Memory and Optogenetic Intervention: Separating the Engram from the Ecphory

Sarah K. Robins*

Optogenetics makes possible the control of neural activity with light. In this article, I explore how the development of this experimental tool has brought about methodological and theoretical advances in the neurobiological study of memory. I begin with Semon’s distinction between the engram and the ecphory, explaining how these concepts present a methodological challenge to investigating memory. Optogenetics provides a way to intervene into the engram without the ecphory that, in turn, opens up new means for testing theories of memory error. I focus on a series of experiments where optogenetics is used to study false memory and forgetting.

1. Introduction. “Optogenetics has ushered in a new era of potent and targeted control over multiple aspects of neural function” (Guru et al. 2015, 1). Such proclamations about optogenetics are rife in contemporary neuroscience. Optogenetics is a new intervention technique that allows neurons to be controlled with light. Its impact belies its age; the tool has only been available for little more than a decade. Within 5 years of the first study demonstrating its use, optogenetics was declared Method of the Year (Nature Methods, vol. 1, no. 8 [2011]) and Insight of the Decade (Science, vol. 330, no. 6011 [2010]). Its inventor, Karl Deisseroth, was profiled in the New Yorker (Colapinto 2015). Optogenetics has been used to aid inquiry into an array of neural systems, ranging from addiction to zebra fish. Given the hoopla, it is unsurprising that optogenetics has captured the interest of philosophers of neuroscience as well. To many, it looks to be a scientific revolution occurring in real time (Bickle 2016; Craver, forthcoming).

*To contact the author, please write to: Philosophy Department, University of Kansas, 3090 Wescoe Hall, 1445 Jayhawk Blvd., Lawrence, KS 66045; e-mail: skrobins@ku.edu.

†Many thanks to John Bickle for organizing the PSA 2016 symposium Intervention and Progress: Optogenetics and Control over Neural Systems, where this article was originally presented. Thanks, too, for comments and feedback from the audience and my fellow symposium participants, Carl Craver and Jacqueline Sullivan.

Philosophy of Science, 85 (December 2018) pp. 1078–1089. 0031-8248/2018/8505-0027$10.00
Copyright 2018 by the Philosophy of Science Association. All rights reserved.
Does optogenetics really have such potential? One way to make progress on this question is to observe how this intervention technique has influenced a particular domain of inquiry. Such is the aim of this article. I focus on the study of memory in cellular and molecular neuroscience. Memory serves as an interesting case study because it is a capacity for which the basic neural processes have long been understood. And yet, as I will show, the arrival of optogenetics has brought about significant methodological and theoretical advances in just a few years.

I begin with a review of the two theoretical posits that German biologist Richard Semon (1921) introduced into the scientific study of memory—the engram and the ecphory—and the methodological challenge their entanglement presents. I then introduce optogenetics and show how its application to engram theory provides a solution to this basic methodological puzzle. I focus on two experimental techniques using optogenetics from an extended research project based in the Tonegawa Laboratory at the Massachusetts Institute of Technology (http://tonegawalab.org/). In each case, optogenetic intervention has led to surprising discoveries that challenge the standard view of memory errors.

2. Engram Theory. Engram is a new word for an old idea. It is the current scientific term for the memory trace, an idea as old as thinking about memory itself. In conversation with Theatetus, Socrates likened the mind to a block of wax into which memories are impressed. These impressions are memory traces, which make possible the retention of information, ideas, and experiences over time. The comparison to wax tablets may no longer be illuminating, but the supposition that memory involves an enduring psychological or physical change to the rememberer continues (Robins 2017).

Semon (1921) coined the term engram in the early twentieth century, refashioning the age-old memory trace as the neurological process by which information is encoded, stored, and retrieved. The engram persists as the central concept guiding the investigation of memory in cellular and molecular neuroscience. Scientists working in this area often explicitly frame their research in terms of engram theory, where the central supposition is that engrams exist.1 There is some change to the brain, as the result of experience, by which the retention of information, ideas, and experiences is made possible. This shared commitment leaves plenty of room for debate and discovery; it alone does not settle what the engram is—whether it refers to a particular neural structure, neural mechanism, or set of neural activities. Semon’s original definition was intentionally vague; when he was writing, the locus engram storage had

