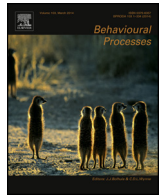




Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc



Mechanism, function, and computation in neural systems

Nathan Insel^{a,b,*}, Paul W. Frankland^{b,c,d,e,*}

^a Department of Biological Sciences, University of Toronto Scarborough, Toronto, ON, Canada M1C1A4

^b Department of Psychology, University of Toronto, Toronto, ON, Canada M5S3GM

^c Program in Neurosciences and Mental Health, The Hospital for Sick Children, Toronto, ON, Canada M6G1X8

^d Institute of Medical Science, University of Toronto, Toronto, ON, Canada M5S1A8

^e Department of Physiology, University of Toronto, Toronto, ON, Canada M5S1A8

ARTICLE INFO

Article history:

Received 9 June 2014

Received in revised form 21 August 2014

Accepted 10 September 2014

Available online xxx

Keywords:

Mechanism

Circuits

Levels of analysis

Hippocampus

Prefrontal cortex

ABSTRACT

What constitutes a “mechanism” of behavior? In this tribute to Jerry Hogan we examine how questions of behavioral mechanism can be reframed as causes and consequences of neural circuit activity. Drawing from our work on the hippocampus and the medial prefrontal cortex we discuss the inherent difficulties of characterizing the behavioral functions of circuits that are many synapses away from sensory reception and motor/visceral expression. We briefly review the advantages of reframing a region's functions according to its *computations*, while also distinguishing those computations from the algorithms by which they are achieved. As an example of how these ideas can be applied, we discuss why the hippocampus and medial prefrontal cortex may have overlapping roles in memory expression in spite of being very different circuits. The present analysis draws inspiration from David Marr, whose framework for describing neural systems can be compared with Aristotle's “causes.” This article is part of a Special Issue entitled: In Honor of Jerry Hogan.

© 2014 Elsevier B.V. All rights reserved.

Contents

1. Introduction	00
2. Challenges in identifying mechanisms of behavior	00
3. Isolating the challenge: Identifying the computations of neural systems	00
4. Translating behavioral functions into computations: Example of the hippocampus and prefrontal cortex	00
5. Computations of the hippocampus and mPFC in memory expression: Same or different?	00
6. Summary	00
Acknowledgement	00
References	00

1. Introduction

A fundamental question in science is “Why?” But the question is ambiguous, as explanations can take any number of reference frames. Aristotle (350 BCE) had a scheme for classifying explanations that has endured history with a vitality customary amongst ancient Greeks. He contended that any given event can be explained in terms of (1) the physical substance of the system (material

cause), (2) the way that the system or event is organized (formal cause), (3) the external forces that put the event in motion (efficient cause), and (4) the purpose, goal, or function of the event (final cause). These divisions are not perfect. They are not entirely unambiguous and may fail to capture some of the most critical features of complex systems (meanwhile raising further metaphysical questions about the nature of contingency). But the framework provides a basic road map that, in spite of its simplicity, is often neglected.

It is easy to see how different types of explanations might become confused in the contemporary scientific setting, and even more so within the specific field of animal psychology. The field has expanded exponentially with the incorporation of neuroscience and genetics techniques, but merging with biology has also meant an intermingling of very different scientific goals. At the same

* Corresponding authors at: The Hospital for Sick Children, 555 University Avenue, Toronto, ON, Canada M5G 1X8. Tel.: +1 416 813 7654x1823; fax: +1 416 813 7717.

E-mail addresses: nathan.insel@utoronto.ca (N. Insel), paul.frankland@sickkids.ca (P.W. Frankland).

time, increased competition for “high impact” real estate offers incentives to interpret findings in as far-reaching ways as possible. Fortunately, the scientific community is a self-correcting system, one that depends on unique individuals like Jerry Hogan who, like Aristotle, has endured with a classical vitality. Jerry reintroduced Aristotle’s “four causes” into ethology as a way of offering a common language for different types of inquiries (Hogan, 1994). Jerry has argued, for example, that learning about the behavioral role of specific brain regions (part of the material cause, or “matter”) does not necessarily explain what stimuli motivate a behavior (the efficient causes; or simply “causation”). Similarly, he has emphasized the fundamental distinction between knowing the variables that motivate a behavior and understanding the implications of the behavior’s consequences. Jerry’s ideas are a reminder of the limits over which scientific findings can be interpreted, but they also provide a system for classifying and thereby promoting the diversity of scientific inquiry.

Our own research has focused less on animal behavior *per se* and more on the neural processes that support it. One way to put this is to say that we study “material causes” (or, arguably, “formal causes”) of behavior, but at a certain point it becomes useful to depart from Aristotle’s language. We can say instead that we study behavioral *mechanisms*; however, as we discuss in more detail below, this term is complicated by ambiguity. By definition, a mechanism is a description of a phenomenon in terms of the actions and interactions of a system’s parts—in our case, the phenomenon of interest is the behavioral expression of memory. But establishing a mechanism requires addressing issues that do not have obvious or unanimously agreed-upon solutions, such as the right balance between detail, accuracy, and generalizability, the best ways to subdivide a system into parts, and defining which components are necessary in a complex system with inherent redundancies. It is probably safer to say that we study neural circuits, specifically a few rather deeply-buried circuits in the mammalian forebrain: the hippocampus and medial prefrontal cortex (mPFC). And what we want to know is: what causal interactions make these circuits so important for memory expression?

The purpose of the present article is to step back from what we are doing and think hard about what we do. What, specifically, are the challenges of explaining behavior according to its mechanisms? What angle of approach in studying brain regions like the hippocampus and mPFC will be most informative for a generalizable understanding of their respective roles in behavior? Our examination of these questions outlines the importance of articulating and testing theories about the computations a brain system performs. A theory of computation does not have to be a computer-based simulation or written in the form of differential equations, although we discuss some of the ways that mathematical and computer-based models can be valuable. Theories should also be consistent with, but not fully dependent upon, the region’s proposed cognitive functions and the underlying algorithms by which the region is thought to work. After discussing these issues in the abstract sense, we see how they have played-out in studies of the hippocampus and mPFC. Although definitions will be provided throughout the text, we also offer a glossary (Table 1) that clarifies how we use some of the most relevant terms.

2. Challenges in identifying mechanisms of behavior

Ecosystems can be described as interacting communities, communities as interacting individuals, individuals as organs, and so on down to the apparently indivisible, “fundamental” particles and forces. These systems vary in spatio-temporal scale, but they also can be thought about as a hierarchy of embedded systems (or embedded *mereological* systems). The term “levels of analysis” is

most often used, which is appealing in that it implies that the purpose of parsing systems into levels is to make them more accessible to our understanding or analysis.

