# 1 The Engram Revisited: On the Elusive Permanence of Memory

**Yadin Dudai**

That memory involves enduring physical changes in the organism has been proposed, using era-dependent metaphors, since antiquity.[1] More than a century ago, Richard Semon introduced the term *engram* to refer to such a change.[2] It gained popularity, however, only when a group of researchers, most notable among them Karl Lashley, embarked on a systematic hunt for engrams, using lesions to determine which parts of the brain impair the ability of animals to form and maintain memories. That search proved futile: "This series of experiments," wrote Lashley in 1950 toward the end of his career, "has yielded a good bit of information about what and where the memory trace is not … I sometimes feel, in reviewing the evidence on the localization of the memory trace, that the necessary conclusion is that learning just is not possible … Nevertheless, in spite of such evidence against it, learning does sometimes occur."[3]

Lashley's disillusion was quickly forgotten, however, or probably suppressed in the collective memory of memory scientists, and the search for the engram was soon revitalized. Reasons were proposed why Lashley failed in his attempts to localize the physical trace of memory, and advanced methodologies were recruited to the game. These included localized brain stimulation, recording of nerve cell activity in the behaving animals, and, ultimately, functional brain imaging in humans. This renewed search yielded an abundance of fascinating data on memory systems in the brain, on the neural circuits that subserve them, and on cellular and molecular mechanisms that might make memory possible.

With this massive accumulation of new data on brain mechanisms of memory in the background, it seems like a proper time to revisit the engram. How enduring is the hypothetical physical change, if at all? What is its relevance to the expression of memory? The *Zeitgeist* in the

neurobiology of memory was until recently that the engram is indeed a lasting change induced by the learning experience, not much unlike the inscriptions on wax tablets proposed by Plato in the dialogue *Theaetetus*. This view is now changing, however, which comes as no surprise to students of human psychology and cognition, who for years have claimed the most intriguing attribute of long-term memory to be its frailty rather than its stability.[4] Yet it is of interest to the science of memory as a whole to appreciate why the aforementioned conceptual change is taking place in brain research.

In analyzing the current transition in the interpretation of the engram, it is useful to spell out at the outset the two major, long-standing hypotheses in the neurobiology of memory. One is the "dual trace hypothesis," the other the "consolidation hypothesis." Derived mostly from the works of William James and Donald Hebb, the dual trace hypothesis posits that memory traces exist in two forms. [5] One is ephemeral (short-term memory, STM), the other long-term and stable (long-term memory, LTM). The consolidation hypothesis posits that for memory to become long-term, it must undergo a maturation process, which renders the trace resistant to some agents and treatments that can impair or erase short-term memory. [6] The consolidation hypothesis further assumed that consolidation occurs just once per item.

The dual trace and consolidation hypotheses are closely related. For the trace to be dual, transition from one form to another has to take place, which is consolidation; and for consolidation to take place, a change in the nature of the memory trace has to be assumed. Both hypotheses embrace a universal unifying concept in biology: Living entities develop and grow. Both consider learning to be an experience-dependent process in which the "teaching stimulus" triggers a local, restricted developmental shift in the relevant areas of the brain, involving local growth processes in interconnected sets of nerve cells. Memory is hence the outcome of growth processes in the neuronal circuits that encode the newly acquired information. Seen this way, brains never reach full maturity (admittedly, not a particularly surprising conclusion, scientific evidence notwithstanding). Edwin Holt epitomized this viewpoint in 1931: "Growth and learning are one continuous process, to the earlier phases of which we give the one name, and to the later … the other."[7] Or, an earlier similar view: "For every act of memory," said Alexander Bain, "…

there is a specific grouping or coordination of sensations and movements, by virtue of specific growth in the cell junctions."[8] Elaborate experience-dependent growth theories paved the way to the discovery that de novo macromolecular synthesis, characteristic of development and growth in all tissues, is indeed required for long-term memory. The concept of trace transition via consolidation hence fits the idea that like organs and organisms, memories mature over time.

