http://www.jstor.org

A Role for Representation in Cognitive Neurobiology

Jacqueline Anne Sullivan[†‡]

What role does the concept of representation play in the contexts of experimentation and explanation in cognitive neurobiology? In this article, a distinction is drawn between minimal and substantive roles for representation. It is argued by appeal to a case study that representation currently plays a role in cognitive neurobiology somewhere in between minimal and substantive and that this is problematic given the ultimate explanatory goals of cognitive neurobiological research. It is suggested that what is needed is for representation to instead play a substantive role.

1. Introduction. The idea that learning and memory require experience-dependent changes in synaptic strength is a cornerstone of modern cognitive neurobiology. The corresponding idea, that learning and memory involve changes in internal *representations* in the brain, however, is often not explicitly acknowledged among cognitive neurobiologists (see Dudai 2007). This is primarily because, as I aim to demonstrate in this article, representations have a peculiar status in cognitive neurobiology. In one sense, cognitive neurobiologists are interested in what organisms trained in experimental learning paradigms actually learn: what the changes in internal representations in the brain actually are. However, their failure to make reference to changes in internal representations as either the causes of observable changes in behavior or the effects of cellular and

molecular activities indicates that they are not interested in knowing precisely what organisms learn, only that learning does or does not occur (see Dudai 2007; Machamer 2009). As I will show, the ambiguous role of representation in cognitive neurobiology is problematic in light of the purported global explanatory aims of research in this area.

I begin, in section 2, by defining representation and distinguishing perceptual representations from memory representations. In section 3, I outline the basic structure of experiments in cognitive neurobiology and appeal to two descriptive accounts of explanation that shed light on how representations may function in cognitive neurobiological explanations. I use these accounts as the basis for introducing a conceptual distinction between what I refer to as minimal and substantive roles for representation. Appealing to this distinction, in section 4, I undertake an analysis of a case study from the recent cognitive neurobiological literature. I demonstrate that representations currently play something more than a minimal role but something less than a substantive role in cognitive neurobiology. I conclude by considering some of the implications of representations playing this role.

2. Memories as Representations. An organism's thoughts, experiences, and memories are said to be about objects and events in the world, and insofar as they are about them, they *represent* them. These aspects of an organism's psychological life are commonly referred to in philosophy and cognitive science as "mental representations." In neuroscience, however, they are referred to as "neural representations." On the most basic definition, a *mental representation* is a mental state that, insofar as it is about some thing in the world, has that thing as its object or content. In contrast, a *neural representation* is a neural state, namely, the firing rate or firing pattern of an individual neuron, which is thought to carry "information" about the object or event that triggered that firing pattern (see Dretske 1981/1999; deCharms and Zador 2000). When the firing pattern of a neuron is significantly correlated with the presence of some feature of a stimulus that an organism is currently experiencing, that pattern is said to represent that feature.

Two assumptions are made about neural representations. The first is that the relationship between a neural representation and whatever it is about is causal: a neuron will only exhibit a significant increase in its firing rate above baseline in response to that stimulus feature that causes it to fire. Whatever stimulus feature causes it to fire in this way, it represents (see deCharms and Zador 2000). Support for this assumption comes from cognitive neurophysiological investigations of predominantly sensory neurons. For example, neurons in auditory cortex fire in response to auditory stimuli, neurons in insular cortex to taste stimuli, and so forth for other

sensory systems. The second assumption is that neural representations are *veridical*. Our sensory systems convey to us what is going on out there in the world; there is a 1 : 1 mapping between objects and events in the world and the representation of those objects and events in our brains (see Akins [1996] for a critique of the veridicality assumption).

One need only consult the theoretical literature in cognitive neurobiology to see how these assumptions shape thinking about memory. Memories may be understood either as mental representations of objects and events that are ascribed to a whole organism and its mind or as representations that are distributed across that set of neurons in the brain that were sufficiently and jointly activated when the organism experienced some stimulus or set of stimuli in the world (see Squire and Kandel 2000; Dudai 2007). It is widely accepted throughout neuroscience that persistent internal representations are the result of experience-dependent changes in synaptic strength between and among those neurons activated during that organism's experience. Insofar as these neurons are taken to comprise the memory trace or engram, they "serve as the physical representation of memory" (Han et al. 2009, 1492). Dudai, for example, claims that "nervous systems . . . have evolved to encode representations of facets of the world, and make computations over these representations while keeping the distinctiveness of things represented (parsing)" as well as "the relationship among things represented (structure)" (2007, 14–15). This is in part why he goes so far as to claim that memories, construed as experience-dependent veridical representations of objects and events in the world, constitute *knowledge*.

