Middlebury

Clinton Cave, PhD
Assistant Professor of Neuroscience
Neuroscience Program
Middlebury College
ccave@middlebury.edu

276 Bicentennial Way, Rm 351
Middlebury, VT 05753
802.443.5450
middlebury.edu

June 14, 2023

Greetings, Alumni College Students!

I'm thrilled that you'll be joining me at Bread Loaf for Our Changing Brain: The Neuroscience of Learning! When picking a subject for the course, I was trying to remember what was most exciting for me when I first started studying neuroscience as an undergraduate. There were lots of interesting topics to be sure, but I was fascinated by the research on memory. Our brain's ability to create, store, and recall memories is such a core component of our conscious experience and our identity. It was truly fascinating to learn about the biological mechanisms (the physical changes that occur in the brain) that allow each of us to experience our unique lives. Importantly, the more we know about those mechanisms, the more we'll be able to help those with memory disabilities. In fact, a great deal of what we know about memory comes from individuals who have suffered brain injuries and have courageously chosen to share their experiences with researchers. I'm excited to be able to share some of their stories with you.

Our course will give you an introduction to neuroscience as a discipline—stated simply, the study of how the nervous system functions to support behavior in health and disease. We'll do a brain dissection to see how the different structures of the brain fit together and where the major functions reside. Then we'll investigate individual neurons (a type of brain cell) and see what makes them special compared to other cells in the body. This will give us the foundation to start our work on memory. We'll focus on three types of memory (though there are many others). Fear memory is a type of emotional memory for traumatic or painful events. Spatial learning is a memory of our surrounding environment that we use to navigate the world. Next, we'll discuss language acquisition and the memory required for speech. My two-year-old son is right in the middle of this process! In our last session we'll explore some of the newest research on memory, which deals with the experimental manipulation of memory in rodents. We now have the genetic tools to artificially install and suppress learned behaviors in mice! Though we are a long way from doing so in humans, we'll consider some of the moral and clinical implications of this research.

To help you prepare, I've assembled a course reader that will support each of our sessions based on the schedule below. They provide an overview of the brain and neuronal function (Linden Ch 1 and 2), a historical account of memory research (Kean Ch 10), a description of memory at the molecular level (Linden Ch 5), a personal narrative of someone living with aphasia (Weiss 2023), and a discussion of recent memory experiments in rodents (Svoboda 2017). I encourage you to read through them this

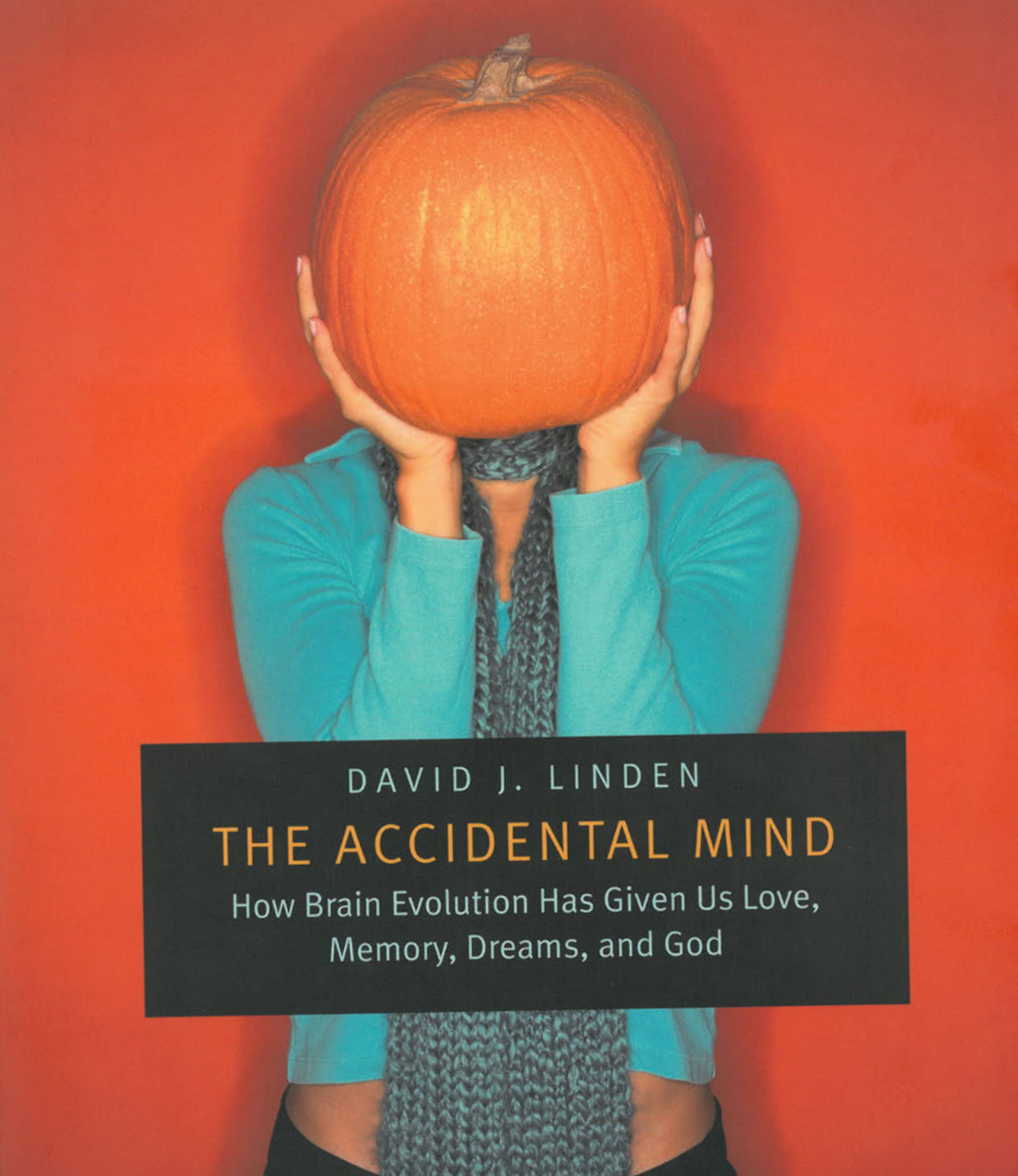
summer without worrying about memorizing all the details. Look for big-picture themes and write down any questions that arise. I hope by the end of our work together, you'll have useful answers to the questions below, and a greater appreciation of what makes your memories truly special!