---

1. It is interesting to ask whether, in the neuroscience of memory, commitment to the existence of discrete memory traces is a pretheoretical commitment or empirical discovery. For a discussion of this issue, see De Brigard (2014b).
not yet been identified. Many memory scientists now believe that engrams are changes to the strength of synaptic connections between neurons, but even this is not universal. There are others who propose, for example, that engrams are established via phase coding of oscillatory patterns over the entire hippocampal formation (Hasselmo 2012). And philosophers of neuroscience continue to disagree as to the conclusions about the nature of explanation that can be gleaned from the search for and discovery of these neural mechanisms (e.g., Theurer 2013). Nonetheless, the neurobiological study of memory remains the study of engrams. The central project is to locate these engrams, identifying the mechanism(s) by which they are formed, retained, and retrieved.

3. The Methodological Challenge to Engram Theory. Understanding the methodological challenge for engram theory requires a return to the theoretical background from which the concept emerged. When theorizing about the nature of memory, Semon introduced two theoretical posits: the engram and the ecphory. The engram is the neural memory trace; the ecphory is the process by which that trace is reactivated to produce remembering. The two are complementary and equally important for successful remembering. Remembering, of course, requires the retention of information, but this alone is not sufficient. In addition, it requires the activation of the ecphoric retrieval process by which the dormant information is revived. Semon emphasized this point by insisting on two laws of memory, one corresponding to each component:

Law of Engraphy: All simultaneous excitations within an organism form a coherent simultaneous excitation-complex which acts engraphically; that is, it leaves behind it a connected engram-complex, constituting a coherent unity.

Law of Ecphory: The partial recurrence of the energetic condition, which had previously acted engraphically, acts ecphorically on a simultaneous engram-complex. Or, more precisely described: the partial recurrence of the excitation-complex, acts ecphorically on the latter, whether the recurrence be in the form of original or mnemic excitations. (1921, 148)

Given how little was known about the molecular processes of memory at the time of Semon’s writing, few have thought that the details of these laws would be applicable to the present-day study of the neural mechanisms of memory. His Law of Engraphy, for example, suggests that the engram involves activity throughout the entire brain, whereas most contemporary views attribute engram storage to the hippocampus exclusively (or include, at most, a few other selective brain regions). What has stood the test of time is Semon’s account of the ecphory and his insistence on the importance of both laws. Semon argued that any understanding of the engram, whatever it may turn out to be, must be paired with an understanding of this process of re-
excitation. Cognitive psychologists who study memory recognize this entanglement; retrieval is not a neutral intervention that can be used to probe the engram without disturbance (Schacter, Eich, and Tulving 1978). Instead, the ephoric process acts on the engram so that the resultant memory is a reflection of both what was stored and how it was retrieved.

The focus of the neurobiological study of memory, as the name engram theory reflects, is on the engram. Returning to Semon’s original account of the engram promotes awareness of the ephory as well, which has received little attention in memory science outside of cognitive psychology. Semon’s work reminds us that, when studying the engram, it is critical to keep the ephory in mind—both as a process in its own right and as a mitigating factor in our ability to access and understand the engram. As we pause to consider the ephoric process, and its deep interconnection with the engram, we become more aware of the difficulty of disentangling its influence on the act of remembering. This is the methodological challenge for engram theory.

While the distinction between the engram and the ephory is clear in principle, the two are difficult to disentangle in practice. How can we distinguish the information that is stored (the engram) from the process by which it is retrieved (the ephory)? Our only insight into the contents of memory comes from the act of remembering. We determine what has been stored in memory by investigating what can be retrieved. Outside of retrieval, there is no way of establishing memory’s contents. In other words, access to the engram is only possible via the ephory. Without the ability to intervene into the engram directly, our understanding of this basic mechanism of memory storage remains limited.

In cases of successful remembering, in which an accurate representation of the past experience is produced, the concern is minimal and can easily go unnoticed. But in cases of memory error, where we want to determine which factors contributed to the error and how, the methodological impasse becomes clearer. Memory errors can be sorted into two general types: errors of commission and errors of omission. Commission errors are false memories; instances where something purporting to be a memory is produced, but some or all of what is produced is inaccurate. Omission errors are forgetting errors. On these occasions, one attempts to remember but is unable to call the desired information or experience to mind.