3. A Minimal versus a Substantive Role for Representation. These general theoretical assumptions about memories prompt a question about whether and how notions of representation shape experimentation and inform explanation in different areas of the neurosciences of learning and memory. I am particularly interested in what role, if any, appeals to representation play in cognitive neurobiology (see Bechtel 2001, 2007; Eliasmith 2009, for the role in cognitive neurophysiology). To get a handle on this, it is worthwhile to first say something very basic about the kinds of experiments in which I am interested.

In cognitive neurobiological learning experiments, animals are trained in experimental learning paradigms, and pharmacological or genetic intervention techniques are used to alter the activity of a molecule before, during, or after training to assess the effects on behavioral performance and molecular activity. An experimental learning paradigm is used to operationally define a form of memory. The basic idea is that an investigator first selects stimuli to be presented to an organism and determines how they ought to be arranged to produce a type of memory in the

laboratory and detect when it occurs. She then identifies the response variables that are to be measured before, during, and after training. Finally, she identifies the measurable changes in the response variables that must be observed in order to claim that an organism has a memory. These measurable changes in behavioral response variables constitute data that are taken as indicative of a phenomenon—the form of memory that the investigator is interested in explaining something about (see Sullivan 2009). So, my question is, what role, if any, does the notion of representation play in these experiments and in the explanations that these experiments are used to generate?

Two accounts of explanation in the philosophical literature are useful to consider insofar as they offer conflicting responses to this question. The self-proclaimed "radical-reductionist" John Bickle (2006), for example, may be interpreted as claiming that notions of representation play at best a heuristic role in cognitive neurobiology. On his account, cognitive neurobiology is what he dubs "reductive-in-practice," insofar as practitioners import learning paradigms from experimental psychology and use them to operationally define types of memory in terms of observable changes in behavior. They then use pharmacological interventions and genetic techniques to determine the causes of those changes. Once a form of memory has been thus redefined, the referent of that category is no longer in the brain (see Machamer 2009). Thus, the causal explanations that investigators generate go directly from cells and molecules to behavior. Internal representations are neither the targets of these explanations nor causes that are appealed to in order to explain observable changes in behavior.

As evidential support for his claims, Bickle (2006) describes research that purportedly establishes a role for a cyclic-AMP response element binding protein (CREB) in social recognition memory. As I have criticized Bickle's treatment of this case previously (Sullivan 2009), I will not consider it here. I will instead apply Bickle's model to a case study below, in the same context in which I will evaluate what the model of neuroscientific explanation offered by Carl Craver (2007) may be taken to imply about the role of representation in cognitive neurobiology.

For Craver, explanations in neuroscience are mechanistic. He defines a mechanism as "a set of entities and activities organized such that they exhibit the phenomenon to be explained" (2007, 5). Mechanisms are multilevel insofar as the entities and activities in which they engage span ontic levels from, for example, the molecular to the systems levels. It is worthwhile to appeal to a case study to understand what the model implies about the role of representation in neuroscientific explanations. Since Craver uses the case of spatial memory to illustrate the basic structure

of mechanistic explanations, I will investigate his treatment of that case here.

In a classic example of a spatial memory task, the Morris water maze (MWM), a rat is placed into a large circular pool of opaque water that contains a hidden platform resting just below the water's surface (see Morris 1981, 1984). The maze is positioned in a room containing a set of distal cues. On the first trial, a normal rat discovers the platform randomly; across subsequent trials, the directionality of its swim path to the platform improves, and the time it takes to locate it ("escape latency") decreases. This is taken to indicate that the rat has spatial memory—memory for where the hidden platform is located in relation to the distal room cues.