It is noteworthy that the term *consolidation* is used in memory research to denote hypothetical memory stabilization processes at different levels of brain organization. Molecular neurobiologists refer to postencoding stabilization of synaptic or cell-wide information storage, which is completed within hours or days after encoding, as "cellular" or "synaptic consolidation." Cellular consolidation is universal and has been identified in all species capable of acquiring long-term memory. But there is also an additional process, called "systems consolidation," which refers to the postencoding reorganization of information in distributed corticohippocampal circuits. This process requires weeks, months, possibly even years to complete.[9] In this chapter, unless otherwise indicated, *consolidation* refers to both synaptic and systems types of memory maturation.

The consolidation hypothesis thus implies two interrelated attributes of long-term memory: *unidirectionality* along the time arrow and *stability* over time. In particular, it was assumed that once consolidation is over, the memory item becomes resistant to a variety of amnesic agents, such as inhibitors of protein synthesis.

## Retrieval Renders Old Memories Malleable Again

Ample evidence from human cognitive psychology indicates that recollection involves reconstruction of information rather than mere replication. In parallel, animal studies have provided intriguing evidence that reactivation of items in long-term memory opens a window of susceptibility to amnesic agents, long after the completion of the postulated post-encoding consolidation. This phenomenon, termed "reconsolidation," was nevertheless practically neglected for decades. There were two reasons for this neglect. First, the interpretation of the data on the apparent susceptibility of reactivated memories to interference in terms of a process of recurrent consolidation was challenged. Second, the dominance of the consolidation

hypothesis tended to interfere with proper discussion of the idea of reconsolidation. Ultimately, the accumulating data made their impact, and in recent years the study of reconsolidation has become a major focus of interest in both human and animal research.[10]

In terms of the cellular and circuit mechanisms, reconsolidation is not a faithful replay of consolidation. Both processes do, however, share dependence on de novo macromolecular synthesis in nerve cells. To date, several boundary conditions have been identified that constrain reconsolidation. These include the degree of "dominance" of the reactivated trace compared to other associations of the same cue (i.e., the ability of that specific trace to control behavior after retrieval); competition with concomitant memory extinction; and, most pertinent to our discussion here, conditions that promote new encoding in or immediately after retrieval.[11]

Close examination of the data and discussions in the field unveils three potential versions of the reconsolidation hypothesis. The "strong version" posits that the regained plasticity applies to all the elements of the original memory and may indeed end up in the erasure of that memory. The "intermediate version" posits that there is a core memory that is stable and unaffected by the reconsolidation, although some stored elements of the original trace can still be modified and even erased. The "weak version" posits that the original memory trace is unaffected, and the transient augmentation of plasticity refers only to new information added to the older memory during or immediately after retrieval. The weak version does not deviate from the classical consolidation hypothesis, as it simply claims that new information consolidates; thus not really "reconsolidation." It is so far unclear which of the other versions fits reality better, the "strong" or the "intermediate." Yet even if upon memory reactivation the core representation becomes susceptible to amnesic agents (i.e., as the "strong version" predicts), related memory associations seem to be spared.[12]

Whereas in laboratory settings reconsolidation is usually unveiled by detecting susceptibility to memory impairment after retrieval, in real life the process might provide an opportunity for the strengthening of the trace. This observation, along with the finding that reconsolidation is promoted by the induction of an encoding state in the retrieval situation, raises the possibility that the role of reconsolidation may be to update memory. That is to say, the process promotes adaptation of the retrieved trace to the retrieval context. However, whereas the consolidation hypothesis postulates

that the original memory is securely consolidated, updating notwithstanding, the reconsolidation hypothesis, even in its "intermediate version" (see above), assumes that at least part of the original trace regains susceptibility to change. Some data support a role for reconsolidation in the updating of long-term memories. The current discrepancy on the role of reconsolidation in updating in different systems and paradigms might be related to boundary conditions on reconsolidation, which are not yet completely understood.