Neuroscience is largely concerned with crossing levels of analysis, often by attempting to explain behavioral phenomena in the form of physiological processes. Claims about the mechanisms of behavior are common and elicit far more excitement – and are correspondingly more publishable – than more conservative “descriptions” of physiological or behavioral processes. When controversies arise, they are often based on disagreements over what empirical and computational evidence is sufficient to demonstrate a mechanism. Why is it so difficult to decide whether one set of processes is a mechanism of another? On the surface, it seems no harder than determining whether the process meets the definition. Craver (2007) has defined mechanism as a “set of entities and activities organized such that they exhibit the phenomenon to be explained”. He has also noted that it is important to include all and only those components and actions that are relevant (Craver, 2008). This all sounds very sensible, in principle. In practice, mechanisms become very difficult to identify due to difficulties or disagreements in answering three main questions:

- (1) How do we define the phenomenon we are attempting to explain? Specifically, what level of detail is necessary to describe a system’s behavior, and what conditions must it generalize across before we can properly attempt to explain it? As an example, if we are interested in understanding the processes that underlie retrieval of memories for experiences, to what extent should the proposed mechanism be specific for, say, visual features of the experience, as compared with spatial? Another example commonly encountered when discussing the memory roles of the mPFC: to what extent are we concerned with memory for the experience as a whole, including its incidental details, versus only those associations that help an animal predict rewards and punishers?
- (2) How much precision and accuracy is necessary for describing the parts of the system and their actions? Any useful understanding will necessarily require filtering out details (by including only the components and actions that are relevant), but the loss of details can also influence the accuracy of the described mechanism. Those who use simulations to study mechanism encounter this question all of the time. It is neither practical nor useful to model a network of 10,000 neurons using highly-detailed, multi-compartment simulations of each neuron, but by simplifying the cells as single-compartment, integrate-and-fire neurons a certain degree of accuracy is lost. In the words of Frances Skinner “We should neither ignore the details nor be consumed by them” (Skinner and Mulloney, 1998). A related question is what “level” of analysis is the right one to build a mechanistic account? In other words: what is the most optimal way to subdivide a complex system to understand and predict its behavior?
- (3) How can descriptions of mechanism handle built-in redundancies in the functional components of the system? In Craver’s formulation of mechanism, all and only those components that contribute to a phenomenon should be included; but when the components of a system have overlapping functions, the removal of any one may have no impact on the system’s behavior. A specific example arose in a recent set of studies on molecules involved in memory (see Frankland and Josselyn, 2013). The question was whether the protein PKM-zeta was responsible for – meaning, part of the mechanism that explained – long-term maintenance of altered synaptic connections. While some manipulations suggested that PKM-zeta was involved, others which specifically isolated the protein suggested that it was not a necessary component. It may be the case

Table 1
Glossary of terms used.

Level of analysis	A specific range within the continuum of embedded, mereological systems. Within a level of analysis, interactions between units can be described as causes and consequences . Relationships between levels of analysis can be described as mechanisms and functions
Behavior	A description of a system's actions, which can also be considered the consequences of a system/behaving unit on external systems/units, along with how those consequences might have been caused by external forces
Mechanism	A description of a phenomenon (behavior of a system) based on the actions and interactions (behaviors ; i.e., causes and consequences) of the system's components. The mechanism of a system's behavior is the set of all functions of the system's components
Function	A description of the role of individual components in a system – or more specifically, the component's actions and interactions – for the behavior of the system
Causes	The external influences that determine the behavior of a system (or behaving unit) within a level of analysis . Also called the efficient or motivational cause
Consequences	The effects that the behavior of a system (or behaving unit) have on other systems/units within a level of analysis
Levels of understanding	Marr's distinction between a system's computation (behavior), algorithm (mechanism), and hardware of the system
Computation	The behavior of a system; i.e., the causes and consequences of a behaving unit. Presently used to describe the behavior of specific neural structures rather than that of the organism
Algorithm	The mechanisms of a system's behavior; i.e., the causes and consequences of components within a system which allow it to achieve a specific computation . Presently used to describe the processes by which a neural structure achieves its computations
Hardware/material	The substance or matter with which the system is implemented. The material can be thought of as the set of interacting components that make-up the components of the system, and can therefore be thought of as the “ mechanisms of the mechanisms ” (or further recursion)

that PKM-zeta is, in fact, an element that participates in the synaptic maintenance system, but because other proteins have sufficiently overlapping function, its removal has no effect in the behavioral paradigms used for its study. If this is determined to be the case empirically, then the mechanistic component could be re-defined in a more abstract way—e.g., a function performed more generally by the PKM protein family. However, if “function” is defined too specifically, this method might also unravel—an issue that we return to below in discussing the distinction between computation and algorithm.

In some cases, the above questions are relatively easy to answer, and the identification of a physiological mechanism for behavior can be very straightforward. The patellar reflex is a simple behavior in which the quadriceps are extended in response to sudden force applied just below the kneecap. Its physiological explanation is well summarized as a mechanical triggering of action potentials in the muscle spindle leading to, by way of a synapse, the stimulation of a motor neuron and subsequent muscular stimulation and contraction. This mechanism is accessible because both the behavior and the components are specifically defined, and the activities of each are conducive to study. Whether or not it is also sufficiently complete will depend on the specific goals of investigating the mechanism: e.g., to predict the behavior across circumstances with a high level of accuracy, to provide information on how to fix the system if some part of it is broken, to build artificial agents with the same behavior, etc.

Tackling behavioral mechanisms for more abstract psychological phenomena, such as episodic memory, is far more challenging. The phenomena are not specifically defined, nor are the actions and interactions of the system's components. They are also not easy to study. Although the behavior is supported by neurons, as in the patellar reflex example, there are *too many*. Any attempt to describe the system based on hundreds of thousands of detailed, heterogeneous cells is not likely to make it much easier to understand and predict the organism's behavior. This is because small changes in neuron activity or connectivity can result in very different circuit-level behavior (Getting, 1989).

One way to approach the question of how to mechanistically explain a system with too many parts is to re-define “parts” at an intermediate level—in this case, as neural circuits or brain regions. This reduces the number of variables and thereby makes it easier (more efficient) to (1) generate theories of the underlying causal interactions that determine a system's behavior, (2) test those theories, and consequentially (3) develop a model (or understanding) that allows accurate prediction of the behavior under a variety of circumstances. The challenge of this approach is that it requires

being able to specify the actions and interactions (causes and consequences) of those intermediate parts.