A number of neuropsychological and neurophysiological studies support the idea that the hippocampus plays an important role in spatial memory (see Kentros 2006; Bird and Burgess 2008). All subfields of the rodent hippocampus contain what are referred to as "place cells," which may exhibit a high rate of firing in response to specific locations in the rodent's environment (see O'Keefe and Dostrovsky 1971). Place cells are said to "represent" these locations and are taken to play a role in the formation of "cognitive or spatial maps" (Tolman 1948; O'Keefe and Nadel 1978; Redish 1999). Spatial maps are thought to be representations that are distributed across place cells, with each place cell contributing that aspect of the environment it represents to the map.

Richard Morris and colleagues (1986) demonstrated that when the N-methyl-D-aspartate (NMDA) receptor antagonist aminophosphonovaleric acid (AP5/APV) is microinfused into the hippocampus of rats before training them in the MWM, their performance in locating the hidden platform fails to improve across trials. Specifically, the escape latency to arrive at the hidden platform never decreases, and their swim paths are random and indirect. This is taken to indicate that "AP5 impairs place . . . learning" in AP5-treated rats and, thus, that NMDA-receptor activation is necessary for spatial memory (Morris et al. 1986, 776).

Bickle would likely claim that this study does establish that NMDA-receptor activation in the hippocampus is a cause of spatial memory defined operationally in terms of observable changes in behavior and that this causal explanation is all that Morris and colleagues intend to provide. Bickle is correct insofar as there is no appeal to "internal representations" in the hippocampus, or the lack thereof, to explain the failure of rats treated with APV to navigate the maze (Morris et al. 1986). Morris and colleagues instead identify the cellular (i.e., place cells) and molecular (i.e., NMDA receptor activation) mechanisms that contribute to spatial memory operationally defined in terms of escape latency and directionality of swim path.

However, given what I have said about representation in section 2, memories are internal representations. This means that memory formation in the MWM may be understood to involve two different kinds of events: (1) changes in synaptic strength and (2) changes in internal representations. If I understand Bickle correctly, Morris and colleagues may be regarded as interested in the former, not the latter. This may be correct and attributable in part to the fact that the basis for Morris and colleagues' study was to determine whether molecules involved in the purported neural correlate of learning and memory, long-term potentiation (LTP), were also involved in a form of hippocampal learning.

Craver's account of mechanistic explanation, however, provides an alternative framework for thinking about the role of representation in the spatial memory case. He claims the mechanistic explanation of spatial memory is multilevel. He identifies four different levels (although he admits there may be more), namely, "the level of spatial memory," "the level of spatial map formation," "the cellular-electrophysiological level," and "the molecular level" (Craver 2007, 165–66). The level of spatial memory corresponds to a rodent navigating a maze (166, fig. 5.I), which is the phenomenon to be explained. The level of spatial map formation corresponds to what purportedly happens in the hippocampus when the rat is trained in the maze. The cellular-electrophysiological level is the level at which changes in synaptic strength are formed across hippocampal place cells. At the molecular level, NMDA receptor activation in place cells results in a calcium signal sufficient to initiate those cellular and molecular changes that result in changes in synaptic strength.

Craver separates the level of spatial memory, which is an organism-level phenomenon, from that of spatial map formation, which is a systems-level (hippocampal) phenomenon. Spatial memory is indicated by the improvement in a rat's maze-navigating behavior across trials in the MWM. However, these changes in behavior are not identical to the changes in spatial representations purportedly taking place in the rat's hippocampus. In fact, on Craver's model, it is more appropriate to say that the formation of the hippocampal spatial map is a contributing factor in the observable changes in the rat's behavior.

Craver's model, therefore, allows for two possible ways in which representations may play a role in experimentation and explanation in cognitive neurobiology. First, changes in internal representations (e.g., spatial map formation) may be understood as phenomena to be explained by appeal to changes in the activity of cells (LTP) and molecules (NMDA receptor activation). Yet, hippocampal neurons may engage in both synaptic plasticity activities and representational activities simultaneously; one set of activities is not identical to the other. A complete explanation of a phenomenon like spatial map formation would demand an expla-

nation of both kinds of activities and how they contribute to spatial memory formation. Second, changes in observable behaviors—a rat's gradual improvement in locating a hidden platform across multiple trials in the MWM (i.e., "spatial memory")—may be understood as the phenomena one of the causes of which is spatial map formation. Here, both the representational and synaptic plasticity activities of hippocampal neurons may be appealed to as causes of the observable changes in behavior. Again, these causes are not identical. The basic idea, then, is that changes in internal representations may be either the targets of multilevel mechanistic explanations or the causes to which an investigator appeals in order to explain downstream consequences including observable changes in behavior.