### Malleability in the Absence of Retrieval

Recent data demonstrate that long-term memory (at least up to a few months after encoding) is susceptible to certain amnesic agents, even in the absence of explicit memory reactivation, that is, in the absence of retrieval. These agents are inhibitors of an enzyme, the atypical isozyme of protein kinase C (PKC) called PKMζ. PKCs are molecules composed of a catalytic subunit, which catalyzes the modification of the substrate proteins in vivo, and a regulatory subunit, which inhibits the catalytic subunit by binding to it via a specific part, termed the *pseudosubstrate domain*. In the absence of the regulatory subunit, the enzyme becomes constitutively active, or "autonomous." PKMζ is the autonomous form of PKCζ. In laboratory experiments, PKMζ can be selectively inhibited by a number of inhibitors, notable among them a cell-permeable form of the pseudosubstrate protein sequence, called the zeta inhibitory peptide (ZIP). PKMζ has been reported to be critical to the maintenance of long-term potentiation (LTP), a popular cellular model of learning in the hippocampus.[13] The persistently active PKMζ acts on specific synaptic substrates, leading to modification of the microstructure of the synapse and, ultimately, to a substantial increase in the number of functional postsynaptic receptors for the major transmitter glutamate (particularly receptors of the subtype called "AMPA"). All this culminates in persistent enhancement of synaptic transmission, presumed to encode the experience-dependent alteration in the activity of the specific neuronal circuit, that is, the cellular manifestation of the memory formed by the specific experience.

Long-term spatial information in the hippocampus, subserved by LTP, was shown to critically depend on persistent activity of PKMζ.[14] This was demonstrated by the microinfusion of the selective inhibitor ZIP into the

hippocampus of the behaving rat. Additional forms of hippocampus-dependent memories and some forms of amygdala-dependent memories were also shown to be impaired by the PKMζ inhibitor. Although the hippocampus is indeed well known to play a critical role in some types of memory, it is the neocortex that is considered to serve as the ultimate repository of many types of long-term memory in the mammalian brain. Microinfusion of ZIP into the neocortex was shown to rapidly erase remote memories in the behaving rat. The affected brain area was, however, still able to reacquire a new memory association, implying that information was depleted from the storage apparatus but the apparatus itself was not damaged .[15] These data thus suggest that PKMζ permanently maintains the cellular machinery that embodies long-term memory. When the enzymatic activity is blocked briefly, the experience-dependent synaptic modifications collapse, and so does the specific memory. One possibility is that the target of PKMζ is a synaptic "tag" that is formed when new information is encoded, but is then degraded rapidly by dephosphorylation. In the absence of this tag, although the enzymatic activity recovers from the inhibition, the enzyme can no longer locate the proper phosphorylation site and therefore the tag is not regenerated and memory is lost.

Two main conclusions emerge from the findings concerning the role of PKMζ in long-term memory persistence. First, that some specific inhibitors (e.g., ZIP) can cause rapid, irreversible amnesia even in the absence of explicit memory reactivation. Thus postretrieval "reconsolidation" is not the only window of opportunity in which an item in long-term memory can be modified. And second, neuronal changes that subserve long-term memory are not indelible modifications in synaptic structure, but remain dependent on ongoing enzymatic activity and thus are capable of rapid and dynamic alterations by experimental manipulations.

