PROJECT MUSE®

Memory

Radstone, Susannah, Schwarz, Bill

Published by Fordham University Press

➡ For additional information about this book

13. Memories Are Made of This

Steven Rose

"If any one faculty of our nature may be called more wonderful than the rest, I do think it is memory. There seems something more speakingly incomprehensible in the powers, the failures, the inequalities of memory, than in any other of our intelligences. The memory is sometimes so retentive, so serviceable, so obedient—at others, so bewildered and so weak—and at others again, so tyrannical, so beyond controul!—We are to be sure a miracle in every way—but our powers of recollecting and forgetting, do seem peculiarly past finding out."

Thus Fanny Price, Austen's long-suffering heroine in *Mansfield Park.* It took more than half a century from the writing of that novel for psychologists to attempt to bring the discipline of the laboratory to bear on this tyrannical and uncontrollable memory, and nearly another before it was to become subject to the molecular probes and confident claims of a resurgent neurobiology. Today's neuroscience seizes not on Austen but on the poet Emily Dickinson for its claim to knowledge, its leading figures gleefully quoting her verse:

> The Brain is wider than the sky,
>   For, put them side by side,
> The one the other will contain
>   With ease, and you beside.

Yet after my own lifetime of research, in charting the biochemical cascades and cellular remoulding that even the simplest of learning experiences seems to generate in my young chicks (the experimental animals which have participated, albeit involuntarily, in more than three decades of my study of memory) I have to confess that I still don't feel we have done more than deepen some of its mysteries. Fifteen hundred years before Fanny Price, in his *Confessions*, St. Augustine listed some

of the phenomena that needed explaining: memory, he says, is a "spacious palace, a store-house for countless images." But memory is capricious. Some things come spilling from the memory unwanted, while others are forthcoming only after a delay. Memory enables one to envisage colors even in the dark, to taste in the absence of food, to hear in the absence of sound. "All this goes on inside me in the vast cloisters of my memory." Memory also contains "all that I have ever learnt of the liberal sciences, except what I have forgotten . . . innumerable principles and laws of numbers and dimensions . . . my feelings, not in the same way as they are present to the mind when it experiences them, but in a quite different way," and things too, such as false arguments, which are known not to be true. Further, he points out, when one remembers something, one can later remember that one has remembered it. No wonder that the mind seemed to soar outside the physical confines of mere brain-goo. For Augustine, unlike Emily Dickinson, it is the mind, not the brain, that is wider than the sky, and I am inclined to agree.

What the experimental sciences have tried to do, of course, is to operationalize memory, to reduce and control "learning experiences" in such a way that their parameters could be studied. The process was begun by Hermann Ebbinghaus, whose book, *Uber das Gedächtnis* (*On Memory*), published in 1885, broke new ground by asking whether there were general laws of memory formation. To explore these general laws, he invented the simple technique that in various forms has been a staple psychologist's tool ever since—that of the nonsense syllable, a series of three letter sets each composed of a vowel between two consonants, as for instance: HUZ; LAQ; DOK; VER; JIX. Using himself as subject, Ebbinghaus then explored the conditions required to remember such lists; numbers of readings, spacing and so forth, until he could make two errorless readings of the entire list. Once the list was learned, he could then test how successful he was at recalling it at various subsequent times, whether minutes or days later. To quantify this process of recall, all that he had to do was to note how many readings of the list were necessary, at any given time after it had been learned, to once again be able to repeat it without error.

A number of general rules could be derived from such observations. For instance, in any such list of a dozen nonsense syllables, some are easier to remember than others—in particular, those at the beginning and at the end of the list. These are the so-called primacy and recency effects. They may seem obvious when described so simply, but what Ebbinghaus did was to demonstrate clearly that, in this case at least, common sense was supported by science. In addition, he showed that if a list is once learned, it becomes easier to relearn subsequently. A comparison of the number of trials required to learn it the second time with those required first time round provides a calculation that has become known in the psychology literature as *savings*—the measure of memory. The use of the savings score enables one to specify more precisely the loss and stabilization of memory with time. Ebbinghaus found that most of the memory loss occurred within the first minutes after training; once the memory had survived that hurdle it seemed much more

stable, leading to the temporal distinction between short- and long-term memory that has become a staple of subsequent research.