Despite the fact that Craver's model allows for the possibility of representations playing these distinct roles in cognitive neurobiological explanations, Bickle is correct insofar as the experimental methods that cognitive neurobiologists use exclusively track behavioral changes. These changes are then used as surrogates for internal representational changes. Although the MWM is designed to produce a specific form of memory (i.e., spatial memory) in a brain area where spatial representations are supposedly forged (i.e., the hippocampus), "what" rats learn in the context of the MWM—that is, the changes in internal representations—cannot be adequately established exclusively by reference to changes in their behavior or differences in the behavior of APV rats compared to normal controls (see D'Hooge and De Deyn 2001). In other words, it is not clear what the representational changes that occur in the hippocampus during spatial map formation actually are or whether there are representational changes in other areas of the brain as well. Just because a rat can find a hidden platform in a maze does not reveal how it did it. As evidential support for this point, investigators continue to struggle with "what" rats trained in the MWM actually learn and which brain structures are actually involved (see Morris and Frey 1997; Redish 1999; D'Hooge and De Deyn 2001).

In light of the aforementioned considerations, I want to draw a distinction between representations playing a minimal role versus a substantive role in the contexts of experimentation and explanation in cognitive neurobiology. These roles correspond roughly to how I have characterized Bickle's and Craver's descriptive accounts of explanation. In cases in which representation plays a minimal role in the context of experimentation, an investigator operationally defines a form of learning in terms of observable changes in behavior, and those changes constitute the targets of explanation. He then intervenes in the activity of cells and molecules, determines the effects on behavior, and explains those effects in terms of the cellular and molecular interventions. No appeals to internal

representations occur in the context of explanation, either as causes of observable changes in behavior or as the phenomena to be explained. This is Carnap's model for converting psychology into a physical science, with insights pertaining to the functional differentiation of the brain from cognitive neuroscience thrown in.

In contrast, in cases in which representations play a substantive role in cognitive neurobiology, changes in internal representations may be construed as the phenomena that one aims to produce and then intervene in or disrupt by undertaking a molecular intervention. In this case, representations would likely be the explanatory targets of a mechanistic explanation. However, it may also be that an investigator aims to explain some change in behavior that she has produced in the laboratory via training an animal in a learning paradigm. So, she may appeal to what the organism has learned—the changes in internal representations—in order to explain the observable changes in behavior. In these experiments, she will be concerned with precisely what an organism is learning when it is trained in an experimental learning paradigm and with what the individual cells are representing. In this latter case, the question is not "does this rat have spatial memory" but rather, "what does that memory amount to?" In posing this question, an investigator has shifted from representation playing a minimal role in experimentation to it playing a substantive role, and if she appeals to what has been learned in order to explain the observable changes in behavior, then representation will be playing a substantive role in the context of explanation, too.

Applying the distinction to the case of spatial memory and the MWM, I take representation in this study (Morris et al. 1986) to have played only a minimal role in shaping the experimental design and no role in the explanation, except that the changes in behavior were attributable to cellular activity in a brain area thought to process spatial information. I think that representations have actually played something between a minimal and a substantive role historically across all research studies in cognitive neurobiology, which is why they have an uncertain status in cognitive neurobiology. Yet, making this case requires additional examples. To this end, in the next section, I apply the conceptual distinction between minimal and substantive roles for representation to a more recent case study from cognitive neurobiology, in order to demonstrate the role that representation is playing there.

4. Case Study. I want to consider two interrelated research studies that take seriously the idea that "ensembles of neurons" constitute "the physical representation of memory (the memory trace)" (Han et al. 2009, 1492). I want to begin by describing the basic features of the experimental learn-

ing paradigm used in both studies before investigating how the two studies differ.