Cheers,

Clinton Cave

Our Changing Brain: The Neuroscience of Learning
Course Schedule

Session	Topic	Reading
1	Introductions—What is neuroscience?	Linden Ch 1
2	What are the parts of the brain?	Linden Ch 1
3	What is a neuron?	Linden Ch 2
4	How do we utilize our memory?	Kean Ch 10
5	How do our brains create memories?	Linden Ch 5
6	How do we find our way home?	Linden Ch 5
7	How do we learn what words to use?	Weiss 2023
8	Are all our memories real?	Svoboda 2017


DAVID J. LINDEN
THE ACCIDENTAL MIND
How Brain Evolution Has Given Us Love, Memory, Dreams, and God

THE ACCIDENTAL MIND

The large brain, like large government, may not be able to do simple things in a simple way.

—Donald O. Hebb

Now, the president says that the jury is out on evolution . . . Here in New Jersey, we're countin' on it.

—Bruce Springsteen

Prologue

Brain, Explained

THE BEST THING about being a brain researcher is that, in a very small number of situations, you can appear to have the power of mind reading. Take cocktail parties. Chardonnay in hand, your host makes one of those introductions where he feels compelled to state your occupation: "This is David. He's a brain researcher." Many people are wise enough to simply turn around at this point and go looking for the bourbon and ice. Of those who stay behind about half can be counted on to pause, look heavenward, and raise their eyebrows in preparation for speech. "You're about to ask if it's true that we only use 10 percent of our brain, aren't you?" Wide-eyed nodding. An amazing episode of "mind reading."

Once we get past the 10-percent-of-the-brain thing (which, I should mention, has no basis in reality), it becomes clear that many people have a deep curi-

osity about brain function. Really fundamental and difficult questions come up right away:

"Will playing classical music to my newborn really help his brain grow?"

"Is there a biological reason why the events in my dreams are so bizarre?"

"Are the brains of gay people physically different from the brains of straight people?"

"Why can't I tickle myself?"

These are all great questions. For some of them, the best scientific answer is fairly clear and for others it is somewhat evasive (me, in my best Bill Clinton voice: "What exactly do you mean by "brain"?). It's fun to talk to non–brain researchers about these kinds of things because they are not afraid to ask the hard questions and to put you on the spot.

Often, when the conversation is over, people will ask, "Is there a good book on brain and behavior for a nonspecialist audience that you can recommend?" Here, it gets tricky. There are some books, such as Joe Le Doux's *Synaptic Self,* that do a great job on the science, but that are rough sledding unless you've already got a college degree in biology or psychology. There are others, such as Oliver Sacks's *Man Who Mistook His Wife for a Hat* and V. S. Ramachandran and Sandra Blakeslee's *Phantoms in the Brain* that tell fascinating and illuminating stories based on case histories in neurology, but that really don't convey a broad understanding of brain function and that largely ignore molecules and cells. There are books that talk about molecules and cells in the brain, but many of them are so deadly dull that you can start to feel your soul depart your body before you finish the very first page.

What's more, many books about the brain, and even more shows on educational television, perpetuate a fundamental misunderstanding about neural function. They present the brain as a beautifully engineered, optimized device, the absolute pinnacle of design. You've probably seen it before: a human brain

lit dramatically from the side, with the camera circling it as if taking a helicopter shot of Stonehenge and a modulated baritone voice exalting the brain's elegant design in reverent tones.

This is pure nonsense. The brain is not elegantly designed by any means: it is a cobbled-together mess, which, amazingly, and in spite of its shortcomings, manages to perform a number of very impressive functions. But while its overall function is impressive, its design is not. More important, the quirky, inefficient, and bizarre plan of the brain and its constituent parts is fundamental to our human experience. The particular texture of our feelings, perceptions, and actions is derived, in large part, from the fact that the brain is not an optimized, generic problem-solving machine, but rather a weird agglomeration of ad hoc solutions that have accumulated throughout millions of years of evolutionary history.

So, here's what I'll try to do. I will be your guide to this strange and often illogical world of neural function, with the particular charge of pointing out the most unusual and counterintuitive aspects of brain and neural design and explaining how they mold our lives. In particular, I will try to convince you that the constraints of quirky, evolved brain design have ultimately led to many transcendent and unique human characteristics: our long childhoods, our extensive memory capacity (which is the substrate upon which our individuality is created by experience), our search for long-term love relationships, our need to create compelling narrative and, ultimately, the universal cultural impulse to create religious explanations.

Along the way, I will briefly review the biology background you will need to understand the things I am guessing you most want to know about the brain and behavior. You know, the good stuff: emotion, illusion, memory, dreams, love and sex, and, of course, freaky twin stories. Then, I'll try my best to answer the big questions and to be honest when answers are not at hand or are incom-

plete. If I don't answer all of your questions, try visiting the book's website, accidentalmind.org. I'll strive to make it fun, but I'm not going to "take all the science out." It will not be, as you might find on a label at Whole Foods, "100 percent molecule free."

Max Delbrück, a pioneer of molecular genetics, said, "Imagine that your audience has zero knowledge but infinite intelligence." That sounds just about right to me, so that's what I'll do. Let's roll.

Chapter One

The Inelegant Design of the Brain

WHEN I WAS IN middle school, in California in the 1970s, a popular joke involved asking someone, "Want to lose 6 pounds of ugly fat?" If the reply was positive it would be met with "Then chop off your head! Hahahaha!" Clearly, the brain did not hold a revered place in the collective imagination of my classmates. Like many, I was relieved when middle school drew to a close. Many years later, however, I have been similarly distressed by the opposite view. Particularly when reading books or magazines or watching educational television, I have been taken aback by a form of brain worship. Discussion of the brain is most often delivered in a breathless, awestruck voice. In these works the brain is "an amazingly efficient 3 pounds of tissue, more powerful than the largest supercomputer," or "the seat of the mind, the pinnacle of biological design." What I find problematic about these statements is not the deep appreciation

that mental function resides in the brain, which is indeed amazing. Rather, it is the assumption that since the mind is in the brain, and the mind is a great achievement, the design and function of the brain must then be elegant and efficient. In short, it is imagined by many that the brain is well engineered.

Nothing could be further from the truth. The brain is, to use one of my favorite words, a kludge (pronounced "klooj"), a design that is inefficient, inelegant, and unfathomable, but that nevertheless works. More evocatively, in the words of the military historian Jackson Granholm, a kludge is "an ill-assorted collection of poorly matching parts, forming a distressing whole."