What might be the role of a mechanism that permits rapid erasure of long-term memory? Several possibilities come to mind. First, in situ, the cellular mechanism that requires persistent phosphorylation by PKMζ might be regulated in selected synapses, possibly in a graded manner, resulting in fast, restricted modulation of local synaptic properties. Such rapid, local modulation of long-term synaptic plasticity might, for example, be useful in the course of fast incorporation of new experience into existing associative knowledge schemas in the neocortex, without necessarily activating other related associations that are accessed at the time of change.[16]

Second, rapid inhibition of PKMζ in specific synapses might result in a rapid shift of synapses to a reduced level of activity or even to a silent state. This might be useful when previous accumulating modifications culminate in catastrophic "freezing" (e.g., a stable local minimum trap) of the computational abilities of the neuronal circuit, a situation that might be remedied by "rebooting." (We can note again the abundance of era-dependent metaphors in discussions of brain function). And third, as some computational models suggest, circuits might saturate, a situation that might benefit from erasure because it releases computational space for processing and storing new information.

Selective inhibitors of PKMζ are, at this time, the only agents found capable of rapidly erasing certain types of long-term and remote memory associations in the mammalian brain in the absence of explicit memory reactivation. Since phosphorylation of synaptic proteins is implicated in many cellular models of memory encoding, and since the phosphorylation of a target protein can be reversed by another type of enzyme, protein phosphatase, further research on protein phosphatase inhibitors may lead to identification of additional types of memory erasers.

## Memories Active and Inactive

The recent findings concerning modifiability of long-term memory have revitalized an alternative conceptual framework to the aforementioned dual-trace and consolidation hypotheses. This alternative conceptual framework portrays memory items in two alternating states: active and inactive.[17] "Active" is the state of the memory trace immediately after encoding and retrieval. Occasionally the memory trace might also become activated independent of encoding and retrieval. Otherwise, the trace is "inactive." Over time, so goes the hypothesis, the trace alternates between the active and inactive states. The data on consolidation and reconsolidation indicate that whenever active, the trace enters a special state ("post-activation state"), in which it is highly plastic and susceptible to interference by amnesic agents. This runs counter to the dual trace hypothesis, which predicts no augmented plasticity after retrieval once consolidation is over.

It is noteworthy that the active/inactive types of models neither nullify the existence of a unique initial consolidation process nor preclude an early maturation phase for each item in memory immediately following

its encoding. As noted above, studies comparing consolidation to reconsolidation show that reconsolidation is not a faithful recapitulation of consolidation. In addition, studies on the role of PKMζ in neural plasticity and memory show that memories are not sensitive to PKMζ inhibitors in the first hours after training. All this implies that the properties of a fresh memory are different from those of an old one. But once long-term memory is established, the active/inactive models assume that the memory is still malleable and not stored as an indelibly consolidated item. Resorting to metaphors, the combination of the dual trace model with the consolidation hypothesis connotes a "storehouse" class of metaphors, [18] whereas the more recent data on the high plasticity of the long-term trace and the "cyclic" models that stem from these data favor a "phoenix" type of metaphor: Occasionally, items in memory get the opportunity to be reborn again and again.

## The Advantage of Instability

The findings that items in long-term memory are prone to change either upon their reactivation in retrieval (i.e., reconsolidation) or even in the absence of such explicit reactivation (e.g., by interfering with persistent activity of the cellular information-keeping machinery), may at first seem counterintuitive. It does seem advantageous to abort the formation of long-term memory in the consolidation window to eliminate new information that is judged by the brain to be superfluous or only of temporary value. But, once information is judged to be valuable for long-term use and hence consolidated, why should it be modified over time? Similar to the answer to any other teleology-driven question in science, the answer to this question as well is supposed to be *a priori* speculative. Still, the intellectual exercise is worth playing. First and foremost, the possibility should not be excluded that this potential frailty of declarative long-term memory reflects an inherent mechanistic shortcoming of the biological system, rather than adaptivity. That said, it is still worthwhile to consider adaptive possibilities. The first that comes to mind is that memories too robust are a potential disadvantage, as they may not fit anymore to guide the proper action and reaction in a changing environment. The updating process, as noted above, is highly valuable. Updating in retrieval can benefit from the existence of the reconsolidation window. Updating outside the time

window of reconsolidation may further facilitate fast incorporation of new experience into existing associative knowledge schemas in the absence of superfluous activation of indirect associations.[19]