The studies were undertaken in Sheena Josselyn's laboratory (Toronto) and published in the journal *Science*. The second study (Han et al. 2009) is a follow-up to the first (Han et al. 2007). In both experiments, the form of learning under study is an auditory fear-conditioning paradigm. The experimental subjects are mice (mutant or transgenic, specified below). The main experimental learning paradigm and protocol used consisted of a 2-minute exposure to the training context (a fear-conditioning chamber) followed by presentation of an auditory stimulus (28,000 Hz, 30 secs., 85 dB tone) that coterminated with a 2-second high-intensity (75 mA) shock. As neurons in the lateral amygdala (LA) are thought to be the primary location in which the memory is formed, Han and colleagues investigated and intervened in molecular activity in this brain area. They operationally defined "fear memory" as "the percentage of time mice spent freezing during subsequent tone presentation 24 hours" after termination of training (2007, 457).

In the first study, they conducted a series of experiments in order to determine what role cyclic-AMP response element binding protein (CREB) plays in recruiting neurons in the LA into a "fear memory trace" (Han et al. 2007). They had shown previously that cells in the mouse LA exhibit activated CREB after training in the learning paradigm, and they wanted to know whether CREB activation selectively determined in which population of cells a fear memory would be stored.

In one experiment, Han and colleagues sought to determine whether mice who are deficient in CREB function throughout the brain, and who show correlative deficits in learning, could actually learn in an auditory fear-conditioning task when a $CREB^{WT}$ vector was microinfused directly into the LA. Their rationale was that the increase in CREB function resulting from this microinfusion would make cells otherwise unavailable for formation of the fear memory trace available. So, they took two groups of mice, a wild-type control group and a CREB-deficient group, and "microinject[ed]" into a portion of neurons in the lateral amygdala a "replication-defective herpes simplex viral vector expressing endogenous CREB ($CREB^{WT}$)" or "dominant negative ($CREB^{S133}$) respectively, fused with green fluorescent protein" (Han et al. 2007, 457). They then trained both groups in the auditory fear-conditioning paradigm. They found that the $CREB^{WT}$ vector microinfused mice "froze at levels similar to those" of wild-type controls, and when they removed the brains to visualize which LA neurons were infected with the $CREB^{WT}$ vector, a significant number of cells were infected. Han and colleagues concluded that microinfusion of the vector "completely rescued the memory impairment" (58) and that the infected LA neurons were causally responsible for the memory trace.

In a second experiment, Han and colleagues wanted to know whether neurons in the LA that were infected with the CREB^{WT} vector were more likely to be "represented in the fear memory trace" (Han et al. 2007, 458). To assess this, they took one group of wild-type mice and microfused the CREB^{WT} vector into a subset of LA neurons and another group of wild-type mice and infused them with a control vector. They trained mice in both groups in the fear-conditioning paradigm and, 5 minutes after determining freezing in response to the tone, removed the brains in order to determine whether activity-regulated cytoskeleton-associated protein (Arc) RNA, which is used as a molecular indicator of both the recent activity of a neuron and its recruitment into a memory trace, was present in the nuclei of the LA neurons containing elevated CREB. They found that these neurons were significantly more Arc$^+$ than both the neurons that did not receive the CREB^{WT} vector and the neurons of wild-type mice infected with only the control vector. They took these results to indicate that in animals that are successfully trained in a fear-conditioning paradigm, there is a high probability that those LA neurons that are CREB active will be recruited into the fear memory trace.

In the second research study, Han and colleagues (2009) hypothesized that if neurons with relatively higher CREB function are recruited during the formation of the fear memory trace, then to ablate them would result in a disruption of the behavioral expression of the fear memory, namely, freezing to the tone in the testing condition. To this end, they "used transgenic mice in which cell death may be induced in a temporally and spatially restricted manner" by microinfusion of diphtheria toxin (Han et al. 2009, 1493). They then trained transgenic mice and controls in the auditory fear-conditioning paradigm, selectively ablated those neurons containing higher CREB activity in the experimental group, and then determined the effect of such ablation on behavior. They found that selectively ablating these neurons before testing for the fear memory resulted in an elimination of freezing to tone, the behavioral expression of the memory trace. They concluded that they had eliminated the memory trace, or at least a substantial proportion of it.