What I hope to show here is that at every level of brain organization, from regions and circuits to cells and molecules, the brain is an inelegant and inefficient agglomeration of stuff, which nonetheless works surprisingly well. The brain is not the ultimate general-purpose supercomputer. It was not designed at once, by a genius, on a blank piece of paper. Rather, it is a very peculiar edifice that reflects millions of years of evolutionary history. In many cases, the brain has adopted solutions to particular problems in the distant past that have persisted over time and have been recycled for other uses or have severely constrained the possibilities for further change. In the words of the pioneering molecular biologist François Jacob, "Evolution is a tinkerer, not an engineer."

What's important about this point as applied to the brain is not merely that it challenges the notion of optimized design. Rather, appreciation of the quirky engineering of the brain can provide insights into some of the deepest and most particularly *human* aspects of experience, both in day-to-day behavior and in cases of injury and disease.

SO, WITH THESE issues in mind, let's have a look at the brain and see what we can discern about its design. What are the organizational principles that emerge? For this purpose, imagine that we have a freshly dissected adult human brain before us now (Figure 1.1). What you would see is a slightly oblong, grayish-pink object weighing about 3 pounds. Its outer surface, which is called the cortex, is covered with thick wrinkles that form deep grooves. The pattern of these grooves and wrinkles looks like it might be variable, like a fingerprint, but it is actually very similar in all human brains. Hanging off the back of the brain is a structure the size of a squashed baseball with small crosswise grooves. This is called the cerebellum, which means "little brain." Sticking out of the bottom of the brain, somewhat toward the back end is a thick stalk called the brainstem. We've lopped off the very bottom of the brainstem where it would otherwise taper to form the top of the spinal cord. Careful observation would reveal the nerves, called the cranial nerves, which carry information from the eyes, ears, nose, tongue, and face into the brainstem.

One obvious characteristic of the brain is its symmetry: the view from the top shows a long groove from front to back that divides the cortex (which means "rind"), the thick outer covering of the brain, into two equal halves. If we slice completely through the brain, using this front-to-back groove as a guide, and then turn the cut side of the right half toward us, we see the view shown in the bottom of Figure 1.1.

Looking at this image makes it clear that the brain is not just a homogeneous blob of stuff. There are variations in shape, color, and texture of the brain tissue across brain regions, but these do not tell us about the functions of these various regions. One of the most useful ways to investigate the function of these locations is to look at people who have sustained damage to various parts of the brain. Such investigations have been complemented by animal experiments in which small regions of the brain are precisely damaged through surgery or the administration of drugs, after which the animal's body functions and behavior are carefully observed.

The brainstem contains centers that control extremely basic regulation of

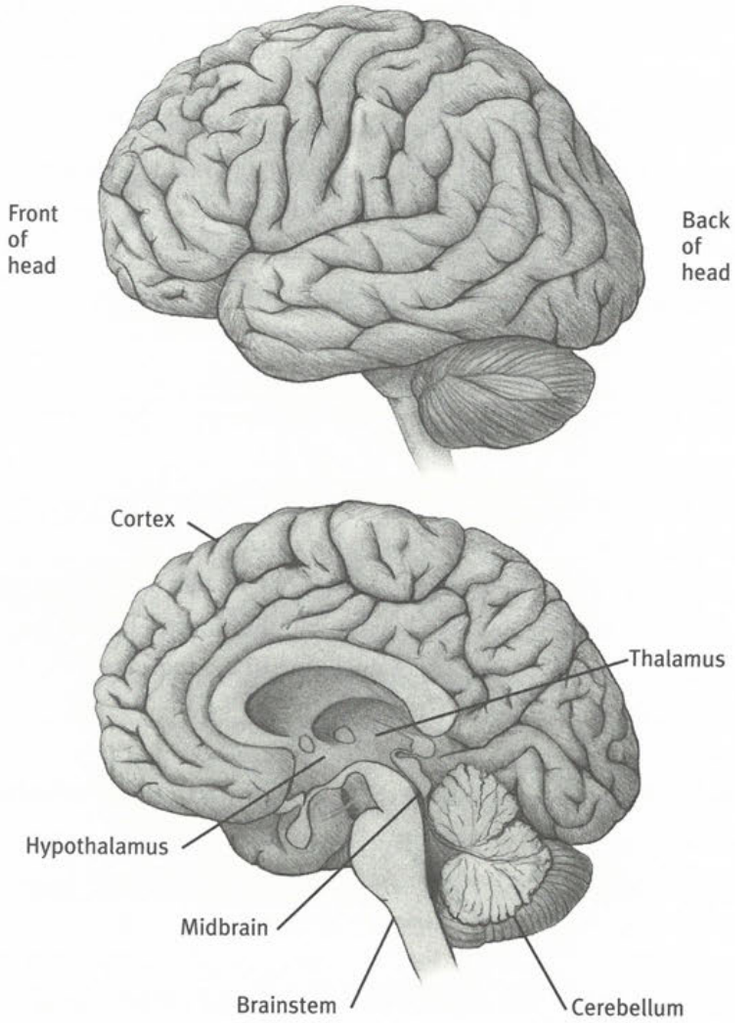


FIGURE 1.1. The human brain. The top shows the intact brain viewed from the left side. The bottom shows the brain sliced down the middle and then opened to allow the right side to face us. *Joan M. K. Tycko, illustrator.*

the body that are not under your conscious control, including vital functions such as regulation of heart rate, blood pressure, breathing rhythm, body temperature, and digestion. It also contains the control centers for some important reflexes, such as sneezing, coughing, and vomiting. The brainstem houses relays for sensations coming up the spinal cord from your skin and muscles as well as for command signals coming from your brain and destined for muscles in your body. It also contains locations involved in producing feelings of wakefulness versus sleepiness. Drugs that modify your state of wakefulness, such as sleeping pills or general anesthetics on the one hand and caffeine or amphetamines on the other, act on these brainstem regions. If you get a small area of damage in your brainstem (from an injury, tumor, or stroke), you could be rendered comatose, unable to be aroused by any sensation, but extensive damage in the brainstem is almost always fatal.

The cerebellum, which is richly interconnected with the brainstem, is involved with coordination of movements. In particular, it uses feedback from your senses about how your body is moving through space in order to issue fine corrections to the muscles to render your movements smooth, fluid, and well coordinated. This cerebellar fine-tuning operates not only in the most demanding forms of coordination such as hitting a baseball or playing the violin, but also in everyday activities. Damage to the cerebellum is subtle. It will not paralyze you, but rather will typically result in clumsiness in performing simple tasks that we take for granted, such as reaching smoothly to grasp a coffee cup or walking with a normal gait; this phenomenon is called ataxia.

The cerebellum is also important in distinguishing sensations that are "expected" from those that are not. In general, when you initiate a movement and you have sensations which result from that movement, you tend to pay less attention to those sensations. For example, when you walk down the street and your clothes rub against your body, these are sensations that you mostly ignore.