The price paid by the organism for this plasticity may be the reduced veracity of stored information. It is noteworthy that for at least one type of memory system, episodic memory, this has been proposed as an advantage rather than an imperfection.[20] It has been postulated that the function of the cognitive system that we dub "episodic memory" is primarily to permit generation of mental time travel and particularly the imagination of scenarios of future events rather than storing the memory of past events. If indeed "episodic memory" is primarily a mental future-time-travel organ, the fact that items in long-term memory change over time is not a disadvantage, but rather an advantage. The reason is that imagination permutates and extends our previous experience, not unlike the recursiveness proposed to underlie the faculty of language; and hence too rigid a memory may lead to poor imagination, one that plays scenarios of the future that are only similar to the past.[21] Thus the elementary neuronal mechanisms that permit recurrent updating of items in long-term memory may have also permitted us to evolve a more effective imagination, clearly a faculty of great phylogenetic value.

## Engrams as Palimpsests

Is the engram permanent then? Given that synapses, cells, and circuits seem to be in constant flux, that experience-dependent modifications in neural systems have ample opportunities to become redone and possibly undone, and that the engram refers to the physical trace formed in encoding, the first answer that comes to mind is no. But this answer is rather simplistic. To arrive at a more realistic one, we need to consider two fundamental issues. The first is the level of details that are valuable for the organism to remember. The second is the distinctiveness or "individuality" of engrams, that is, the distinct set of informational attributes that reflect the unique event that had led to the formation of the engram and are supposed to be "stored" in that specific engram.

Memory systems seem to differ from one another in the resolution of the information that makes the memory item valuable to the organism. For example, some skills are critically dependent on fine details. Yet to

remember superfluous details of events could be counterproductive, as illustrated in the biographies of the real-life mnemonist Solomon Shereshevsky and his fictional counterpart, Funes, from Jorge Luis Borges's fascinating story.[22] That fine details concerning the content and timing of events become unreliable as time goes by is thus not necessarily a design flaw in our memory systems. On this issue, the cognitive evolutionary perspective is different from that of the individual struggling to remember a pertinent detail. We indeed feel uncomfortable and embarrassed when attempting in vain to recall exactly what happened during last year's vacation. But from a phylogenetic point of view, this information is rarely critical. The gist of the experience, or the processed and distilled mental narrative, is usually more important than the accuracy of the details. Narratives can assimilate and parsimoniously represent the valuable impact of experience. Giving up on superfluous details could allow the brain to promote generalization and facilitate appropriate response to both expected and unexpected cues. It is also probably easier on the capacity of the memory system.

In the process of forming mental narratives, engrams merge, losing much of their original individuality. They join the distributed, large and dynamic "society of engrams" that comes to constitute our memory. To consider an engram as a discrete, well-defined long-term physical trace is hence a bit naive. In real life, engrams are palimpsests, reflecting physical traces of many layers of past events.[23] Molecular, synaptic, and cellwide mechanisms, of the types described earlier in this chapter, allow the engrams to do just that.

The permanence of memory traces is hence evasive. Discrete, fine-grained mnemonic traces, assumed to be formed in encoding, are likely to be ephemeral. What persists is their increasingly diluted contribution to memory palimpsests that keep metamorphosing so long as the brain endures. But stripping engrams of their individuality doesn't render the concept of the engram useless. Keeping this concept alive, over a century after Semon had proposed it, reminds us that items in memory are indeed embedded in some type or another of physical change in the biological material and drives us to search for the algorithms used by biological learning machines and for their implementation in identified biological nuts-and-bolts. And as the recent exciting developments in neuroscience demonstrate, the search for the engram can indeed unveil surprising new

and useful properties of memory systems at all levels, from the molecular to the behavioral.