Is representation playing a minimal or a substantive role in these experiments? For several reasons, I take it to be playing a role somewhere in-between. First, Han and colleagues operationally define fear memory (i.e., the representation) as freezing to the tone 24 hours after training in the auditory fear-conditioning paradigm. However, they do not take mice freezing to the tone to be identical to the fear memory; the freezing is rather what they take to indicate that such a memory is present in the brain. This immediately suggests that they take the representation to be separate from its behavioral expression.

For representation to play a minimal role in this case, two criteria would

have to be satisfied. First, representation would play a heuristic role in the experimental design. I think it does play this role since the choice of learning paradigm is informed by findings from cognitive neuroscience pertaining to the role of the amygdala in fear learning. Yet, according to my definition, if representation was playing a minimal role, representations would be redefined in terms of observable behaviors, and the observable behaviors would be the targets of causal explanations. This is clearly not what is happening in this case since Han and colleagues take the memory representation to be separate from its behavioral expression. This opens the possibility that the representation can play a role as either a target of explanation or a cause of the observable changes in behavior. Thus, representation is not playing a minimal role in this case. But is it then playing a substantive role?

To answer this question, it is necessary to consider whether changes in representation are regarded as separate from changes in molecular activity and whether such changes play a role in the context of explanation. Han and colleagues operationally define the memory trace (i.e., the representation) as consisting of those LA neurons that overexpress CREB and are Arc^+ after the fear memory test. So, they identify the memory trace exclusively by appeal to molecular activity (i.e., CREB expression), which is related primarily to plastic changes in LA neurons that are thought to subserve memory formation (i.e., representational changes). This alone is insufficient for establishing that these neurons constitute the memory trace. They then ablate those neurons that overexpress CREB before the fear memory test and identify these neurons as constituting the memory trace primarily because freezing to tone, the behavioral expression of the memory, subsequently disappears. They use this correlation as a basis to establish a causal relation between the memory trace and observable changes in behavior. Thus, representations do not play a substantive role in this case. Instead of representations being redefined in terms of observable changes in behavior, they are redefined as observable changes in molecular activities. Han and colleagues then attempt to establish a causal link exclusively between these molecular activities and the behavioral expression of the fear memory. This strategy essentially skirts dealing with what is being represented in the brain.

However, Han and colleagues do recognize two aspects that determine which LA neurons get recruited into a fear memory trace, which suggests that they do acknowledge a role for representation that veers in the direction of being more substantive. First, as they take themselves to have demonstrated, the neurons must overexpress CREB. Second, and more interestingly, they claim that the neurons must also "receive the necessary tone and shock information" such that they are "correctly 'wired' to form a fear memory" (Han et al. 2009, 13). Although they claim that being

"anatomically well positioned" is sufficient for receiving the relevant information, they clearly rely on molecular activity to infer that there is informational content in the LA—that the tone and shock are being represented. If representation were playing a substantive role, the aim may be to establish the relationship between the plasticity-related molecular activities and the informational activities of LA neurons. As I have demonstrated, this is clearly not the aim of this study.

5. Conclusion. The cognitive neurobiologist Yadin Dudai recently has suggested that "reductionistic-oriented experts" should be "relieve[d] from the chagrin of trying to figure out the relevance of findings in synapses and individual cells to the actual output of the behaving organism" (2007, 14). However, one conclusion of this article is that it will be a difficult feat to merge the results of one area of science that takes seriously "what" individual neurons represent with those of another that for the most part strips neurons of their representational capacities or at least emphasizes only those capacities relevant for changes in synaptic strength. The only way to overcome this situation is for cognitive neurobiologists to actively instigate a shift from representation playing a minimal to it playing a more substantive role in experimental practice and explanation. What this shift will amount to in practice will have to be left for another paper.