By contrast, if you were standing still and you started to feel similar rubbing sensations on your body, you would probably pay a lot of attention. You would probably whirl around to see who was groping you. In many situations, it is useful to ignore sensations produced by your own motion and pay close attention to other sensations that originate from the outside world. The cerebellum receives signals from those brain regions that create the commands that trigger body motion. The cerebellum uses these signals to predict the sensations that are likely to result from this motion. Then the cerebellum sends inhibitory signals to other brain regions to subtract the "expected" sensations from the "total" sensations and thereby change the way they feel to you.

This may all sound a bit abstract, so let's consider an example. It is well known that you can't tickle yourself. This is not just true in certain cultures; it is worldwide. What's different about having someone else tickle you, which can result in a very strong sensation, and self-tickling, which is ineffective? When researchers in Daniel Wolpert's group at University College, London, placed people's heads in a machine that can make images showing the location and strength of brain activity (called functional magnetic-resonance images, or fMRI) and then tickled them, they found strong activation in a brain region involved in touch sensation called the somatosensory cortex and no significant activation in the cerebellum. When people were then asked to tickle themselves on that same part of the body, it was seen that there was a spot of activation in the cerebellum and reduced activity in the somatosensory cortex. The interpretation of this experiment is that commands to activate the hand motions in self-tickling stimulated the cerebellum, which then formed a prediction of the expected sensation and sent signals encoding this prediction to inhibit the somatosensory cortex. The reduced activation of the somatosensory cortex was then below the threshold necessary to have the sensation feel like tickling. Interestingly, there are now reports that some humans who sustain damage to the

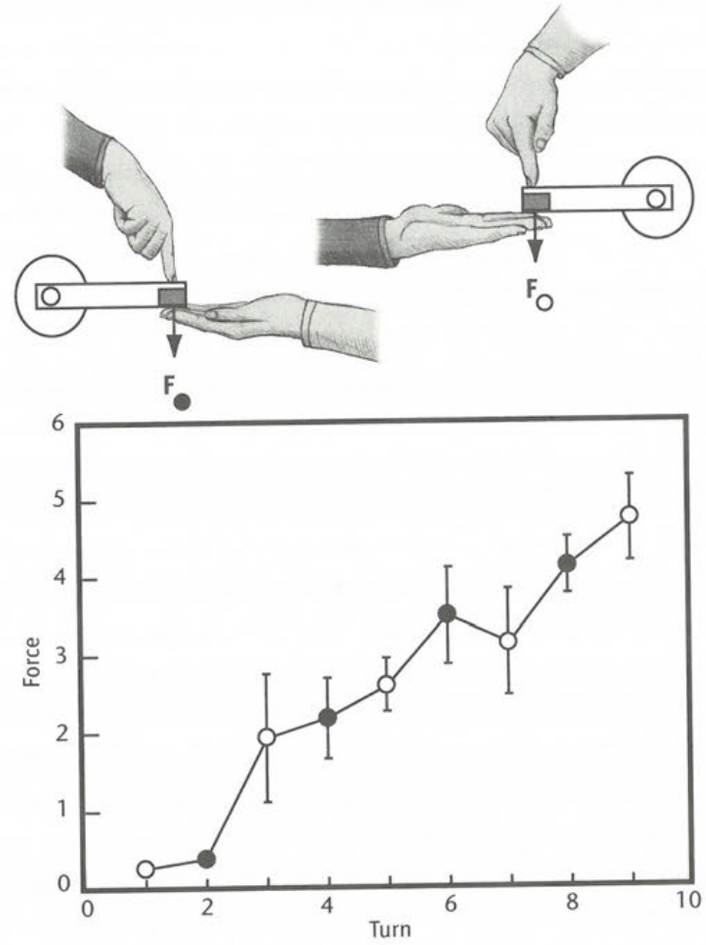


FIGURE 1.2. Force escalation in a tit-for-tat finger-tapping task. The white circles show the force of finger taps delivered by one subject, the black circles the force from the other subject. In 9 tit-for-tat exchanges, the force increased almost 20-fold. Adapted from S. S. Shergill, P. M. Bays, C. D. Frith, and D. M. Wolpert, Two eyes for an eye: the neuroscience of force escalation, *Science* 301:187 (2003); copyright 2003 AAAS. *Joan M. K. Tycko, illustrator.*

cerebellum cannot generate predictions of expected sensations and therefore can actually tickle themselves!

Daniel Wolpert and his colleagues at University College, London, have also devised a simple and elegant experiment to explain the cerebellum's involvement in the escalation of a shoving match (Figure 1.2). When a shoving match starts between two people the force of the shoving tends to escalate, often to the point of a full-blown brawl. Typically, we have thought of this solely in terms of social dynamics: neither participant wants to show weakness by backing down. That may explain why the conflict continues, but it does not necessarily shed light on why the force of each shove increases in a tit-for-tat exchange.

What Wolpert and his colleagues did was have two adult subjects face each other, each resting the left index finger, palm up, in a molded depression. A small metal bar on a hinge was then rested lightly on top of each subject's finger. The hinge was fitted with a sensor to measure the force delivered when the bar was pressed down. Both subjects were given the same instructions: exactly match the force of the tap on his finger that he receives with an equivalent tap when his turn comes. Neither subject knew the instructions given to the other.

Despite explicit instructions to the contrary, when the subjects took turns pressing on each other's fingers, the force applied always escalated dramatically, just as it does in schoolyard or bar-room confrontations. Each person swore that he matched the force of the other's tap. When asked to guess the instructions given to the other person, each said, "You told the other guy to press back twice as hard."

Why does this happen? Several clues point to the answer. First, it is not specific to social situations. When a person is asked to match the force of a finger tap which comes from a machine, he or she she will also respond with greater force. The second line of evidence comes from modifying the tit-for-tat experiment so that the tap is produced not by pressing on a bar but rather by moving a

joystick that controls the pressure by activating a motor. The important difference between these two situations is that when the force is generated by bar pressing, making a stronger tap requires generating more force with the fingertip. When the joystick is used, however, the motor does the work and there is only a weak correlation between the force generated by the tapping finger and the force produced on the upturned finger of the other subject. When the tit-for-tat experiment is then repeated with joysticks there is very little force escalation. The interpretation here is similar to that offered for self-tickling: The cerebellum receives a copy of the commands to produce the finger tap (using the bar) that are proportional to the force applied. It then creates a prediction of the expected sensation that is sent to the somatosensory cortex to inhibit feedback sensations from the fingertip during tapping. To overcome this inhibition, the subject presses harder to match the force perceived from the last tap he or she received, thus escalating the force applied.