Ebbinghaus's was the first step in developing the taxonomy of memory that has provided much of the focus of subsequent psychological research. In the 1930s Frederic Bartlett famously showed how the content of even remembered items becomes transformed and simplified over time. And in the 1980s and 90s, Alan Baddeley drew a distinction between working memory—that is, memory dredged up from past experience for current use, so to say—and the more deeply stored reference memory. Meanwhile, based in part on evidence from patients with identifiable brain lesions, Endel Tulving, and later Larry Squire, added a further taxonomical distinction, that between various classes of memory. Procedural memory is remembering *how* to do something—to ride a bicycle, for instance. Declarative memory is remembering *that*—that a particular two-wheeled drivable object is called a bicycle. Declarative itself becomes divided into semantic (Augustine's numbers and dimensions) and episodic or autobiographical memory (recall of episodes in one's own life).

For a neurobiologist, the crucial questions are about how these forms of memory are instantiated in the brain. Do the categories reflect the engagement of different brain regions, and different molecular processes, or are they higher-level distinctions, without matching brain correlates? Until recent decades, the only effective way of addressing these questions in humans was by observing the effects of various forms of brain lesion and disease on memory. Classical disease-induced losses of memory, notably from what used to be called senile dementia but is now more frequently called Alzheimer's disease, can't answer these questions because the brain damage such maladies cause is both progressive and very general. But some consequences of strokes or accidents—or surgically induced damage—can be instructive. The most famous case of iatrogenic, or surgeon-induced, memory loss is an epileptic patient, known to every neuroscientist by his initials, HM, who was operated on in the 1950s to remove regions of his temporal cortex and hippocampus so as to eliminate the epileptic focus. The result was a catastrophic loss in his ability to transfer memory from short to long term. HM, who has continued to be a subject of research for the subsequent half-century, can remember events up to the time at which he was operated on, but forgets any new experience within minutes. Although he can show some procedural learning of new skills, he cannot retain declarative—especially episodic, autobiographical—memory. Events, as he himself puts it, simply fade away; he says "every day is by itself." This observation, soon matched by studies in animals, suggests that the hippocampus has a crucial role to play in the registration of new experience, and that without it and adjacent brain regions, items can no longer be transferred into longer-term memory.

Within the last decades, the possibility of studying ongoing memory processes in the living human brain has been transformed by the advent of new technologies, notably the

windows opened by functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG). The former makes possible the measurement of changes in blood flow to small regions of the brain, the assumption being that the higher the rate of blood flow the more active that region is under any particular circumstance, such as performing a learning or memory task. The latter takes advantage of the fact that signaling within the brain is primarily electrical and that electrical current flow is accompanied by tiny changes in the magnetic field surrounding the current. Both techniques require rather formidable instrumentation; fMRI is better at localizing sites of change, MEG is most helpful in charting the temporal dynamics, making it possible to plot changes in brain activity millisecond by millisecond.

Two examples reveal what can be learned from such techniques. Eleanor Maguire and her colleagues studied London taxi drivers who were asked, while under fMRI, to recall a complex journey within the city. The act of recalling the route activated their hippocampi.[1] In our own experiments, using MEG, we took subjects on a virtual supermarket tour, asking them to make choices of items to purchase based on their past experience and preferences. Faced with a choice, say of three brands of coffee, subjects took about two seconds to press a key indicating their choice. But in those two seconds, there was a flurry of brain activity. Within 80 milliseconds, the visual cortex became active; by 300 milliseconds, the left inferotemporal cortex, assumed to be a site of memory storage. At 500 milliseconds, Broca's area, a region associated with speech, was engaged, as the subjects silently vocalized the range of choice items—and at 800 milliseconds, as they made their final decision as to which item they preferred—assuming they preferred any— the right parietal cortex, associated with affect-laden decisions, was active. These dynamics reveal the many regions of the brain involved in even a simple act of episodic and semantic memory; even primary sensory regions like the visual cortex are more active when people are performing a memory-related task than when they see the same images but are asked simply to make a cognitive choice—for instance of which item is the shortest of those displayed. Thus Baddeley's working memory does not seem to be simply localizable to one brain region.