How can the actions and interactions of brain networks be identified and described? The problem is a hard one. Some philosophers have even questioned whether the nervous system can be decomposed into separate modules (Simon, 1969; Fodor, 1983). Efforts to model brain processes have come at the problem from very different angles. Some scientists argue that brain processes are best revealed through “bottom-up” methods. An extreme version of this is the effort to simulate a human brain by simulating the actions and interactions of as many connected neurons (Markram, 2012). Previous studies that have taken this direction have yielded computationally intensive simulations that lacked simulations of perception, action, or other meaningful behaviors (Izhikevich and Edelman, 2008). At the other end of the spectrum are “top-down” approaches. Our broad and historically rich field of psychology often considers mechanistic explanations for behavior by parsing those behaviors into apparent component processes. The absence of physiological anchoring in these methods can greatly limit the explanatory power of the proposed mechanisms.

A more direct approach to addressing the intermediate level is to parse the brain according to specific structures, or circuits with high internal connectivity, and then determine the actions of and interactions between them. This approach is the basis for most systems neuroscience research, but the field is handicapped by the inherent difficulty in identifying the behaviors of individual networks. Few regions directly receive input from the external world, and similarly few directly synapse onto muscles or glands. In other words, the inputs and outputs are not directly, experimentally accessible by observing or stimulating an animal behaviorally. By analogy with the ethology field, it would be like studying the behavior of individuals through their communities: how does one animal react when a visual stimulus is presented to its littermate? What can be said about one animal's vocalizations based on migrations of the herd before and after its removal? The methods are not fruitless, but because they are indirect they also have the potential to be misleading. Advances in this field depend on an often slow but ultimately effective cycling of trial and error—involving the “Popperian” cycling of observation, theory, and testing in ways that rely on both bottom-up and top-down knowledge to determine brain circuit computations.

3. Isolating the challenge: Identifying the computations of neural systems

Recognizing the difficulties of explaining psychological phenomena, David Marr and Tomaso Poggio developed a

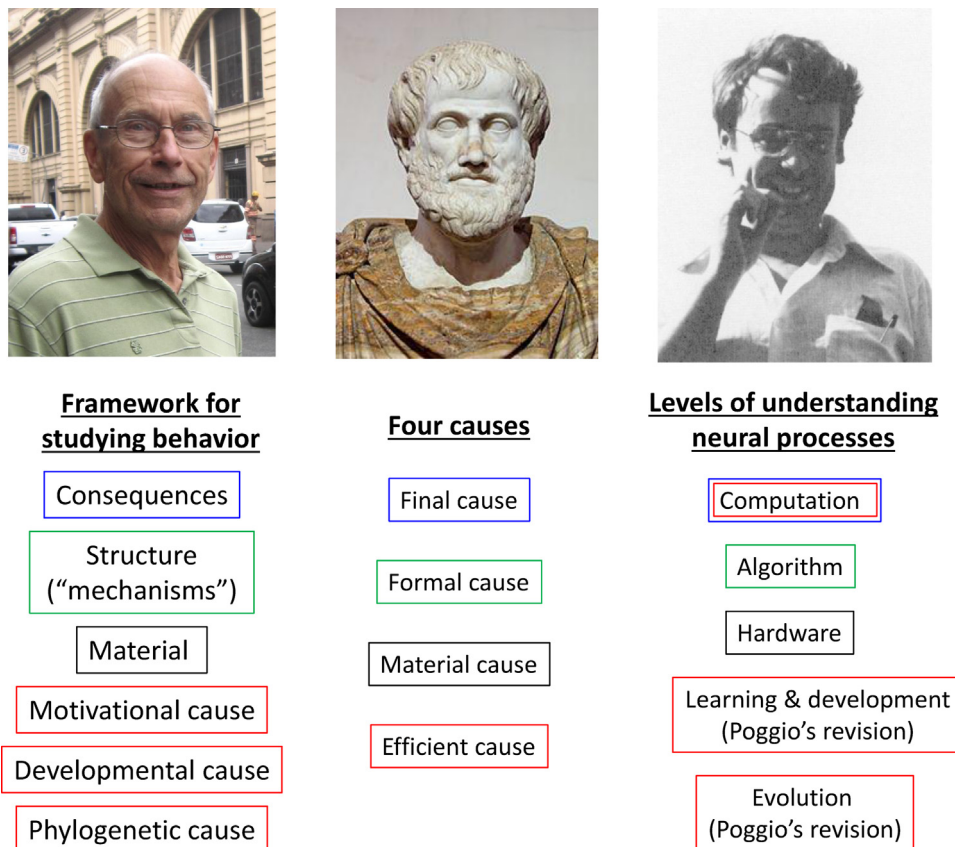


Fig. 1. Different classes of explanations. Aristotle subdivided explanations into four categories of “causes” (middle column). Jerry Hogan applied a similar classification scheme to explanations of animal behavior (left column), and these groupings can also be compared to David Marr’s scheme for understanding neural system operations (right column). Boxes with the same color reflect categories with related features, although mappings are not exact. Marr’s “computation” is highlighted in both blue and red because the term refers to both the output (consequences) of a system as well as how those outputs are driven by inputs (the efficient or motivational causes).

classification scheme that is in many ways comparable to that of Aristotle’s (Marr and Poggio, 1977; Marr, 2010). Marr called these “levels of understanding,” which is slightly confusing in that the levels overlap with, but are logically distinct from, “levels of analysis.” The three levels are computation, algorithm, and hardware. Computation, which shares elements with Aristotle’s efficient and final causes, refers to a system’s behavior, with emphasis on how the system’s actions (outputs) influence other systems (as inputs). A system’s *computation* is, essentially, its “function” in the mathematical sense of describing a relationship between inputs and outputs, or in the sense that Jerry Hogan may use the term as a description of the consequences of the system’s behavior. The *algorithm*, a variation of Aristotle’s formal cause, is the method for accomplishing a computation—essentially the mechanism. The *hardware* is synonymous with Aristotle’s material cause and, in spite of its name, does not have to be hard in the sense of being physically unyielding. It is the substance which implements the algorithms; in the case of the brain, this is the neural tissue itself. In a sense, the material is the set of components that make-up the system’s parts, and can therefore be thought of as the “mechanisms of the mechanisms.” Approximate relationships between the terminology used by Aristotle, Hogan, and Marr are illustrated in Fig. 1.