So, in most situations, the cerebellar circuit that allows us to pay less attention to sensations that result from self-generated movement and more attention to the outside world is a useful mechanism. But as any 8-year-old coming home with a black eye and a tale of "But Mom, he hit me harder!" will tell you, there is a price to pay for this feature. This is a common brain design flaw. Most systems, like the cerebellar inhibition of sensations from self-generated movement, are always on. They cannot be switched off even when their action is counterproductive.

Moving up and forward from the cerebellum, the next region we encounter is called the midbrain. It contains primitive centers for vision and hearing. These locations are the main sensory centers for some animals, such as frogs or lizards. For example, the midbrain visual center is key for guiding the tongue-thrust frogs use to capture insects in flight. But in mammals, including humans, the midbrain visual centers are supplemented and to some degree sup-

planted by more elaborate visual regions higher up in the brain (in the cortex). Even though we make only limited use of a frog-like visual region in our brains (mostly in orienting our eyes to certain stimuli), this evolutionarily ancient structure has been retained in human brain design and this gives rise to the fascinating phenomenon called blindsight.

Patients who are effectively blind owing to damage to the higher visual parts of the brain will report that they have no visual sense whatsoever. When asked to reach for an object in their visual field, such as a penlight, they will say, "What can you possibly mean? I can't see a thing!" If however, they are told to just take a guess and try anyway, they can usually succeed at this task at a rate much higher than would be due to pure chance. In fact, some patients can grasp the penlight 99 percent of the time, yet will report each time that they have no idea where the target is and they are guessing randomly. The explanation seems to be that the ancient visual system in the midbrain is intact in these patients and guides their reaching, yet because this region is not interconnected with the higher areas of the brain, these people have no conscious awareness of the penlight's location. This underscores a general theme that is emerging here. The functions of the lower portions of the brain such as the brainstem and the midbrain are generally performed automatically, without our conscious awareness. As we continue our tour to those parts of the brain that are both literally and metaphorically higher, then we will begin to make the transition from subconscious to conscious brain function.

Furthermore, the midbrain visual system is a lovely example of brain kludge: it is an archaic system that has been retained in our brains for a highly delimited function, yet its action can be revealed in brain injury. As an analogy, imagine if your present-day audio electronics, let's say that sleek handheld MP3 player, still contained a functional, rudimentary 8-track tape player from the 1960s.

Not too many of those would get sold, even with a really urban-hip, edgy ad campaign.

Moving a bit upward and forward, we reach two structures called the thalamus and the hypothalamus (which just means "below the thalamus"). The thalamus is a large relay station for sending sensory signals on to higher brain areas and also relaying command signals from these areas out along pathways that ultimately activate muscles. The hypothalamus has many smaller parts, each of which has a separate function, but one general theme of this region is that it helps to maintain the status quo for a number of body functions, a process called homeostasis. For example, when you get too cold, your body begins to shiver reflexively in an attempt to generate heat through muscular activity. The shivering reflex originates within the hypothalamus.

Perhaps the most well-known homeostatic drives are those that control hunger and thirst. Although the urge to eat and drink can be modulated by many factors, including social circumstances, emotional state, and psychoactive drugs (consider the well-known phenomenon of "the munchies" from smoking marijuana and the appetite-suppressing action of amphetamines), the basic drives for hunger and thirst are triggered within the hypothalamus. When tiny holes are made surgically in one part of the hypothalamus of a rat (called the lateral nucleus; a "nucleus" in the brain is just a name for a group of brain cells), it will fail to eat and drink, even after many days. Conversely, destroying a different part of the hypothalamus (the ventromedial nucleus) results in massive overeating. Not surprisingly, a huge effort is under way to identify the chemical signals that trigger feelings of hunger and fullness, with the hope of making a safe and effective weight-loss drug. So far, this has proven to be much more difficult than anticipated because multiple, parallel signals for both beginning and ending feeding appear to play a role.

In addition to its involvement in homeostasis and biological rhythms, the hypothalamus is also a key controller of some basic social drives, such as sex and aggression. I will talk about these functions in detail later. A point that must be made here, though, is that the hypothalamus exerts some of its effects on these drives by secreting hormones, powerful messenger molecules that are carried in the bloodstream throughout the body to cause many varied responses. The hypothalamus secretes two types of hormones. One type has direct actions on the body (such as the hormone called vasopressin, which acts on the kidney to limit the formation of urine and thereby increase blood pressure), and the second type, the so-called master hormones, directs other glands to secrete their own hormones. A good example of the latter is growth hormone, secreted by the pituitary gland in growing children and adolescents but stimulated by a master hormone released by the hypothalamus. After much careful scientific thought, this master hormone was given the compelling name "growth hormone releasing hormone" (endocrinologists, like many scientists, are not known for their literary flair).

Up to this point, we have been looking at the brain sliced exactly down the middle. Many areas inside the brain are revealed with this view, but others are buried deep within the tissue and are not visible either from the outside surface or from the cut surface at the midline. Particularly important are two deeply buried structures called the amygdala ("almond") and the hippocampus ("seahorse") that constitute part of a larger circuit in the center of the brain called the limbic system (which also contains portions of the thalamus, cortex, and other regions). The limbic system is important for emotion and certain kinds of memory. It is also the first place in our bottom-to-the-top tour where automatic and reflexive functions begin to blend with conscious awareness.

The amygdala is a brain center for emotional processing that plays a particular role in fear and aggression. It links sensory information that has already been

highly processed by the cortex (that guy in the ski mask jumping out of that dark alley at me can't be up to any good) to automatic fight-or-flight responses mediated by the hypothalamus and brainstem structures (sweating, increased heart rate, dry mouth). Humans rarely sustain damage to the amygdala alone, but those who do often have disorders of mood and appear to be unable to recognize fearful expressions in others. Electrical stimulation of the amygdala (as sometimes occurs during neurosurgery) can evoke feelings of fear, and the amygdala also appears to be involved in storing memories of fearful events.