Revealing though such studies are—and they are as yet in their infancy as the techniques and instrumentation mature—there are limits to the types of answer they can provide. If learning and the making of memory demand cellular and molecular changes, these cannot be studied except in animals. To do this demands developing models of learning and memory in these animals that may serve in some sense as a surrogate for the same processes in humans. The doors to such an approach were opened early in the last century by Ivan Pavlov's well-known experiments with dogs. Pavlov trained them to associate the ringing of a bell with the arrival of food, and hence to salivate (a learned or, in the jargon, conditioned reflex). In the 1930s B. F. Skinner developed a different learning model (operant conditioning), in which animals had to perform some act, such as pressing a lever to obtain food, or to escape an electric shock. If after one or more trials the

animal's behavior changes appropriately—for instance, by salivating to the bell, pressing the lever sooner in response to a signal, or running a maze faster and with fewer errors—the animal is said to have learned from the experience. And, when it performs the learned task in an error-free way, it is said to be remembering the experience. The unspoken assumption is that whatever the brain processes that are involved in such changes in the animal's behavior may be, they are similar to those occurring in the human brain when we learn and remember. The Skinnerian view was that all creatures learn and remember in the same way, that there are general laws of learning that are as universally applicable as the gas laws or gravitation in physics.

Of course, there are problems with such an assumption. What and how an animal will learn is species-specific. Some food-storing birds, such as scrub jays, can recall during winter the many thousands of sites at which they cached edible seeds the previous summer. Others—songbirds like zebra finches—cannot learn such tasks but readily acquire new songs. Further, an animal can only inform a human experimenter that it is learning or remembering by way of some change in its performance of a task. It may "remember" its previous experience but choose not to perform the task appropriately—a point made in the 1950s in a famous critique of Skinnerian approaches, a paper called simply "The Misbehaviour of Animals." Despite heroic attempts at complex experimental designs, the taxonomic distinctions between procedural and declarative learning and memory are always going to be confounded in animal studies.

Yet if learning from some new experience results subsequently in a change in the behavior of the animal when presented with a similar situation, one must assume that something has changed in the brain to support the changed behavior. This inferred intervening variable is regarded as a memory "store," "trace," or "engram," which is formed when learning is taking place and reactivated when that learning is later recalled. The challenge for neurobiologists then became that of identifying the anatomical, cellular, molecular, or physiological nature of the trace. The temporal distinction between short- and long-term memory, the evidence that short-term memory is labile and easily disrupted, whereas long-term memory seems relatively protected, suggested that it must depend on some structural remodeling of the patterns of neural connection within the brain, engraving the memory in the brain in a manner analogous to that of inscribing a magnetic trace on a tape or a CD that can subsequently be replayed, invoking the original material. The seductive metaphorical power of computer "memory" has been influential in shaping thought on this question.

The myriad nerve cells in the brain (a hundred billion in the human cortex alone) communicate by way of up to ten-thousandfold (a hundred trillion) more junctions, known as synapses. It is at the synapses that electrical signals traveling down one nerve axon trigger the release of chemical signals—neurotransmitters—that in turn carry the message across a small gap to an adjacent nerve cell, stimulating a response in the second cell. Maybe learning results in some change in synaptic connections, so as to create novel

signaling pathways? In 1948 the Canadian psychologist Donald Hebb framed the hypothesis that has shaped all subsequent biochemical and physiological research in the field, that learning involves the remodeling of such synaptic junctions. In his own words:

> Let us assume then that the persistence or repetition of a reverberatory activity (or "trace") tends to induce lasting cellular changes that add to its stability. The assumption can be precisely stated as follows: *When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.*
>
> The most obvious and I believe much the most probable suggestion concerning the way in which one cell could become more capable of firing another is that synaptic knobs develop and increase the area of contact between the afferent axon and efferent [cell body]. There is certainly no direct evidence that this is so. . . . There are several considerations, however, that make the growth of synaptic knobs a plausible perception.[2]