Each of Marr’s levels depend on one another, but in principle the same computation can be implemented with different algorithms or hardware. This is no different than saying that a phenomenon might have multiple possible mechanisms. A recent example of this concept was presented by Prinz et al. (2004). The authors created a 3-neuron network simulation of a circuit in the lobster stomach and systematically changed cell parameters and connection strengths

to generate 20 million different variations. Many of the model variants exhibited indistinguishable network outputs. One might point out that the degrees of freedom for behavior in this simulation environment was highly constrained. But it is a proof of principle that, extended further, implies that a mechanistic understanding of an organism’s behavior can be obtained without an understanding of the mechanisms of the underlying circuits.

Referring to the input–output relationships of a circuit as a computation raises the question of whether the computational theory requires mathematics or computers (meaning, computers that are not our brains). Clearly both can be helpful. Math is an efficient language for describing relationships between parts of a system, and computers provide a resource for offloading some of the otherwise overwhelming complexity and memory demands of a theory, helping to ensure against computational paradoxes. Human language has a number of “bugs” (to borrow from Buonomano, 2012), so formalizing simulations in math or computer operations can help make hypotheses explicit and communicable. At the same time, simulations are not exempt from questions about how generalizable and accurate an explanation must be, and no amount of mathematics will transmute trash into gold.

We have not, at this point, solved the challenges about mechanism raised in the previous section. All we have done is introduced the language of Marr and Poggio’s levels of understanding, descended from mathematics and computer science, as a substitute for levels of analysis. However, applying this language selectively to the intermediate level between an animal and its single neurons helps define the table space on which to construct mechanistic explanations of behavior. Somewhere between the activity of individual neurons and the behavior of organisms are

the *computations* of specific brain regions and their subcircuits. These computations can be described as input–output relationships, which contribute in particular ways – i.e., have specific functions – for behavior. They are also implemented by algorithms defined by the interactions between neurons, which can be inferred in part by the connection patterns, the circuit architectures. As will be described in the next section, decomposing behaviors into sets of interacting computations can dramatically improve our ability to understand the circumstances in which certain regions of the brain will be important; in principle, it may also improve our ability to fix disrupted brain processes or create artificial, intelligence systems.

4. Translating behavioral functions into computations: Example of the hippocampus and prefrontal cortex

The hippocampus and prefrontal cortex are in some ways ideal examples of brain regions that demand theories of computation. The hippocampus is a double-fold of cortex in the mammalian temporal lobe with a particular, regular circuitry; the medial prefrontal cortex (mPFC) is the rostral third of cingulate cortex that exhibits much of the circuit structure found throughout the rest of the cortex. Both regions are relatively distant, in terms of number of synapses, from those parts of the nervous system that transduce sensory stimuli or which stimulate muscle contraction or hormone release. Perhaps because of this, their behavioral functions are diverse and often elusive. In this case “behavioral functions” refers to the behavioral/cognitive capacities a brain region contributes to, which can be inferred by anatomy, effects of damage, and physiological activity. The two regions can both be divided into sub-regions that differ in architecture and connectivity, revealing a heterogeneity of form that also helps to explain the heterogeneity of function.

Ideas about the contribution of the hippocampus to behavior have taken several notable historical turns. The hippocampus has been thought to be involved in olfaction (Broca, 1878), emotion (Papez, 1937), and in spatial navigation (O’Keefe and Nadel, 1978). Most investigations of the hippocampus have targeted the region’s role in learning and memory, a function that was first brought into the spotlight by Scoville and Milner (1957) and which we return to in the last section. A complete description of how the hippocampus contributes to behavior would have to include elements of all of these, which also vary in degree across the length of the structure (Fanselow and Dong, 2010).

Behavioral functions that have been associated with the mPFC are even more extensive than those associated with the hippocampus. Gage et al. (2008) point out that the number of paper abstracts which mention the cingulate cortex has grown exponentially, far faster than those which mention motor or auditory cortex. The authors applaud human imaging technology for revealing that “The cingulate cortex is responsible for everything.” In a sense it must be true that the mPFC contributes to everything, as damage to the region can lead to a dramatic dimming of emotion and behavior known as “akinetic mutism” (Vogt et al., 1992). Differences across subregions of the mPFC have been extensively studied (Vogt et al., 1992; Bush et al., 2000), with more dorsal regions specifically implicated in the ability of animals to maintain pursuit of a task goal (Hadland et al., 2003; Walton et al., 2002), more ventral regions linked to the ability to regulate emotion (Milad and Quirk, 2002; Amat et al., 2005), and middle-regions seemingly important for an animal’s ability to select behaviors that are appropriate to a context (Marquis et al., 2007).

This plethora of behavioral functions can be simplified by an understanding of the region’s computations. This is because the behaviors are a manifestation of the computations applied to a specific set of inputs and outputs; thus, the same basic principles

of input–output relationships can be used to explain, say, spatial navigation, memory, and certain properties of language. In the final chapters of the classic *The Hippocampus as a Cognitive Map*, O’Keefe and Nadel (1978) propose that the computational principles by which the hippocampus represents physical space may also be applied to navigation through semantic space, thus explaining some of the memory and language deficits that can arise from hippocampal damage. Redish (1999), in his book *Beyond the Cognitive Map*, continued this tradition with updated and more formal descriptions of the computational operations performed by the hippocampus, explaining how these operations, applied to particular inputs, could support memory, navigation, and other behavioral functions. Within the prefrontal cortex field, a relatively early “theory of everything” was described in Joaquin Fuster’s book, *The Prefrontal Cortex*, in which he discusses the region’s role in temporally organizing behavior through the representation of schemas. He defined schema as an abstracted representation of the program of action, and describes how these are built from, and act back upon, lower levels of the action–perception hierarchy (Fuster, 2008)¹. Fuster’s hypothesis may appear vague to the point of having limited predictive value, and therefore also difficult to disprove. But it can be distinguished from many other hypotheses in the sense that explaining the region’s behavioral roles becomes secondary to explaining its input–output relationships with other brain structures. Importantly, the actual algorithms by which the prefrontal cortex functions may not be substantially different from other regions of cortex, but the behavioral functions are achieved by how the region is connected to its inputs and outputs (i.e., the *computation* is based on the *consequences*, or interactions, of the region on other brain regions).