Consider a case of false memory. Suppose you are reminiscing about a past dinner party and recall your uncle telling many funny jokes during the meal, only later to learn that, while the dinner did feature many jokes, they were not told by your uncle as he was not there. How does such an error occur? The problem could be with the engram. You could have encoded the event in-

2. For a defense of Semon’s work on this point, see Schacter (2001).
correctly, or the information may have degraded over time. But the problem could also be with the ephory; the retrieval process could have distorted an otherwise well-preserved, accurate engram. A similar problem arises for forgetting errors. Suppose a friend asks you about this dinner party and your mind goes blank. You have forgotten. This could occur because the engram has been lost or because the ephoric process is damaged so that it no longer activates the engram. If the only way to assess the engram is by activating the ephoric process of retrieval, then there is no way to sort between these alternatives.

4. Optogenetics and Engram Theory. In this section, I introduce optogenetics, an intervention technique that offers a way around the methodological puzzle introduced above. Optogenetics is an intervention technique by which living cells, particularly neurons, become light responsive. Opsins, light-sensitive proteins, make this possible. Neuroscientists genetically engineer model organisms (mice, rabbits, fruit flies, nematodes, etc.) so that they possess a transgene expressing a particular opsin. This transgene is then introduced into a particular type of neuron in the brain of the model organism, so that the selected cells become light responsive when the transgene is activated and the opsin is expressed. There are many different kinds of opsins. Some respond to yellow light, others to blue or red. Some opsins excite the cell; others inhibit. Using optical fiber implants, neuroscientists can shine colored light onto specific cells to activate or inhibit them. There are many reasons to be excited about optogenetics: it allows intervention into living systems, rather than inert tissue. The response to light application is instantaneous, making interventions temporally precise. Optogenetics is especially exciting for the study of memory, I suggest, because it offers a way to intervene into the engram directly, circumventing the ephory. In this way, it provides a novel workaround for the methodological challenge that plagues engram theory.

To illustrate this point, I discuss a series of findings from the Tonegawa Laboratory, where optogenetic intervention is used to identify, activate, and manipulate engrams. In these experiments, mice explore a novel conditioning chamber. The neurons active during this exploration encode the mouse’s experience of the environment—they constitute the engram for the spatial memory. Previous research has established that these engrams are encoded by neurons in the dentate gyrus (DG), a portion of the hippocampal formation. The mice in these studies are engineered so that the engrams they form will be light sensitive. The Tonegawa group uses the opsin Channelrhodopsin-2 (ChR2), a membrane protein found in algae. Mice are engineered to possess the ChR2 transgene in all DG neurons; the light-responsive ChR2 protein will be expressed whenever a DG neuron is activated. To ensure that light-responsive

DG neurons are only activated during the experimental condition in which the engram is formed, the mice are also engineered to be sensitive to doxycycline (dox, an antibiotic given through the animal’s water supply). Mice will only express the transgene when they are not being treated with dox.

The mice are given dox at all times except for during the experimental condition. During the experiment, the mouse enters the conditioning chamber and, as a result of its exploration, activates a set of DG neurons that encode a spatial memory for that context. Since all of the animal’s DG neurons have the ChR2 transgene and the animal is no longer taking dox, whichever of these neurons is recruited to form the resultant engram will be light responsive. To reactivate this engram, the researchers do not need to instigate retrieval, which for mice would mean returning them to the original conditioning chamber. Instead, the researchers can bypass the ecphory and proceed directly to the engram by simply turning on a light. That is, when the mouse’s DG is exposed to blue light (via the insertion of a fiber optic cable), the DG neurons that encoded the spatial memory will be reactivated.4

Activating the engram directly and overcoming this basic methodological challenge is, in and of itself, an exciting breakthrough. The impact on engram theory, and the neurobiological study of memory more broadly, becomes clearer once this basic technique is used to explore the influence of further manipulations of the engram and how they contribute to memory error. Doing so provides novel insight into our understanding of both false memory and forgetting errors, as I illustrate in the two case studies below.