By the 1960s, neuroscientists had become sufficiently confident in the power of their technologies to attempt to verify Hebb's hypothesis experimentally. But after an initial burst of enthusiasm the field became mired in controversy. Claims that training rats on some simple task resulted in increases in RNA (ribonucleic acid) and protein synthesis in their brains and, even more extravagantly, that when the RNA was extracted and injected into the brain of a recipient, the memory was transferred too, achieved great publicity but were technically flawed. Research funds dried up and even to suggest that one was working on the biochemistry of memory became somewhat disreputable. More patient experiments in the 1970s began to revive confidence, and Hebb's plausible perception became tangible evidence. Two disparate approaches helped. One seems very remote from memory as we might understand it. The physiologists Tim Bliss and Terje Lomo placed stimulating and recording electrodes into cells in the rat hippocampus and found that if they fired a train of electrical pulses into the cells, their output properties were permanently modified; the cells showed a "memory" of their past experience.[3] The phenomenon, called long-term potentiation, became intensely studied as either a mechanism or a model for memory over the succeeding decades. At around the same time, the psychiatrist-turned-neuroscientist Eric Kandel began exploring the physiological properties of the neurons in the giant sea slug *Aplysia californica*. Aplysia has two useful properties. One is that it can be trained on a simple task, to contract its gills (rather as a land slug rolls into a ball) in response to a jet of water being applied to its tail. The second is that many of its nerve cells are giant and the "same" cell is easily identifiable from slug to slug. Kandel was able to map the neural circuitry involved in the withdrawal reflex and to identify some specific synapses whose electrical properties and biochemistry changed as the slug learned. With

a reductionist rhetorical flourish, Kandel offered the research community "memory in a dish."[4]

Over the decades that followed, evidence from a variety of labs, including my own, showed that indeed, when an animal—in my case a young chick—is trained on some novel task, there are increases in the size and strength of specific synaptic connections in particular brain regions. Under the microscope, the connections are structurally larger and the efficacy of the neurotransmitters within them is enhanced. The experimental problem was to prove that these changes were in some way associated with the storage of the putative memory trace, rather than a consequence of other aspects of the task and its performance. For example, in the task we use, young chicks are offered a small bright bead. Almost invariably, they will peck at such a bead within a few seconds of seeing it. If the bead is made to taste unpleasant (we dip it in a rather bitter, curryish-tasting liquid), the chick will peck once, and then demonstrate its distaste by shaking its head energetically and wiping its bill on the floor of its pen. If it is subsequently—any time up to several days later—offered a similar but dry bead, the chick will not peck it, but back away, sometimes replicating the earlier pattern of head shaking and bill wiping. We infer that the chick has learned, after a single experience, that this particular color, shape, and size of bead tastes unpleasant—at least in this specific context—and that when the bead is presented once more, this memory is reactivated. For the initiates, this is described as one-trial passive avoidance learning.

The passive avoidance task has a number of experimental advantages. It is quick and reproducible and builds upon a normal aspect of the young chick's behavioral development—that is, to spontaneously explore its environment by pecking at small objects. Because the training event—the peck at the bead—takes only a few seconds, one can readily separate the immediate consequences of the bitter taste from the subsequent cascade of events during the transitions between shorter- and longer-term memory. Advantages have corresponding disadvantages. Is what we discover about learning in such a young animal, where the brain is developing rapidly, relevant to learning in adulthood? Do the molecular events involved when a chick learns in a single trial not to peck a bead in any way correspond to those during the many trials a rat needs to learn to run a maze—still less those when a child learns the names of the days of the week or what to expect on its birthday?

Even setting these queries aside, can we be sure, even for the chick, that the change we find in the synapses is actually some form of memory trace? That is, that it is a necessary, sufficient, and exclusive change in the brain that in some way "represents" the memory, enabling it later to be recalled? Could the change not have occurred simply as a result of some aspect of the initial experience, such as the taste or sight of the bead, or the learned motor activity of pecking? Or, as we cannot know whether the chick has learned the task without testing it, maybe it is a consequence of the recall experience rather than the learning itself? I don't intend here to reprise the decade-long series of

control experiments that enabled us to distinguish between non-specific experience-induced and learning-induced changes. I have discussed these at some length in my book *The Making of Memory*.[5] But it may be of more than merely technical interest to outline the sorts of approach one can use.