Some of the most established theories about what these regions compute are based on simulations that use an architecture known as parallel distributed processing (PDP), built by McClelland and Rumelhart (1987) and Rumelhart and McClelland (1987). In these models, large-scale networks are built from simplified representations of single neurons and synapses. PDP provided a bridge between observed network dynamics described by the physicist Hopfield (1982) and theories about the behavioral functions of brain regions. One apropos example is a computational description of how the hippocampus helps reinstate, and thereby interleave, activity patterns in the cortex (McClelland et al., 1995). Another is a set of simulations describing how the prefrontal cortex may represent combinations of contextual information, about both situations and goals within those situations, to bias or override more direct sensory–motor connections (Cohen et al., 1990, 1996). The underlying concepts of the prefrontal cortex models have been disseminated to the broader scientific community through a highly cited review by Miller and Cohen (2001).

Theories of computation have also taken smaller scales, focusing on the specific operations of sub-circuits within the regions. Many efforts have been made to reduce cortex connections into a simplified, “canonical” microcircuit, the behavior of which may constitute a unit of cortical processing (e.g., Douglas et al., 1989). Although most of these studies have focused on regions of primary sensory cortex (Douglas and Martin, 2004), it is likely that the same principles can be extended to understand the medial prefrontal cortex. For example, in a phenomenon often referred to as input normalization, or gain control, inhibitory components of a circuit help compensate for increased input to a region, thereby

¹ Action–perception hierarchy, or more properly the hierarchy of action–perception cycles, refers to what Sherrington would have called the hierarchy of reflex arcs; meaning, the neural connections that link sensory signals to motor. These are organized in the nervous system hierarchically, with direct connections at lower levels being controlled by more indirect connections at higher levels.

maintaining activity levels among the population of projection neurons (see Carandini and Heeger, 2012). It has been suggested that a disruption of this process in the medial prefrontal cortex may take place in schizophrenia, producing, in dorsal regions, disproportionate activity levels (and consequent attention) under situations of increased sensory stimulation and, in more ventral regions disproportionate activity levels (and consequent overvaluation) of temporal lobe activity corresponding to certain cognitive states (Insel and Barnes, 2014). The hippocampus, meanwhile, is made-up of several, mostly feed-forward layers of processing, including the dentate gyrus, area CA3, and area CA1. It has long been thought that the recurrent network of CA3 is capable of rapidly storing associations, and then completing the patterns when later presented with a subset of the inputs (Marr, 1971; McNaughton and Morris, 1987). Empirical investigations across a range of techniques have found support for this theory (Guzowski et al., 2004; Neunuebel and Knierim, 2014). The dentate gyrus, meanwhile, has been proposed to generate very different output patterns relative to its inputs, called “pattern separation” (Treves and Rolls, 1992; O'Reilly and McClelland, 1994; Treves et al., 2008 offers a thorough and scholarly review on this topic). This theory has also found empirical support (McHugh et al., 2007; Niibori et al., 2012; Neunuebel and Knierim, 2014). It is worth briefly mentioning that the pattern separation theory of the dentate gyrus has in some cases been conflated with the possible behavioral functions of a hippocampal pattern separator (see Santoro, 2013), and in other cases with the originally proposed mechanism of pattern separation, known as “expansion recoding.” Expansion recoding, in which two overlapping input patterns become less overlapping when mapped onto a field of many more nodes (Marr, 1969, proposed that this algorithm was implemented in the cerebellum), does not appear to explain the role of the dentate gyrus in pattern completion (Leutgeb et al., 2007; Alme et al., 2010).

To summarize, behavioral functions of the hippocampus and the mPFC appear to be highly varied, without links between them that are obvious in some *a priori* way. Attempts to articulate specific computations performed by the regions, and their underlying circuits, have begun to provide frameworks for understanding the diversity of behavioral functions. In the prefrontal cortex, these computations rely on the same network phenomena observed in all areas of cortex, which have been formalized in simplified fashion by parallel distributed processing simulations. Applying these principles to specific inputs and outputs – i.e., to connections with cortical regions involved in perception and action – reveals functions such as context- and goal-based biasing of actions. Computations in cortex also may rely on input normalization (or gain control), the disruption of which may result in neuropsychiatric symptoms. In the hippocampus, computations may be based on a sequence of pattern separation and pattern completion processes that enable a network to both form new activity patterns for new input streams, and subsequently re-instate these for the formation of memories or execution of context-appropriate actions. In the final section we take a closer look at how these apparently different underlying processes found in the mPFC and hippocampus may support overlapping memory functions.

5. Computations of the hippocampus and mPFC in memory expression: Same or different?

Both the hippocampus and mPFC are known to contribute to the expression of learned memories. The hippocampus is required to learn and is at least temporarily required for the behavioral expression of certain types of associations (a scattering of examples from the literature include memories about spatial goal locations: Morris et al., 1982; trace eyeblink associations: Solomon et al., 1986; socially-transmitted food preferences: Winocur, 1990; and

arbitrary stimulus–response mappings: Brasted et al., 2003). The mPFC is also necessary for expressing these memories, although its function seems to be acquired over the course of weeks or months (examples across these same memory domains include: spatial goal locations: Teixeira et al., 2006; trace eyeblink associations: Takehara et al., 2003; trace fear: Blum et al., 2006; contextual fear: Frankland et al., 2004). One could therefore say that the hippocampus and mPFC share *behavioral functions* at different time periods of the memory's life. This raises the question of whether the hippocampus and mPFC also share computations and algorithms.

The hippocampus and mPFC likely implement very different algorithms. This is suggested by differences in circuit architecture, but also more directly by the different representations observed in the two regions, as measured by single neuron activity patterns. Neural activity, such as single neuron firing patterns, can be studied with respect to an animal's behavior and environment. A firing field (in sensory systems, “receptive field”) is the range of behavior or stimulus space over which the neuron will typically increase its firing frequency. The firing fields of neurons in the hippocampus and mPFC are highly complex, and are influenced by a combination of many environmental and behavioral variables. But the two regions differ in some key ways. Neurons in the hippocampus can suddenly alter their firing fields from one set of circumstances to another. When this happens, it often happens across many or all neurons of the network simultaneously, causing a scrambling of the selectivity map called “remapping” (Muller and Kubie, 1987). This property of the hippocampal output can be thought of as a process of pattern separation combined with pattern completion, presumably the combined effect of circuitry within the dentate gyrus and CA3 regions (discussed above). The same effect has not been reported in the mPFC. Neurons in the mPFC are typically selective for many environmental and behavioral variables, but they do not appear to spontaneously acquire different firing fields, and the changes are not thought to take place coherently throughout the network (Jung et al., 1998; Takehara-Nishiuchi and McNaughton, 2008). Although these results do not constitute descriptions of the algorithms used by the hippocampus and mPFC, or comprehensively specify their differences, they do provide evidence that the algorithms differ. The important distinction is not so much *what* information is coded for by the two networks, which will vary according to differences in their inputs (although, notably, inputs from entorhinal cortex to the two areas are highly overlapping, as reviewed by Insel and Takehara-Nishiuchi (2013)), it is that the two networks differ in how they respond to changes in some or part of the environment, which suggest different underlying processes.