Case Study 1: False Memory. False memories are memory errors of commission, inaccurate representations of past experiences that the rememberer herself takes to be genuine memories—like the example of the uncle at the dinner party from section 2. Decades of research in cognitive psychology reveal that such false memories are produced surprisingly easily and often. Memory scientists want to explain these errors, and when doing so, they face the methodological challenge. Are false memories the result of a missing engram or a distorting ecphory? Most contemporary philosophers of memory explain false memories in terms of the absence of an engram. In fact, many endorse constructivism, according to which memory does not make use of engrams at all, relying instead on a general network of information. The process of remembering, on this view, is making use of ecphoric resources—information available and of interest at the time of retrieval—to produce a plausible representation of a past experience (De Brigard 2014a; Michaelian 2016).

4. The Tonegawa group demonstrated the ability to produce and reactivate light-responsive engrams in mice by use of this technique (Liu et al. 2012). Given space limitations, I focus my discussion of experimental detail on the subsequent studies of engram manipulation.
Theorizing about false memory has not been constrained by evidence about the underlying neural mechanisms because there has been no such evidence. False memory is a human phenomenon; evidence has thus been restricted to that which is available with the experimental methods that can be used on humans—those of cognitive psychology and cognitive neuroscience. Without the ability to translate these questions to studies of nonhuman animals, using the methods of cellular and molecular neuroscience, there has been little hope of ever discovering such a mechanism. That is, until recently. Using optogenetic intervention to create and manipulate engrams, the Tonegawa research group has been able to produce an animal model of false memory.

In Ramirez et al. (2013), the experimental technique described above is used to create and then distort light-responsive engrams in mice, where a context the mouse has previously encountered is remembered as fearful even though the initial experience in the context did not include any fearful stimuli. First, the genetically modified mice form a light-responsive engram—while removed from dox, they are each given a novel conditioning chamber to explore. Each mouse is then dosed with dox, preventing the formation of any new engrams, and taken to a second novel conditioning chamber. Once in the second chamber, the optical implant is turned on, reactivating the original engram from the previous chamber. While this light-responsive engram is active, the mouse is given a set of foot shocks, instilling a fear memory. The result of this two-chamber process is a set of mice that each have a light-responsive engram with spatial information about one conditioning chamber and information about a fearful encounter in a second conditioning chamber.

Mice with this manipulated engram are then exposed to one of three test conditions—they are returned to the original chamber, returned to the second chamber, or taken to a third, novel conditioning chamber. When returned to the second chamber, where they received foot shocks initially, mice respond by freezing, a characteristic fear response. Interestingly, mice returned to the original chamber also display this fear response, freezing in place even though they did not receive foot shocks in this chamber. Mice taken to a novel chamber behave differently—they explore the novel environment, suggesting against the idea that the manipulated engrams have made the animals generally fearful.

In a second study (Redondo et al. 2014), the Tonegawa research group reproduced this result with more elaborate memories, demonstrating the ability to produce false memories where the emotional response is reversed. Using the same basic method, this study added valence to the initial engram: when mice were introduced to the first conditioning chamber, their exploration was paired with either a positive stimulus (exposure to a female mouse) or a negative stimulus (foot shocks). Next, for a set of these mice, reactivation of the light-responsive engram in the second chamber was paired with a stimulus of the opposite valence. Mice who previously received foot shocks were exposed
to a female mouse and vice versa. When later returned to the original chamber, the mice display behavior reflective of their experience in the second chamber—that is, mice that received foot shocks in the first chamber display exploratory, pleasure-seeking behavior and mice that were exposed to a female display a freezing fear response. Their behavior contradicts their experience in the chamber, indicating a reversal of the original engram’s valence.

These studies offer examples of distorted, false memories in mice. The mice returned to the original chamber exhibit this most clearly. These mice respond to the environment as familiar or remembered but behave in a way that fails to reflect their previous experience in this context. The information they have retained is distorted and inaccurate. The Ramirez et al. (2013) and Redondo et al. (2014) studies provide the first animal model of false memory. The Tonegawa research group interprets their findings accordingly, making an explicit connection between these results and studies of false memories in cognitive psychology (Ramirez et al. 2013, 290). Their willingness to claim such a connection is not, in and of itself, sufficient reason to believe that one exists. For current purposes, however, the claim is illuminating, as it is not one often made by researchers investigating the neural mechanisms of memory.