There are broadly two approaches to identifying the molecular processes, such as passive avoidance, that occur in the minutes to hours following training on a simple task and that are presumed to be required for the maintenance of short-term memory and to underpin the transition to long-term memory, a process called memory consolidation. One can train the animal on the task and look for changes in some putative biochemical measure—the activity of an enzyme, the concentration or rate of synthesis of a molecule. Or one can attempt to disrupt the consolidation process by administering an inhibitor—an antimetabolite, or drug known to block a specific biochemical process believed to be necessary for consolidation. If the drug blocks such a process, then the animal should subsequently not recall the task; that is, it should show a specific amnesia. Observing the changes in the suspected biochemical measure over time or the time window during which the administered amnesic agent is effective makes it possible to plot a temporal sequence of molecular events—a biochemical cascade—occurring over the hours following training that seem to culminate in the lasting modulation of synaptic strengths. Since the 1980s I have used both methods in tandem in elucidating this cascade in my chicks.

Within the minutes following the onset of the training experience, there are changes in the release of neurotransmitters at the synapses in specific brain regions. As well as activating the postsynaptic nerve cell to fire, these increases also stimulate a wave of biochemical activity in the cell, which in due course results in the synthesis of a family of proteins, called cell adhesion molecules, destined to be transported to the synapses. Cell adhesion molecules are a bit like Velcro. They are located in the cell membrane, for instance at the synapse, with one end (the Velcro end) sticking out into the space between one nerve cell and the next, holding the two sides, pre- and post-synaptic, together. The newly synthesized adhesion molecules that are produced as a result of the training experience are dispatched to the activated synapses (a process that takes some 4–6 hours in chicks and rats), and inserted into their membranes, altering the strength of connections between the two sides of the synapse. This would seem precisely to confirm Hebb's hypothesis for how memory might be coded and stored in the brain.

The distinguished Nobel Prize–winning biochemist, Hans Krebs, in whose Oxford lab I was based during a postdoctoral period in the early 1960s, once told me that for every biological problem, God had chosen an appropriate organism in which to tackle the problem. I have argued that, for the study of the molecular processes involved in memory formation, the chick is indeed God's organism. Others have made different choices, ranging from fruit flies and sea slugs to the more familiar laboratory rats and mice. In 2001, the Nobel Committee opted for the slug—although the prize they gave its developer, Eric

Kandel, was not so much for his memory work with the slug as for his studies of its neurotransmitters. What is interesting and encouraging is that despite the differences and learning paradigms, a sequence of broadly similar molecular processes has been shown to occur in the brains or nervous systems of these varying species during and following the training experience.

So can we conclude that we have found the engram—or at least identified the processes whereby engrams are constructed? The suggestion that memories are encoded in terms of changed synaptic connectivities has certainly proved attractive to a new breed of researcher, who call themselves computational neuroscientists, interested in making mathematical and computer models for how learning might occur in a distributed neural network connected by a mesh of synapses. In such a theoretical network, each memory (or association) is represented by activity in a specific set of synapses, a unique pattern, but any one synapse can be involved in many different such associations. On this basis, and estimating the number of synapses that it contains, Edmund Rolls has calculated that the hippocampus can store some 36,500 memories.

But a calculation of this sort is based on a prior set of assumptions: that biological memories can be decomposed into isolated monads and measured in terms of the bits and bytes with which computer people calculate the power of their machines. It is this that is so unrealistic; how many bits of information does the variety of memories listed by St. Augustine require? For that matter, how many bits of information do my chicks need to remember to avoid pecking at a small red bead but know that it is safe to continue pecking at a yellow one? The chick categorizes the experience of pecking the bitter bead in terms of the color, shape and size of the bead, the context in which it was pecked, its own past experience of pecking other beads, and probably many other features as well, any one of which may provide the cue for its subsequent behavior. I am far from sure that, for the chick, this complex of meanings within which any subsequent sight of and response to the bead is embedded is simply decomposable into information theory's bits. Indeed, this theoretical concern is rapidly confirmed by experiment. The linear cascade that the biochemical and pharmacological experiments demonstrate, leading from transient changes in the release of neurotransmitters to seemingly permanent structural changes in the synapse, was no sooner established than paradoxes began to appear in the data. Memory traces apparently firmly located in one brain region seem over the subsequent hours and days to migrate to others—as indeed might have been suspected from HM's experience. His hippocampal damage did not erase old memories, only prevent new ones being formed. Furthermore, there is no single site for "the memory" as if it constituted a discrete entity. The MEG experiment I referred to above shows that many brain regions are involved in the dynamic process of recalling and responding to prior experience. And even for my chicks we have been able to show that different aspects of the memory of the bitter bead—its color, shape, size—engage different ensembles of nerve cells and synapses distributed in different regions of the brain.