If the memory-related behavioral functions are sometimes the same, but the algorithms are (apparently) different, are the computations performed by the hippocampus and mPFC during memory expression the same or different? Data that are currently available suggest that the computations may have some overlap, but differ in key respects. Both the hippocampus and mPFC may “bind” memory elements together, perhaps by serving as a type of index to those elements. This is supported by the fact that memory functions depend upon learning-related plasticity processes in both the hippocampus (Morris et al., 1986) and mPFC (Takehara-Nishiuchi et al., 2006; Restivo et al., 2009; Vetere et al., 2011). The hippocampus may have access to these distributed representations through the entorhinal cortex, which is connected to regions throughout the neocortex (Insausti et al., 1997). The mPFC may make these connections more directly (Frankland and Bontempi, 2005), or alternatively may also support memory through connections with the entorhinal cortex (Insel and Takehara-Nishiuchi, 2013). At the same time, the amount and type of associative memory may differ between the hippocampus and mPFC. Nadel and Moscovitch (1997) describe a loss of episodic detail in patients with damage to the medial temporal lobe, and subsequent studies in rats also suggest

that the absence of a hippocampus leads to a transformed and in some ways impoverished structure of associations (Winocur et al., 2007). These investigations are also relevant for a recent behavioral experiment in our lab, which demonstrated that spatial search behaviors based on previous experience changes as that experience becomes more distant (and potentially less dependent on the hippocampus; Richards et al., 2014).

6. Summary

In this article we have followed Jerry Hogan's lead, by temporarily putting aside the question "Why do animals behave the way they do?" to orient ourselves on what this question means. Specifically, what does it mean to build a mechanistic explanation of behavior? Mechanistic explanations face a number of challenges, including (1) specifying the behavior in a way that is concrete enough, and yet also generalizable enough, to be useful, (2) specifying the system's parts and interactions in a way that is sufficiently accurate to be predictive while also simple enough to be understandable (3) finding a description that takes into account functional redundancies in complex systems. Mechanisms also depend on the ability to identify the appropriate level of analysis, so that the behavioral function of a neural region is not confused with the region's computation, computation confused with algorithm, nor algorithm with hardware. A history of research on the hippocampus and prefrontal cortex provides an example of how these challenges have been addressed and how ideas about behavioral function, computation, algorithm, and hardware have developed—in some cases becoming confused. Multiple approaches have been used to simplify the complex neural structures, and with the help of computer simulations a number of computational principles have been described. Although a great deal remains unknown about the degree to which the hippocampus and medial prefrontal cortex share computations, it is clear that debates about whether they perform the same memory function must be framed within theories about the their computations.

One topic that was not addressed in this essay was another form of explanation: the process by which the behaving system is built. These might be called "generative" explanations, which include developmental processes as well as day-to-day plasticity. And there is also the underlying question of how a system capable of development and plasticity evolved in the first place. These are fundamental topics for ethology and neuroscience. Jerry Hogan has discussed the generative processes under the heading of developmental and phylogenetic causes. Although Marr did not address them, Poggio (2012) explicitly includes development, learning, and evolution in his revision of Marr's level's of understanding framework. Although these topics are not taken-up presently, they should not be omitted from any discussion on how to classify explanations.

As we began by invoking Aristotle, so we end. Aristotle considered the processes that generate our behavior and cognition to be among the most important, if most difficult, fields of knowledge, which he states in the introduction of *De Anima*. Throughout the treatise he struggles with the question of what the causes and "parts" of mind are—how to subdivide a system that seems to have many modes of operation but also seems to operate as a unit. If it is divisible, he asks, then what holds the parts together? The past century and a half of brain science has revealed that the brain's behavioral functions can be subdivided across many forms and scales, but weaving these into satisfying mechanistic explanations will depend on first spinning the threads of circuit computation.

Acknowledgement

We would like to thank Jerry Hogan for illuminating discussion.