## Notes

1. Yadin Dudai, *Memory from A to Z: Keywords, Concepts, and Beyond* (Oxford: Oxford University Press, 2002).

2. Richard Semon, *The Mneme* (1904; reprint, London: George, Allen and Unwine, 1921). Throughout this chapter, the term *engram* will be used interchangeably with *memory trace*, hence implying that the *trace* is a physical change in the nervous system.

3. Karl S. Lashley, "In Search of the Engram," *Symposia of the Society for Experimental Biology* 4 (1950): 454–482.

4. Daniel L. Schacter, ed., *Memory Distortion* (Cambridge, MA: Harvard University Press, 1995).

5. William James, *The Principles of Psychology* (1890; reprint, New York: Dover, 1950); Donald O. Hebb, *The Organization of Behavior: A Neuropsychological Theory* (New York: Wiley, 1949).

6. Yadin Dudai, "The Neurobiology of Consolidations, or, How Stable Is the Engram?" *Annual Review of Psychology* 55 (2004): 51–86.

7. Edwin B. Holt, *Animal Drive and the Learning Process* (New York: Holt, 1931).

8. Alexander Bain, *Mind and Body: The Theories of Their Relation* (London: Henry King, 1872).

9. James L. McClelland and Nigel H. Goddard, "Considerations Arising from Complementary Learning Systems Perspective on Hippocampus and Neocortex," *Hippocampus* 6 (1996): 654–665.

10. Karim Nader and Oliver Hardt, "A Single Standard for Memory: The Case for Reconsolidation," *Nature Reviews Neuroscience* 10 (2009): 224–234.

11. Richard G. M. Morris et al., "Memory Reconsolidation: Sensitivity of Spatial Memory to Inhibition of Protein Synthesis in Dorsal Hippocampus during Encoding and Retrieval," *Neuron* 50 (2006): 479–489.

12. Jacek Debiec et al., "Directly Reactivated, but Not Indirectly Reactivated, Memories Undergo Reconsolidation in the Amygdala," *Proceedings of the National Academy of Sciences USA* 103 (2006): 3428–3433.

13. Douglas S. Ling, Larry S. Benardo, and Todd C. Sacktor, "Protein Kinase Mzeta Enhances Excitatory Synaptic Transmission by Increasing the Number of Active Postsynaptic AMPA Receptors," *Hippocampus* 16 (2006): 443–452.

14. Eva Pastalkova et al., "Storage of Spatial Information by the Maintenance Mechanism of LTP," *Science* 313 (2006): 1141–1144.

15. Reut Shema, Todd C. Sacktor, and Yadin Dudai, "Rapid Erasure of Long-Term Memory Associations in Cortex by an Inhibitor of PKMz," *Science* 317 (2007): 951–953.

16. Dorothy Tse et al., "Schemas and Memory Consolidation," *Science* 316 (2007): 76–82.

17. Donald J. Lewis, "Psychobiology of Active and Inactive Memory," *Psychological Bulletin* 86 (1979): 1054–1083.

18. Henry L. Roediger III, "Memory Metaphors in Cognitive Psychology," *Memory & Cognition* 8 (1980): 231–246.

19. D. Tse et al., "Schemas and Memory Consolidation"; J. Debiec et al., "Directly Reactivated, but Not Indirectly Reactivated, Memories Undergo Reconsolidation in the Amygdala."

20. Yadin Dudai and Mary Carruthers, "The Janus Face of Mnemosyne," *Nature* 434 (2005): 567.

21. Marc D. Hauser, Noam Chomsky, and William T. Fitch, "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" *Science* 298 (2002): 1569–1579.

22. Alexander. R. Luria, *The Mind of a Mnemonist* (London: Jonathan Cape, 1969); Jorge Luis Borges, "Funes, His Memory," in *Collected Fictions* (1944; reprint, New York: Viking: 1998).

23. Yadin Dudai, *Memory from A to Z*.