This breakthrough in the understanding of the mechanism of false memory is made possible by optogenetic intervention. By allowing for the direct excitation and then manipulation of an engram, researchers could explore possible changes to the engram that occur without influence from the ephory. In this way, the results from the Tonegawa research group provide the beginning sketches of a mechanism for false memory formation. This initial sketch is enough to shake up theoretical explanations of false memory. The results challenge constructivism, which explains false memory by denying the existence of traces. They show, instead, false memories as the result of manipulating but retaining traces. In this way, they suggest an alternative theoretical approach to false memory, one that retains an engram.

Case Study 2: Consolidation. Forgetting is a familiar and frustrating memory error. As introduced in section 3, attempts to explain forgetting raise a question complicated by the methodological blend of the engram and ephory. In cases of forgetting, we want to know whether the error is due to a failure of the engram or a failure of the ephory. Is the information unavailable or merely inaccessible (Tulving and Pearlstone 1966)? Because the engram can only be investigated via the ephory, this question is difficult to answer directly.

Progress on this question is also complicated by the fact that forgetting is hard to study experimentally, as it is difficult to predict when it will occur. One of the best avenues for studying forgetting is consolidation. Consolidation is
the process by which information moves from temporary to long-term memory storage. If the consolidation process is disrupted, forgetting results. The neural mechanisms of consolidation have been studied for more than a century and are well understood (McGaugh 2000). Put simply, learning involves changes to the interactions between neurons. Transitioning from learning to long-term memory storage requires making these structural alterations permanent. Consolidation is the process by which this occurs. The synthesis of proteins stabilizes the synaptic connections between neurons. If there is interference into the consolidation process—by administration of a protein synthesis inhibitor, for example—the result is forgetting.

The definition and investigation of consolidation presupposes an answer to the methodological question above. The consolidation hypothesis just is the claim that the engram moves, gradually, from an initial, fragile state to a more stable form (McGaugh 2000). Researchers who study consolidation simply assume that forgetting, at least in the case of failure to consolidate, is the result of damage to the engram rather than damage to the ecphory.

Recently, the Tonegawa research group has applied optogenetic methods to the study of consolidation-based forgetting. As with the studies of false memory discussed above, the Ryan et al. (2015) article begins with the creation of light-responsive engrams in genetically modified mice. In this study, these engrams are created for mice who are introduced to a novel conditioning chamber where they receive foot shocks. Immediately after the encounter, half of the mice are given anisomycin (ANI), a protein synthesis inhibitor, and the other half are given saline solution as a control. Following previous studies of consolidation interference, the expectation is that this manipulation will cause the mice given ANI to forget the foot shocks when they are later returned to the chamber, but the mice given only the sham intervention of saline will remember and respond accordingly (i.e., freeze when returned to the chamber). And this is what Ryan et al. found. At both 1 and 3 days after the initial encounter in the conditioning chamber, the mice given saline froze when they were returned, while the mice given ANI did not. By disrupting consolidation, the exposure to ANI appears to have induced forgetting in one set of mice.

Returning the mouse to the original conditioning chamber activates retrieval; it provides access to the engram ecphorically. Since the engram involved in the foot shock conditioning is light responsive, Ryan et al. (2015) were able to conduct a second set of experiments exploring the effect of activating the engram directly via optogenetic intervention. These experiments began in the same way as the first consolidation study: mice formed a light-responsive engram for the experience of receiving foot shocks in a conditioning chamber, and then half of the mice were given ANI and half were given saline. One day after this intervention, the mice were tested in the environment to ensure consolidation disruption (i.e., forgetting) in the mice that received ANI. The next day (2 days after the initial training), each mouse was placed in a distinct
chamber and the foot shock engram was activated via the mouse’s optical implant. In this condition, all mice—those who had received ANI and those who had received saline—displayed the characteristic fear response. Ryan and colleagues performed several variations of this experiment, testing for the effect of activating distinct portions of the engram optogenetically. For each manipulation, the result was the same. There was no significant difference between the recall behavior for mice whose consolidation had been disrupted and for those whose consolidation had not. By activating the engram directly, the researchers were able to reinstate the memory, or, optogenetic intervention allowed them to undo the forgetting caused by disruption to the consolidation process. The ability to recover the memory of the foot shock experience optogenetically persisted for 8 days after the original training session.