References

- Alme, C.B., Buzzetti, R.A., Marrone, D.F., Leutgeb, J.K., Chawla, M.K., Schaner, M.J., Bohanick, J.D., Khoboko, T., Leutgeb, S., Moser, E.I., Moser, M.-B., McNaughton, B.L., Barnes, C.A., 2010. Hippocampal granule cells opt for early retirement. *Hippocampus* 20, 1109–1123, <http://dx.doi.org/10.1002/hipo.20810>.
- Amat, J., Baratta, M.V., Paul, E., Bland, S.T., Watkins, L.R., Maier, S.F., 2005. Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. *Nat. Neurosci.* 8, 365–371, <http://dx.doi.org/10.1038/nn1399>.
- Aristotle, 2014. *Metaphysics*, 350 BCE. (<http://classics.mit.edu/Aristotle/metaphysics.html>) (last referenced Aug. 21, 2014).
- Blum, S., Hebert, A.E., Dash, P.K., 2006. A role for the prefrontal cortex in recall of recent and remote memories. *NeuroReport* 17, 341–344, <http://dx.doi.org/10.1097/01.wnr.0000201509.53750.bc>.
- Brasted, P.J., Bussey, T.J., Murray, E.A., Wise, S.P., 2003. Role of the hippocampal system in associative learning beyond the spatial domain. *Brain* 126, 1202–1223.
- Broca, P., 1878. *Anatomie comparee des ciconvolutions cerebrales: le grand lobe limbique et la scissure limbique dans la serie des mammiferes*. *Rev. Anthropol.* 382, 1, n.d.
- Buonomano, D., 2012. *Brain Bugs: How the Brain's Flaws Shape Our Lives*. W.W. Norton, New York; London.
- Bush, Luu, Posner, 2000. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci. (Regul. Ed.)* 4, 215–222.
- Carandini, M., Heeger, D.J., 2012. Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* 13, 51–62, <http://dx.doi.org/10.1038/nrn3136>.
- Cohen, J.D., Dunbar, K., McClelland, J.L., 1990. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* 97, 332–361.
- Cohen, J.D., Braver, T.S., O'Reilly, R.C., 1996. A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. *Philos. Trans. R. Soc. London, B: Biol. Sci.* 351, 1515–1527, <http://dx.doi.org/10.1098/rstb.1996.0138>.
- Craver, C.F., 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Carlandon Press, Oxford, UK.
- Craver, Craver, F., 2008. Physical law and mechanistic explanation in the Hodgkin and Huxley model of the action potential. *Philosophy of science* 75, 1022–1033.
- Douglas, R., Martin, K., Witteridge, D., 1989. A canonical microcircuit for neocortex. *Neural Computation* 1, 480–488.
- Douglas, R.J., Martin, K.A.C., 2004. Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451, <http://dx.doi.org/10.1146/annurev.neuro.27.070203.144152>.
- Fanselow, M.S., Dong, H.-W., 2010. Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* 65, 7–19, <http://dx.doi.org/10.1016/j.neuron.2009.11.031>.
- Fodor, J.A., 1983. *The Modularity of Mind: An Essay on Faculty Psychology*. MIT Press, Cambridge, MA.
- Frankland, P.W., Bontempi, B., 2005. The organization of recent and remote memories. *Nat. Rev. Neurosci.* 6, 119–130, <http://dx.doi.org/10.1038/nrn1607>.
- Frankland, P.W., Josselyn, S.A., 2013. Neuroscience: memory and the single molecule. *Nature* 493, 312–313, <http://dx.doi.org/10.1038/nature11850>.
- Frankland, P.W., Bontempi, B., Talton, L.E., Kaczmarek, L., Silva, A.J., 2004. The involvement of the anterior cingulate cortex in remote contextual fear memory. *Science* 304, 881–883, <http://dx.doi.org/10.1126/science.1094804>.
- Fuster, J.M., 2008. *The Prefrontal Cortex*. Academic Press/Elsevier, Amsterdam; Boston, MA.
- Gage, Gregory J., Parikh, Hirak, Marzullo, Timothy C., 2008. The cingulate cortex does everything. *Ann. Improbable Res.* 14, 12–15.
- Gettings, P.A., 1989. Emerging principles governing the operation of neural networks. *Annu. Rev. Neurosci.* 12, 185–204, <http://dx.doi.org/10.1146/annurev.ne.12.030189.001153>.
- Guzowski, J.F., Knierim, J.J., Moser, E.I., 2004. Ensemble dynamics of hippocampal regions CA3 and CA1. *Neuron* 44, 581–584, <http://dx.doi.org/10.1016/j.neuron.2004.11.003>.
- Hadland, K.A., Rushworth, M.F.S., Gaffan, D., Passingham, R.E., 2003. The anterior cingulate and reward-guided selection of actions. *J. Neurophysiol.* 89, 1161–1164, <http://dx.doi.org/10.1152/jn.00634.2002>.
- Hogan, J.A., 1994. The concept of cause in the study of behavior. In: Hogan, J.A., Bolhuis, J.J. (Eds.), *Causal Mechanisms of Behavioural Development*, second ed. Cambridge University Press, Cambridge, UK, pp. 3–15.
- Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558.
- Insausti, R., Herrero, M.T., Witter, M.P., 1997. Entorhinal cortex of the rat: cytoarchitectonic subdivisions and the origin and distribution of cortical efferents. *Hippocampus* 7, 146–183.
- Insel, N., Takehara-Nishiuchi, K., 2013. The cortical structure of consolidated memory: a hypothesis on the role of the cingulate–entorhinal cortical connection. *Neurobiol. Learn. Mem.* 106, 343–350, <http://dx.doi.org/10.1016/j.nlm.2013.07.019>.
- Insel, N., Barnes, C.A., 2014. Differential Activation of Fast-Spiking and Regular-Firing Neuron Populations During Movement and Reward in the Dorsal Medial Frontal Cortex. *Cereb. Cortex.*, <http://dx.doi.org/10.1093/cercor/bhu062>.
- Izhikevich, E.M., Edelman, G.M., 2008. Large-scale model of mammalian thalamo-cortical systems. *Proc. Natl. Acad. Sci. U.S.A.* 105, 3593–3598, <http://dx.doi.org/10.1073/pnas.0712231105>.