The hippocampus (which, when dissected out of the brain, actually looks more like a ram's horn than the seahorse for which it is named) is a memory center. Like the amygdala, it receives highly processed sensory information from the cortex lying above it. Rather than mediating fear, however, the hippocampus appears to have a special role in laying down the memory traces for facts and events, which are stored in the hippocampus for a year or so but are then moved to other structures. The most compelling evidence for this model comes from a small number of people who have sustained damage to their hippocampus and some surrounding tissue on both sides of the brain. The most famous of these cases is called H.M. (initials used to protect privacy), a man who in 1953 underwent surgical removal of the hippocampus and some surrounding tissue on both sides of his brain in order to control massive seizures that had not responded to other treatments. The surgery was successful in controlling his epilepsy and did not impair his motor functions, language, or general cognitive abilities, but there were two disastrous side effects. First, H.M. lost his memory of everything that occurred 2–4 years before the surgery. He had extensive, detailed, and accurate recall of earlier events, but his memory of his life in the years just before the surgery is lost forever. Even more devastating is that since the surgery H.M. has been unable to store new memories for facts and events. If you were to meet him on Monday, he would not remember you

on Tuesday. He can read the same book every day and it will be new to him. Although he has short-term memory that can span tens of minutes, his ability to store new permanent memories for facts and events is gone.

The seminal insights about memory and the hippocampus that came from H.M.'s case have since been reinforced many times, both by other patients who, for a variety of reasons, have sustained similar damage, and by animal studies in which the hippocampus has been surgically destroyed or had its function disrupted by drugs. A consistent and simple conclusion comes from this work: without a hippocampus, the ability to store new memories for facts and events is severely impaired.

Finally, moving to the outer surface of the brain, we reach the cortex. The cortex of the human brain is massive. The functions of some areas in the cortex are well understood, but others are terra incognita. A portion of the cortex analyzes the information coming from your senses. The very back of your cortex is where visual information first arrives, and another strip of tissue just behind the main sideways groove in your brain (called the central sulcus) is where touch and muscle sensation first arrives. Similar maps can be drawn for other senses. If we stimulate these areas with an electrode we can mimic activation of the sensory system involved: stimulating the primary visual cortex will cause a flash of light, or something similar, to be seen. Likewise, there is a strip of cortex just in front of the central sulcus that sends out command signals that ultimately cause contraction of muscles and consequent body movement. Electrically stimulating this motor cortex results in muscular contraction. This is a standard technique for making a functional map of the brain when surgery must be performed in this area. What's most interesting about the cortex are those regions for which the functions are not obviously either sensory or motor. Brain researchers have sometimes called these regions association cortex. Association areas are most plentiful in the front of the brain (the frontal cortex), a region that is highly developed in humans.

I have offered a number of examples where people (and experimental animals) sustain damage to various brain regions and suffer various losses of function ranging from amnesia to overeating. Yet, to this point, though many of these brain insults have had devastating effects, none of them has changed the personality, the essential core identity of the sufferer. H.M., for example, has the same unique personality that he had before his epilepsy surgery. A far different picture emerges when we consider damage to the frontal cortex.

Here, the most well known example is Phineas Gage, a foreman on a Vermont railway gang in 1848. Railway construction, then and now, uses blasting to remove obstacles and level the roadbed. Phineas, aged 25, had the unenviable task of jamming the explosive charge into place using a long metal rod known as a tamping iron. You can imagine what happened. As he stood over a borehole, tamping the charge, there was a spark that ignited a horrible explosion. The explosion drove the tamping iron through Phineas's left cheek and eye at a steep upward angle, piercing his skull through the eye socket tearing a huge hole in his left frontal cortex, and exiting his skull through the top. Figure 1.3 shows a drawing based on a scan of his skull made long after his death, with the tamping iron in place. Amazingly, after a few weeks in bed, Phineas made a full recovery. The infection of his wound abated. He could walk, talk, and do arithmetic in his head. His long-term memory was fine. What *had* changed was his personality and his judgment. By all reports, before the accident he was kind, level-headed, friendly, and charismatic. After his recovery he became arrogant, opinionated, impulsive, rude, and selfish. Not to put too fine a point on it, damage to his frontal cortex changed him from a nice guy into a jerk. His former coworkers couldn't stand him. "He's just not Gage anymore," one

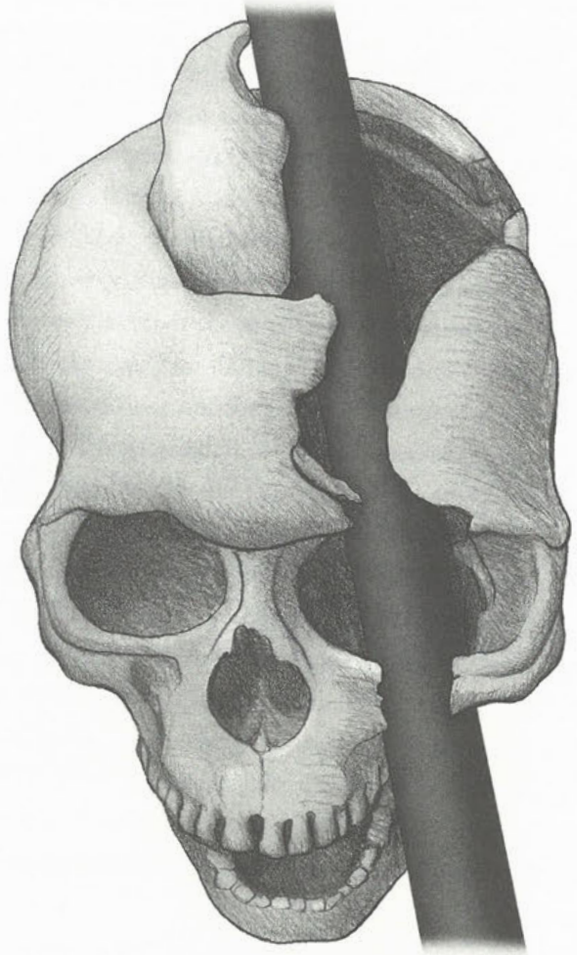


FIGURE 1.3. The skull of Phineas Gage, with the famous tamping iron, reconstructed by computer from scans made long after his death. Derived with permission from P. Ratiu and I.-F. Talos, Images in clinical medicine: the tale of Phineas Gage, digitally remastered, *The New England Journal of Medicine* 351:e21 (2004). *Joan M. K. Tycko, illustrator.*

friend reportedly said. Tragically, he ended up in a carnival freak show, reinserting the tamping iron through the healed but still present hole in his head to the morbid fascination of onlookers. He died 12 years after the tamping iron accident.

As shown by the case of Phineas Gage, and documented many times since, the frontal cortex is the substrate of our individuality, determining our social interactions, outlook, and perhaps even our moral sense. Not just our cognitive capacities but our character—our personhood, so to speak—resides in this most recently evolved region of our brains.