These results provide a novel answer to the methodological question posed above, one that challenges standard assumptions about consolidation. Forgetting, the failure to retrieve information, could be the result of engram decay or ecphoric failure. At least in the case of consolidation failure, memory scientists have supposed that forgetting is due to engram loss. But when nonconsolidated engrams are activated directly, as optogenetic intervention makes possible, remembering is successful. This suggests that the deficit in these cases of forgetting is due to the ecphory instead. As Susumu Tonegawa explained in an interview about this study, “The majority of researchers have favored the storage theory, but we have shown in this paper that this majority theory is probably wrong” (Knight 2015). Forgetting, at least in the case of failure to consolidate, is an ecphoric deficit, not an engraphic one.

Objection: Artificial Ecphory. I have attempted to illustrate ways in which the arrival of optogenetic intervention has revolutionized the study of memory in molecular neuroscience, claiming that it allows researchers to activate the engram directly, circumventing the ecphory. One might worry that I have overstated the advance being made. While optogenetics allows researchers to bypass the standard ecphoric process of retrieval, it may not allow them to avoid the ecphory in its entirety. Instead, what has been created is a new, artificial ecphory. This is an intriguing possibility, one that deserves far more consideration than I have space for here. What I can say is that my claims about the engram-ecphory entanglement have involved certain assumptions about the nature of retrieval. I have assumed that the ecphory’s influence is problematic because it is indirect. During retrieval, engram neurons are reactivated by other neurons, and, since this activation can come from many different directions, this opens up the possibility for at least some of those routes of influence to have a distorting influence. My hope was that optogenetic intervention avoided this because the engram neurons were activated by light instead. But if the

6. I am grateful to an anonymous reviewer for raising this objection.
ecphoric influence is built in to the very act of reactivation, however it occurs, then the studies being discussed here have not circumvented the ecphory as I have claimed—indeed, because nothing could. This would mean the methodological puzzle cuts even deeper than I suggested initially; there is no hope of escape. This possibility warrants further attention. But even if this turns out to be the case, the merits of optogenetic intervention are still clear. The method creates an artificial ecphory that comes as close as any could to direct engram activation.

5. Conclusion. Many neuroscientists and philosophers of neuroscience are excited by the potential for optogenetics to revolutionize our understanding of the neural mechanisms underlying a host of behaviors and cognitive processes. The application of optogenetic techniques to the neurobiological study of memory illustrates how such a revolution occurs. The ability to activate the engram directly, as optogenetic intervention allows, makes it possible to separate the engram from the ecphory. Previously, the conceptual distinction between these components of memory was difficult to reflect in experimental practice.

By manipulating the engram to produce false memories in Ramirez et al. (2013) and Redondo et al. (2014) and to repair forgetting in Ryan et al. (2015), the Tonegawa research group has demonstrated how optogenetic techniques can be used to gain new insight into the neural mechanism(s) responsible for memory errors. Even from these initial studies, optogenetic intervention produces findings that challenge the received view of each of these errors. Many memory scientists have assumed that false memories are the result of a failure to retain memory traces, but the Ramirez et al. and Redondo et al. studies demonstrate the production of false memory via retained, distorted engrams. Similarly, memory theorists have long supposed that disrupting the consolidation process produces forgetting by erasing the original engram. Ryan et al. show that, at least in mice, traces are not erased during consolidation; they merely become inaccessible.

The use of optogenetics to explore the neural mechanisms of memory is still in its early days. It is too early, even, to know whether the results discussed in this article can be successfully replicated and extended. One can hope that the continued development and use of this technique will have implications for the study of cognition well beyond memory. Deisseroth, the pioneer of optogenetics, has repeatedly expressed hope that it could be used to treat people with debilitating psychiatric illness (Colapinto 2015). The advances discussed in this article suggest ways that many cognitive processes could be studied more directly, if the involved brain regions could be identified and activated optogenetically. The consideration of such possibilities invites a host of new challenges into philosophy of mind and neuroscience, perhaps most prominently questions about reduction and levels of explanation. Whether such
challenges are forthcoming is still to be determined. What is clear, however, is that the arrival of this experimental tool has reenergized the neurobiological study of memory, by providing the previously impossible means for distinguishing the engram from the ephory.

REFERENCES