Furthermore, memory involves more than just synapses, more even than just brains. How well a person or a chick learns and remembers depends on many other aspects of body state. Alertness and attention depend on physiological processes such as blood flow and hormonal level. Memory involves emotion as well as cognition, and hormones produced outside the brain, notably adrenaline and its neurotransmitter relative noradrenaline, are engaged in determining what is remembered. When chicks peck a bitter bead, a surge of steroid (corticosterone, the chick's equivalent of cortisol in humans) is released into the bloodstream. Too little or too much corticosterone, and the chick will not remember the experience and will peck the previously bitter bead when tested later. In this sense, learning and remembering—memory—is a property not of individual synapses or nerve cells or brains but of the entire organism, the person.

Nor is this all. Hebb's model is one of learning, of what happens when an animal, or human, registers some new experience. Implicit in it is also a theory of recall: that remembering the experience involves reactivating the novel pathways that learning has generated. Memories are stored as in computer files, and remembering would seem to be no more than pulling these files out of deep storage and reopening them. But this mechanical model won't do. Each act of recall is itself a new experience. Reactivated memories are subtly changed each time we recall them. Classroom experiments beautifully illustrate what we all know to be the case. Thus in the aftermath of the disaster that destroyed the *Challenger* space shuttle and killed the astronauts on it, a group of psychology students were encouraged to write down their recollections of the event. The records were stored, and a year later they were asked to write the account again. The huge discrepancies between their first and second accounts indicated just how labile memories of quite dramatic events are. Far from passively recording the past, we in our memories actively reconstruct it.

Very recently, neurobiology has begun to catch up with common experience and the psychologists. Many labs, including our own, have now shown that when an animal is given a reminder of a previously learned experience, the memory becomes labile once more and can be disrupted by drugs and biochemical inhibitors rather as it can be during initial consolidation. Some researchers have begun to speak of this as ''reconsolidation.'' However, the temporal dynamics of reconsolidation are rather different from those of consolidation; different brain regions are involved, and the biochemical changes do not exactly recapitulate those of consolidation.

Of course, being reminded of a past experience is itself to some extent a novel experience. We don't step into the same stream twice, and memory depends on history. That neurobiologists have only recently come to realize this shows just how blinkered and reductionist their—our—paradigms have been. We are trapped by the experimental need for simple and reproducible designs, for operationalizing our definitions of ''learning,'' ''memory,'' and so on as if these complex processes could be trapped within small boxes,

sealed off from everything else that is going on in a living, behaving, learning, and remembering organism throughout every moment of its existence. Our experiments capture only a small part of such complexity, and we are at fault if we mistake this small part for the whole.

Half a century ago, neuroscience saw the brain as composed of discrete centers, regions responsible for vision, audition, pain, memory, and so forth. Superimposed on all these different regions was a super-coordinating center, the association cortex. Separate regions reported upward to this coordinator, which assembled them and instructed the motor regions of the brain how to respond. This homunculus inside the head was the source of identity, individuality, the ''I'' located a few centimeters behind our eyes.

Alas for simplicity, there is no such homunculus. As Gertrude Stein said about, I believe, Oakland, there is no there there. Brains don't have a central processor, a super-manager controlling everything. Rather they are distributed networks of cellular ensembles, richly interconnected, which between them create the illusion of coherent experience that we all in our normally functioning moments share. The enigma of memory, as with so many aspects of brain processes, seems to be that it is both localized and nonlocalized. Remembering is at once sure and certain, as when we recall the names of the days of the week, or mount and ride off on a bicycle for the first time in many years, and as evanescent and elusive as a soap bubble, as when we try to remember the first moment we saw a lover and compare our own memory of that event with his or hers.

Fanny Price was surely right. Which is why we neurobiologists of memory must from time to time come out of our labs, reflect on our own varied procedural, declarative, episodic, and autobiographical memories, and turn to the work of those philosophers, poets, and novelists who can illuminate and interpret our experience so much more richly and meaningfully than can the most ingenious experimenter.