HAVING COMPLETED OUR whirlwind tour from the bottom to the top of the brain (leaving out a few areas), what can we conclude about the overall principles of brain design? Guiding Principle One: The highest functions of our brain, involving conscious awareness and decision making, are located at the very top and front, in the cortex, and the lowest functions, supporting basic subconscious control of our body functions such as breathing rhythm and body temperature, are located in the very bottom and rear, in the brainstem. In between are centers that are engaged in higher subconscious functions such as rudimentary sensation (midbrain), homeostasis and biological rhythms (hypothalamus), and motor coordination and sensory modulation (cerebellum). The limbic system, including the amygdala and hippocampus, is the crossroads where the conscious and unconscious parts of the brain meet and initiate the storage of certain types of memories.

Guiding Principle Two: The brain is built like an ice cream cone (and you are the top scoop): Through evolutionary time, as higher functions were added, a new scoop was placed on top, but the lower scoops were left largely unchanged. In this way, our human brainstem, cerebellum, and midbrain are not very different in overall plan from that of a frog. It's just that a frog has only rudimen-

tary higher areas in addition (barely more than one scoop). All those structures plus the hypothalamus, thalamus, and limbic system are not that different between humans and rats (two scoops), which have a small and simple cortex, while we humans have all that plus a hugely elaborated cortex (three scoops). When new, higher functions were added, this did not result in a redesign of the whole brain from the ground up; a new scoop was just added on top. Hence, in true kludge fashion, our brains contain regions, like the midbrain visual center, that are functional remnants of our evolutionary past.

You probably have seen those quaint charts from the nineteenth century (Figure 1.4), in which the surface of the brain is divided into neat regions, each labeled with a cognitive function (such as calculation) or a personality trait (say combativeness). The phrenologists who used these charts believed not only that those functions could be mapped to those particular brain regions but also that bumps on the skull resulted from the overgrowth of a particular brain region. Indeed, there was a cottage industry in the nineteenth and early twentieth centuries of professional head-bump feelers, who, armed with charts, plaster models, and even a mechanical bump-measuring helmet, would analyze the skull-and-mind of anyone willing to pay.

The phrenologists were wrong on two counts. First, bumps on the skull don't indicate anything about the underlying brain tissue. Second, their diagrams equating particular regions with cognitive functions and personality traits were pure fantasy. But on a more general issue, the phrenologists were right: the brain is not an undifferentiated mass of tissue where each region contributes equally to all functions. Rather, particular brain functions often are localized to distinct brain regions.

This brings us to Guiding Principle Three: Localization of function in the brain is straightforward for basic subconscious reflexes such as vomiting and is fairly straightforward for the initial stages of sensation (we know where signals

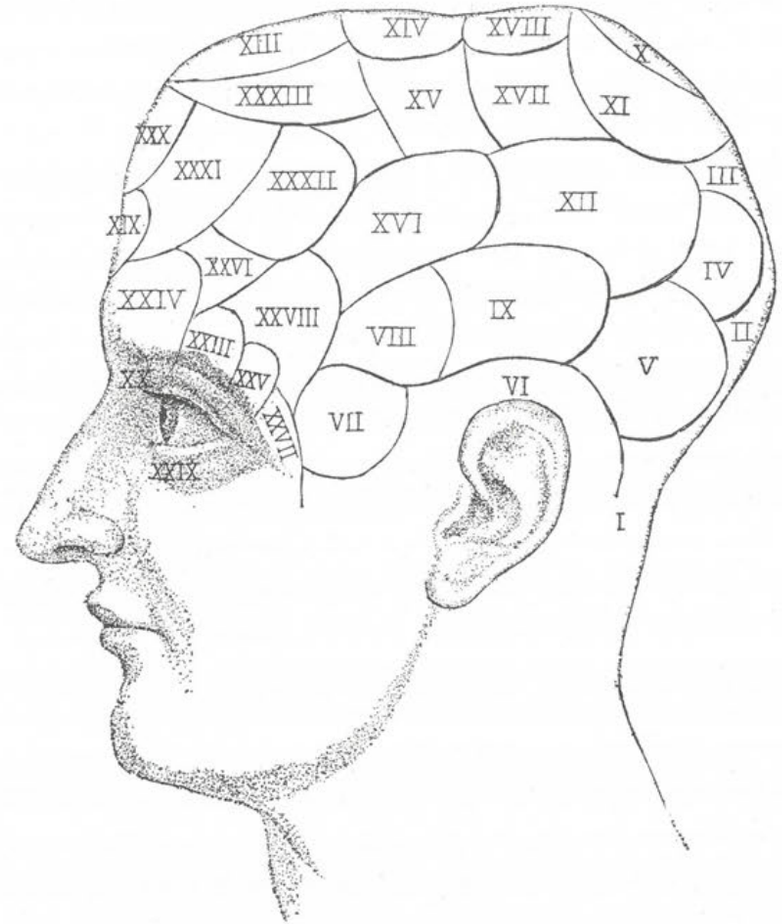


FIGURE 1.4. A phrenologist's chart from the nineteenth century, equating head bumps with particular mental traits. In this case, for example, XIV = veneration, XVII = hope, XIII = benevolence, XXI = imitation, XIX = ideality, VIII = acquisitiveness, XVIII = marvelousness, and XX = wit. *From W. Mattieu Williams, A Vindication of Phrenology (Chatto & Windus, London, 1894).*

first arrive in the cortex for vision, hearing, smell, and so forth) But localization of function is much more difficult for more complex phenomena such as memory of facts and events and is really hard for the highest functions such as decision making. In some cases it becomes complicated because the location of a function in the brain is not fixed over time: memories for facts and events seem to be stored in the hippocampus and some immediately adjacent regions for 1–2 years but are then exported to other locations in the cortex. Decision making generally is such a broad function, and generally requires such a convergence of information, that it may be broken into smaller tasks and distributed to a number of places in the cortex. We may have to define functions more precisely in order to achieve a greater understanding of functional localization.

So, given these Guiding Principles, what is it about this organ that makes us so clever? What is it about our brain that enables language and the ability to understand the motivations of others (the so-called theory of mind) and other capacities that humans have developed far beyond the abilities of other animals? We don't have the biggest brains (an elephant's is bigger) and we don't even have the biggest brain-to-body-weight ratio (small birds beat us on that measure). We don't have the most wrinkled brain surface (whales and dolphins' are more wrinkled). In fact, we don't even have the largest brains among our hominid kin: estimates derived from skull volumes indicate that Neanderthals had brains that were, on average, somewhat larger than ours today. And, although I haven't talked about it yet, we can assume that, overall, the shape and chemical composition of the cells that make up our brains are not fundamentally different from those of a rat (more on this to come). What we do have is the largest association cortex, that which is not strictly sensory or motor, most of it packed into the front half of our brain. Somehow, this is the elaboration that appears to have given humans their cognitive advantages.