- Jung, M.W., Qin, Y., McNaughton, B.L., Barnes, C.A., 1998. Firing characteristics of deep layer neurons in prefrontal cortex in rats performing spatial working memory tasks. *Cereb. Cortex* 8, 437–450.
- Leutgeb, J.K., Leutgeb, S., Moser, M.-B., Moser, E.I., 2007. Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* 315, 961–966, <http://dx.doi.org/10.1126/science.1135801>.
- Markram, Henry, 2012. The human brain project. *Sci. Am.* 306, 50–55.
- Marquis, J.-P., Killcross, S., Haddon, J.E., 2007. Inactivation of the prelimbic, but not infralimbic, prefrontal cortex impairs the contextual control of response conflict in rats. *Eur. J. Neurosci.* 25, 559–566, <http://dx.doi.org/10.1111/j.1460-9568.2006.05295.x>.
- Marr, D., 1969. A theory of cerebellar cortex. *J. Physiol. (Lond.)* 202, 437–470.
- Marr, D., 1971. Simple memory: a theory for archicortex. *Philos. Trans. R. Soc. London, B: Biol. Sci.* 262, 23–81.
- Marr, D., 2010. Vision: a Computational Investigation into the Human Representation and Processing of Visual Information. MIT Press, Cambridge, MA.
- Marr, D., Poggio, T., 1977. From understanding computation to understanding neural circuitry. *Neurosci. Res. Prog. Bull.* 15, 470–488.
- McClelland, J.L., Rumelhart, D.E., 1987. Parallel Distributed Processing, Vol. 2: Psychological and Biological Models. MIT Press, Cambridge, MA [u.a.].
- McClelland, J.L., McNaughton, B.L., O'Reilly, R.C., 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–457.
- McHugh, T.J., Jones, M.W., Quinn, J.J., Balthasar, N., Coppari, R., Elmquist, J.K., Lowell, B.B., Fanselow, M.S., Wilson, M.A., Tonegawa, S., 2007. Dentate gyrus NMDA receptors mediate rapid pattern separation in the hippocampal network. *Science* 317, 94–99, <http://dx.doi.org/10.1126/science.1140263>.
- McNaughton, B.L., Morris, R.G., 1987. Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends Neurosci.* 10, 408–415.
- Milad, M.R., Quirk, G.J., 2002. Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* 420, 70–74, <http://dx.doi.org/10.1038/nature01138>.
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202, <http://dx.doi.org/10.1146/annurev.neuro.24.1.167>.
- Morris, R.G., Garrud, P., Rawlins, J.N., O'Keefe, J., 1982. Place navigation impaired in rats with hippocampal lesions. *Nature* 297, 681–683.
- Morris, R.G., Anderson, E., Lynch, G.S., Baudry, M., 1986. Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. *Nature* 319, 774–776, <http://dx.doi.org/10.1038/319774a0>.
- Muller, R.U., Kubie, J.L., 1987. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J. Neurosci.* 7, 1951–1968.
- Nadel, L., Moscovitch, M., 1997. Memory consolidation, retrograde amnesia and the hippocampal complex. *Curr. Opin. Neurobiol.* 7, 217–227.
- Neunuebel, J.P., Knierim, J.J., 2014. CA3 retrieves coherent representations from degraded input: direct evidence for CA3 pattern completion and dentate gyrus pattern separation. *Neuron* 81, 416–427, <http://dx.doi.org/10.1016/j.neuron.2013.11.017>.
- Niibori, Y., Yu, T.-S., Epp, J.R., Akers, K.G., Josselyn, S.A., Frankland, P.W., 2012. Suppression of adult neurogenesis impairs population coding of similar contexts in hippocampal CA3 region. *Nat. Commun.* 3, 1253, <http://dx.doi.org/10.1038/ncomms2261>.
- O'Keefe, J., Nadel, L., 1978. The Hippocampus as a Cognitive Map. Clarendon Press; Oxford University Press, Oxford: New York, NY.
- O'Reilly, R.C., McClelland, J.L., 1994. Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus* 4, 661–682, <http://dx.doi.org/10.1002/hipo.450040605>.
- Papez, J.W., 1937. A proposed mechanism of emotion. *Arch. Neur. Psych.* 38, 725–743, <http://dx.doi.org/10.1001/archneurpsyc.1937.02260220069003>.
- Poggio, T., 2012. The levels of understanding framework, revised. *Perception* 41, 1017–1023.
- Prinz, A.A., Bucher, D., Marder, E., 2004. Similar network activity from disparate circuit parameters. *Nat. Neurosci.* 7, 1345–1352, <http://dx.doi.org/10.1038/nn1352>.
- Redish, A.D., 1999. Beyond the Cognitive Map: from Place Cells to Episodic Memory. MIT Press, Cambridge, MA.
- Restivo, L., Vetere, G., Bontempi, B., Ammassari-Teule, M., 2009. The formation of recent and remote memory is associated with time-dependent formation of dendritic spines in the hippocampus and anterior cingulate cortex. *J. Neurosci.* 29, 8206–8214, <http://dx.doi.org/10.1523/JNEUROSCI.0966-09.2009>.
- Richards, B.A., Xia, F., Santoro, A., Husse, J., Woodin, M.A., Josselyn, S.A., Frankland, P.W., 2014. Patterns across multiple memories are identified over time. *Nat. Neurosci.* 17, 981–986, <http://dx.doi.org/10.1038/nn.3736>.
- Rumelhart, D.E., McClelland, J.L., 1987. Parallel Distributed Processing, Vol. 1: Foundations. MIT Press, Cambridge, MA [u.a.].
- Santoro, A., 2013. Reassessing pattern separation in the dentate gyrus. *Front Behav. Neurosci.* 7, 96, <http://dx.doi.org/10.3389/fnbeh.2013.00096>.
- Scoville, W.B., Milner, B., 1957. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* 20, 11–21.
- Simon, Herbert A., 1969. The Sciences of the Artificial. MIT Press, Cambridge, MA.
- Skinner, F.K., Mulloney, B., 1998. Intersegmental coordination in invertebrates and vertebrates. *Curr. Opin. Neurobiol.* 8, 725–732.
- Solomon, P.R., Vander Schaaf, E.R., Thompson, R.F., Weisz, D.J., 1986. Hippocampus and trace conditioning of the rabbit's classically conditioned nictitating membrane response. *Behav. Neurosci.* 100, 729–744.
- Takehara, K., Kawahara, S., Kirino, Y., 2003. Time-dependent reorganization of the brain components underlying memory retention in trace eyeblink conditioning. *J. Neurosci.* 23, 9897–9905.
- Takehara-Nishiuchi, K., McNaughton, B.L., 2008. Spontaneous changes of neocortical code for associative memory during consolidation. *Science* 322, 960–963, <http://dx.doi.org/10.1126/science.1161299>.
- Takehara-Nishiuchi, K., Nakao, K., Kawahara, S., Matsuki, N., Kirino, Y., 2006. Systems consolidation requires postlearning activation of NMDA receptors in the medial prefrontal cortex in trace eyeblink conditioning. *J. Neurosci.* 26, 5049–5058, <http://dx.doi.org/10.1523/JNEUROSCI.4381-05.2006>.
- Teixeira, C.M., Pomedli, S.R., Maei, H.R., Kee, N., Frankland, P.W., 2006. Involvement of the anterior cingulate cortex in the expression of remote spatial memory. *J. Neurosci.* 26, 7555–7564, <http://dx.doi.org/10.1523/JNEUROSCI.1068-06.2006>.
- Treves, A., Rolls, E.T., 1992. Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus* 2, 189–199, <http://dx.doi.org/10.1002/hipo.450020209>.
- Treves, A., Tashiro, A., Witter, M.P., Moser, E.I., 2008. What is the mammalian dentate gyrus good for? *Neuroscience* 154, 1155–1172, <http://dx.doi.org/10.1016/j.neuroscience.2008.04.073>.
- Vetere, G., Restivo, L., Cole, C.J., Ross, P.J., Ammassari-Teule, M., Josselyn, S.A., Frankland, P.W., 2011. Spine growth in the anterior cingulate cortex is necessary for the consolidation of contextual fear memory. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8456–8460, <http://dx.doi.org/10.1073/pnas.1016275108>.
- Vogt, B.A., Finch, D.M., Olson, C.R., 1992. Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. *Cereb. Cortex* 2, 435–443.
- Walton, M.E., Bannerman, D.M., Rushworth, M.F.S., 2002. The role of rat medial frontal cortex in effort-based decision making. *J. Neurosci.* 22, 10996–11003.
- Winocur, G., 1990. Anterograde and retrograde amnesia in rats with dorsal hippocampal or dorsomedial thalamic lesions. *Behav. Brain Res.* 38, 145–154.
- Winocur, G., Moscovitch, M., Sekeres, M., 2007. Memory consolidation or transformation: context manipulation and hippocampal representations of memory. *Nat. Neurosci.* 10, 555–557, <http://dx.doi.org/10.1038/nn1880>.