of psychological phenomena to the psychological construction approach as the transition from simple localization to more complex and dynamic accounts of the neural mechanisms responsible for these psychological phenomena, and potentially to a reconceptualization or reconstitution of mental phenomena themselves.

1. Is your work consistent with the hypothesis that emotions are psychological events constructed from more basic ingredients? Is it plausible that there are key ingredients from which emotions emerge? Are these ingredients specific to emotion or are they general ingredients of the mind? Which, if any, are specific to humans?

The mechanistic approach to philosophy of science we describe is not itself a theory of psychological phenomena such as emotion. Rather, it offers a framework for how discovery and explanation operate in sciences such as biology and psychology. Such a framework is thus not itself committed to the psychological constructionist hypothesis that emotions are constructions from more basic ingredients. But what philosophers have noticed about the processes of discovery and explanation for other biological phenomena do point in favor of psychological constructionism about psychological phenomena. It is common for scientists to initially identify a single entity as the mechanism responsible for a phenomenon, describing this part using the vocabulary originally reserved for the phenomenon of interest. Subsequent research, however, often reveals that this simple, direct localization of the phenomenon is false; they then find the mechanism consists of a host of interacting parts, each of which performs activities that are distinct from the phenomenon to be explained. This general pattern of finding complex, multi-part mechanisms is at a minimum consistent with, and can be seen to provide indirect support for, the psychological constructionist approach. In the case of psychological phenomena, mechanists agree with the constructionists that psychology needs to work to discover the more basic operations of the mind that are realized in regions and networks of the brain. Which if any of these basic psychological operations are unique to humans, and whether human-specific psychological phenomena involve the unique combination of psychological operations present in some nonhuman animals, are questions that cannot be answered a priori simply by adopting a mechanistic approach.

2. Which processes bring these ingredients together in the construction of an emotion? Which combinations are emotions and which are not (and how do we know)?

The mechanistic approach is agnostic about how scientists should conceptualize emotions, and what particular parts, operations, and organization should be identified in a mechanistic explanation of emotional phenomena. Like in many other scientific inquires, everyday categories are a starting point for delineating mental phenomena; and if neural mechanisms do not respect these folk categories, conceptual revision of the type proposed by psychological constructionism could be necessary. A host of neuroscientific research does support the constructionists' view that the brain does not respect the traditional division between "mental faculties" such as cognition and emotion. But researchers are only at the stage of proposing initial hypotheses about the basic psychological operations that will go into mechanistic accounts of mental phenomena such as emotion, so it is not yet clear how drastically we may need to revise our folk psychological categories. We do, however, believe a mechanistic approach offers some lessons about how this inquiry should proceed. First, we emphasize the importance of looking to neuroscience to guide such hypotheses, even if neuroscientific research will at times be at too low a level to immediately impact our mechanistic understanding of agents' psychological phenomena. Second, we should expect our emotion mechanisms' parts and operations to be endogenously active and organized in a complex, nonlinear fashion.

3. How important is variability (across instances within an emotion category, and in the categories that exist across cultures)? Is this variance epiphenomenal or a thing to be explained? To the extent that it makes sense, it would be desirable to address issues like universality and evolution.

Mechanistic accounts treat regularities as the phenomena to be explained in terms of underlying mechanisms. This is not inconsistent with recognizing that each instance of a phenomenon differs on a variety of dimensions, but it does require finding regularity even in variable instances. If there is too much variability across instances of a proposed category of phenomena, or great regularity across instances of different categories, this points toward a need to redescribe or reconstitute the phenomena to the explained—possibly dissolving boundaries between phenomena considered distinct, or even more radically recategorizing the phenomena to be explained. Discovering that phenomena have overlapping mechanisms is one major reason for uniting phenomena previously considered distinct, and being able to identify distinct mechanisms is a reason to continue to differentiate phenomena, even if those phenomena are always in fact interconnected.

4. What, in your view, constitutes evidence from your own work to unique support a psychological construction to emotion (i.e., what evidence would convince you)?

Evidence that specific emotion categories are not grounded in discrete, localized neural mechanisms would speak in favor of the psychological constructionist's general approach of finding more basic psychological operations from which emotional episodes are constituted. This would follow the general pattern of researchers finding that biological phenomena often are explained by dynamical mechanisms consisting of many parts organized in complex ways. Further, evidence that regularities at the level of the actions and physiological responses whole agents do not track folk emotion categories would suggest revising these categories, and the relation between emotion and other mental phenomena. Mechanistic research often involves such reconstitution of the phenomena to be explained.