REFERENCES

Akins, Kathleen. 1996. "Of Sensory Systems and the 'Aboutness' of Mental States." *Journal of Philosophy* 93 (7): 337–72.
Bechtel, William. 2001. "Representations: From Neural Systems to Cognitive Systems." In *Philosophy and the Neurosciences: A Reader*, ed. William Bechtel, Pete Mandik, Jennifer Mundale, and Robert S. Stufflebeam, 332–48. Oxford: Blackwell.
———. 2007. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. New York: Erlbaum.
Bickle, John. 2006. "Reducing Mind to Molecular Pathways: Explicating the Reductionism Implicit in Current Cellular and Molecular Neuroscience." *Synthese* 151:411–34.
Bird, Chris, and Neil Burgess. 2008. "The Hippocampus and Memory: Insights from Spatial Processing." *Nature Reviews Neuroscience* 9 (3): 182–94.
Craver, Carl F. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Oxford University Press.
deCharms, Christopher R., and Anthony Zador. 2000. "Neural Representation and the Cortical Code." *Annual Review of Neuroscience* 23:613–47.
D'Hooge, Rudi, and Peter De Deyn. 2001. "Applications of the Morris Water Maze in the Study of Learning and Memory." *Brain Research Reviews* 36:60–90.
Dretske, Fred I. 1981/1999. *Knowledge and the Flow of Information*. Stanford, CA: Center for the Study of Language and Information.
Dudai, Yadin. 2007. "Memory: It's All about Representations." In *Science of Memory: Concepts*, ed. Henry L. Roediger III, Yadin Dudai, and Susan M. Fitzpatrick, 13–16. Oxford: Oxford University Press.
Eliasmith, Chris. 2009. "Neurocomputational Models: Theory, Application, Philosophical Consequences." In *The Oxford Handbook of Philosophy and Neuroscience*, ed. John Bickle, 346–69. Oxford: Oxford University Press.
Han, Jin-Hee, Steven A. Kushner, Adelaide P. Yiu, Christy J. Cole, Anna Maynia, Robert

A. Brown, Rachael L. Neve, John F. Guzowski, Alcino J. Silva, and Sheena A. Josselyn. 2007. "Neuronal Competition and Selection during Memory Formation." *Science* 316: 457–60, (Supporting Online Material) 1–20.

Han, Jin-Hee, Steven A. Kushner, Adelaide P. Yiu, Hwa-Lin (Liz) Hsiang, Thorsten Buch, Ari Waisman, Bruno Bontempi, Rachael L. Neve, Paul W. Frankland, and Sheena A. Josselyn. 2009. "Selective Erasure of a Fear Memory" *Science* 323:1492–96, (Supporting Online Material) 1–25.

Kentros, C. 2006. "Hippocampal Place Cells: The 'Where' of Episodic Memory?" *Hippocampus* 16 (9): 743–54.

Machamer, Peter K. 2009. "Neuroscience, Learning and the Return to Behaviorism." In *The Oxford Handbook of Philosophy and Neuroscience*, ed. John Bickle, 166–78. Oxford: Oxford University Press.

Morris, Richard. 1981. "Spatial Localization Does Not Require the Presence of Local Cues." *Learning and Motivation* 12:239–60.

———. 1984. "Developments of a Water-Maze Procedure for Studying Spatial Learning in the Rat." *Journal of Neuroscience Methods* 11:47–60.

Morris, Richard G., Elizabeth Anderson, Gary Lynch, and Michel Baudry. 1986. "Selective Impairment of Learning and Blockade of Long-Term Potentiation by an N-Methyl-D-Aspartate Receptor Antagonist, AP5." *Nature* 319:774–77.

Morris, Richard G., and Uwe Frey. 1997. "Hippocampal Synaptic Plasticity: Role in Spatial Learning or the Automatic Recording of Attended Experience?" *Philosophical Transactions of the Royal Society of London* B 352 (1360):1489–1503.

O'Keefe, John, and John Dostrovsky. 1971. "The Hippocampus as a Spatial Map: Preliminary Evidence from Unit Activity in the Freely-Moving Rat." *Brain Research* 34 (1): 171–75.

O'Keefe, John, and Lynn Nadel. 1978. *The Hippocampus as a Cognitive Map*. Oxford: Oxford University Press.

Redish, A. David. 1999. *Beyond the Cognitive Map: From Place Cells to Episodic Memory*. Cambridge, MA: MIT Press.

Squire, Larry R., and Eric R. Kandel. 2000. *Memory: From Mind to Molecules*. New York: Holt.

Sullivan, Jacqueline A. 2009. "The Multiplicity of Experimental Protocols: A Challenge to Reductionist and Non-reductionist Models of the Unity of Neuroscience." *Synthese* 167:511–39.

Tolman, Edward C. 1948. "Cognitive Maps in Rats and Men." *Psychological Review* 55 (4): 189–208.