References

Abrahamsen, A., & Bechtel, W. (2011). From reactive to endogenously active dynamical conceptions of the brain. In K. Plaisance & T. Reydon (Eds.), *Philosophy of behavioral biology* (pp. 329-366). New York: Springer.

Barrett, L. F. (2006). Are Emotions Natural Kinds? *Perspectives on Psychological Science, 1*, 28-58.

Barrett, L. F. (2009). The Future of Psychology: Connecting Mind to Brain. *Perspectives on Psychological Science, 4*, 326-339.

Barrett, L. F. (2012). Emotions are real. *Emotion, 12*, 413-429.

Barrett, L. F., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*.

Barrett, L. F., Wilson-Mendenhall, C. D., & Barsalou, L. W. (in press). A psychological construction account of emotion regulation and dysregulation: The role of situated conceptualizations. In J. J. Gross (Ed.), *The handbook of emotion regulation* (2nd ed.). New York: Guildford.

Bechtel, W. (2008). *Mental mechanisms*. London: Routledge.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 421-441.

Bechtel, W., & Richardson, R. C. (1993/2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Cambridge, MA: MIT Press. 1993 edition published by Princeton University Press.

Bickle, J. (2003). *Philosophy and neuroscience: A ruthlessly reductive account*. Dordrecht: Kluwer.

Bogen, J., & Woodward, J. (1988). Saving the phenomena. *Philosophical Review, 97*, 303-352.

Broca, P. (1861). Remarques sur le siége de la faculté du langage articulé, suivies d'une observation d'aphemie (perte de la parole). *Bulletin de la Société Anatomique, 6*, 343-357.

Brodmann, K. (1908). Beiträge zur histologischen Lokalisation der Grosshirnrinde. VI Mitteilung: Dei Cortexgliederung des Menschen. *Journal für Psychologie und Neurologie, 10*, 231-246.

Buchner, E. (1897). Alkoholische Gärung ohne Hefezellen (Vorläufige Mittheilung). *Berichte der deutschen chemischen Gesellschaft, 30*, 117-124.

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences, 32*, 121-138.

Churchland, P. M. (1981). Eliminative materialism and propositional attitudes. *The Journal of Philosophy, 78*, 67-90.

Craver, C. F. (2007). *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 explanation. In F. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 117-144). Cambridge, MA: MIT Press.

Darden, L. (2006). *Reasoning in biological discoveries: Essays on mechanisms, interfield relations, and anomaly resolution*. Cambridge: Cambridge University Press.

Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience, 8*, 700-711.

Gabrieli, J. D. E., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences, USA, 95*, 906-913.

Gendron, M., & Barrett, L. F. (2009). Reconstructing the Past: A Century of Ideas About Emotion in Psychology. *Emotion Review, 1*, 316-339.

Godfrey-Smith, P. (2004). On folk psychology and mental representation. In H. Clapin, P. Staines & P. Slezak (Eds.), *Representation in mind: New approaches to mental representation* (pp. 147-162). Amsterdam: Elsevier.

Godfrey-Smith, P. (2005). Folk psychology as a model. *The Philosopher's Imprint, 5*, 1-16.

Gross, J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review, 3*, 8-16.

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.

Hempel, C. G. (1966). *Philosophy of natural science.* Englewood Cliffs, NJ:: Prentice-Hall.

Langley, P., Simon, H. A., Bradshaw, G. L., & Zytkow, J. M. (1987). *Scientific discovery: Computational explorations of the creative process*. Cambridge: MIT Press.

Lindquist, K. A., & Barrett, L. F. (2012). A functional architecture of the human brain: emerging insights from the science of emotion. *Trends in Cognitive Sciences, 16*, 533-540.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences, 35*, 121-143.

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

Mantini, D., Perrucci, M. G., Del Gratta, C., Romani, G. L., & Corbetta, M. (2007). Electrophysiological signatures of resting state networks in the human brain. *Proceedings of the National Academy of Sciences, 104*, 13170-13175.

Moussa, M. N., Steen, M. R., Laurienti, P. J., & Hayasaka, S. (2012). Consistency of network modules in resting-state FMRI connectome data. *PLoS ONE, 7*, e44428.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature, 331*, 585-588.

Prinz, J. J. (2004). *Gut reactions: A perceptual theory of emotion*. New York: 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.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*, 145-172.

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition & Emotion, 23*, 1259-1283.

Sporns, O. (2010). *Networks of the brain*. Cambridge, MA: MIT Press.

Thagard, P. (1988). *Computational philosophy of science*. Cambridge, MA: MIT Press/Bradford Books.

van den Heuvel, M. P., Mandl, R. C. W., Kahn, R. S., & Pol, H. E. H. (2009). Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Human Brain Mapping, 30*, 3127-3141.

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