Can we take this one step farther? Humans have varying cognitive abilities.

Can human cognitive capacity be predicted by the overall size of the brain or by the size of particular brain regions? Diseases (both inherited and acquired) and trauma, both of which produce gross anatomical disruptions to the brain, can clearly impair cognition. But what about normal variation, excluding obvious mishaps such as trauma or disease? Recent studies relating normal human variation in cognitive ability to brain size or shape have used brain-scanning techniques that provide more accurate measures than older studies that relied upon skull measurements. In general, these newer studies have found statistically significant correlations between brain size (adjusted for body weight) and cognitive ability. But this correlation, while real, accounts for only about 40 percent of the variation in cognitive ability of normal humans. Thus one can find people at the small end of the range of normal brain sizes (say, 1,000 cubic centimeters) who will score highly on a so-called test of general intelligence. Conversely, one can find individuals with unusually large brains (1,800 cubic centimeters) who score well below average.

The large variation in the relationship between human brain size or shape and cognitive capacity has not stopped the continual trickle of publications in which the preserved brains of famous historical figures have been analyzed anatomically. Lenin's brain was studied in Germany in the late 1920s and, while it was of average weight, in some regions a particular subset of cells in the brain (called layer 3 cortical pyramidal cells) were purported to be unusually large compared to other postmortem samples. Einstein's brain actually was smaller than average (but well within the normal range). Recently, there has been a claim that a region of his brain called the inferior parietal cortex was slightly (15 percent) enlarged relative to a sample of men's brains of a similar age. That caused some interest because this region has been associated with spatial and mathematical cognition, areas in which Einstein clearly excelled. But one must be cautious in interpreting this sort of finding. First, it's very hard to make a

claim based on a single sample (Einstein). A more convincing study would need a whole group of mathematical/spatial geniuses compared with controls carefully matched for age, lifestyle, and other factors. Second, and more important, there's a problem of causality at work. If, indeed, a part of Einstein's brain involved in mathematical/spatial thinking was significantly larger than appropriate control brains, does that mean that this variation endowed him with mathematical ability that he was then able to exploit? Or did his lifelong engagement in mathematical and spatial pursuits cause this part of his brain to grow slightly?

Failure up to now to strongly associate gross anatomical features of the brain with normal variation in human cognition should not be taken to mean that variation in human cognition has no measurable physical correlate in brain structure. It's very likely that such a relationship does exist. But this correlation will be only weakly reflected in crude measures such as brain size. Most of human cognitive variation is more likely to be manifest as changes in the microscopic anatomy, the connectivity of brain cells, and the patterns of brain electrical activity.

WE'VE UNCOVERED THREE Guiding Principles of Brain Design and these highlight a few of the ways in which the human brain is poorly organized. The brain has primitive systems that developed in our distant evolutionary past (before mammals) and that have been supplemented by newer, more powerful structures. These primitive structures persist in the lower parts of our brain, giving rise to interesting phenomena such as blindsight. Also, the brain has regions that perform functions that are often useful, such as cerebellar inhibition of the sensations from self-originated movements, but that cannot be turned off in the appropriate circumstances, a fact that contributes to problems such as force escalation in tit-for-tat conflicts.

To put this in perspective, imagine that you are an engineer in charge of building the latest and most efficient car. Only after you agree to take the job do you learn that there are two weird stipulations. First, you are given a 1925 Model T Ford and told that your new car must take the form of adding parts to the existing structure while taking almost nothing of the original design away. Second, most of the new complex control systems you will build, such as the device that rapidly pumps the antilock brakes, must remain on all of the time (not just when a skid is detected). These are some of the types of constraints that have influenced the design of the human brain as it has evolved. Together with the engineering flaws of the component parts (the cells of the brain, which I will consider in Chapter 2) and the assembly process (brain development, covered in Chapter 3), these aspects of suboptimal design are central to brain function. By the end of this book I hope to have convinced you that almost every aspect of transcendent human experience, including love, memory, dreams, and even our predisposition for religious thought, ultimately derives from the inefficient and bizarre brain engineered by evolutionary history.

Chapter Two

Building a Brain with Yesterday's Parts

IT IS A CLICHÉ to be awed by the microscopic complexity of the human brain. Any scientist who talks about this topic inevitably hears the kindly, avuncular ghost of Carl Sagan whispering: "Bill-yuns and bill-yuns of tiny brain cells!" Well, it is rather impressive. There are a hell of a lot of cells in there. The two main cell types in the brain are: neurons, responsible for rapid electrical signaling (the brain's main business), and glial cells, important for housekeeping functions that create an optimal environment for neurons (and that directly participate in some forms of electrical signaling as well). The famous numbers: approximately 100 billion (100,000,000,000) neurons in the adult human brain and approximately one trillion (1,000,000,000,000) glial cells. To put this in perspective, if you wanted to give your neurons away to all humanity, everyone on earth would receive about 16 of them.

Neurons are not a recent development in evolution. They are soft and therefore not well preserved in fossils, so we don't know exactly when the first neurons appeared. But we do know that modern jellyfish, worms, and snails all have neurons. Some other modern animals, such as sea sponges, don't. Therefore, our best guess is that neurons appeared at about the time when jellyfish and their relatives, a group of animals called *Cnidaria,* first appeared in the fossil record, in the Pre-Cambrian era, about 600 million years ago. Incredibly, with few exceptions, the neurons and glial cells in a worm are not substantially different from those in our own brains. In this chapter, I hope to show you that our brain cells have an ancient design that makes them unreliable and slow, and limits signaling capacity.

Neurons come in a variety of shapes and sizes (see Figure 2.1), but have certain structures in common. Like all cells, neurons are bounded externally by a sort of skin, the outer membrane (also called the plasma membrane). All neurons have a cell body, which contains the cell nucleus, the storehouse of genetic instructions encoded in DNA. The cell body can be round, triangular, or spindle shaped and can range from 4 to 100 microns across (20 microns is typical). Perhaps a more useful way to think about this is that five average-sized neuronal cell bodies could be placed side by side in the width of a typical human hair. Thus the outer membranes of neurons and glial cells are incredibly tightly packed with very little space in between.

Sprouting from the cell body are dendrites (from the Greek word for "tree"), large, tapering branches of the neuron that receive chemical signals from neighboring neurons. I'll discuss how this happens soon. Dendrites can be short or long, spindly or bushy or, rarely, even completely absent. High magnification shows that some are smooth while others are covered with tiny nubbins called dendritic spines. Typical neurons have several branching dendrites, but they also have a single long thin protrusion growing from the cell body. This is the

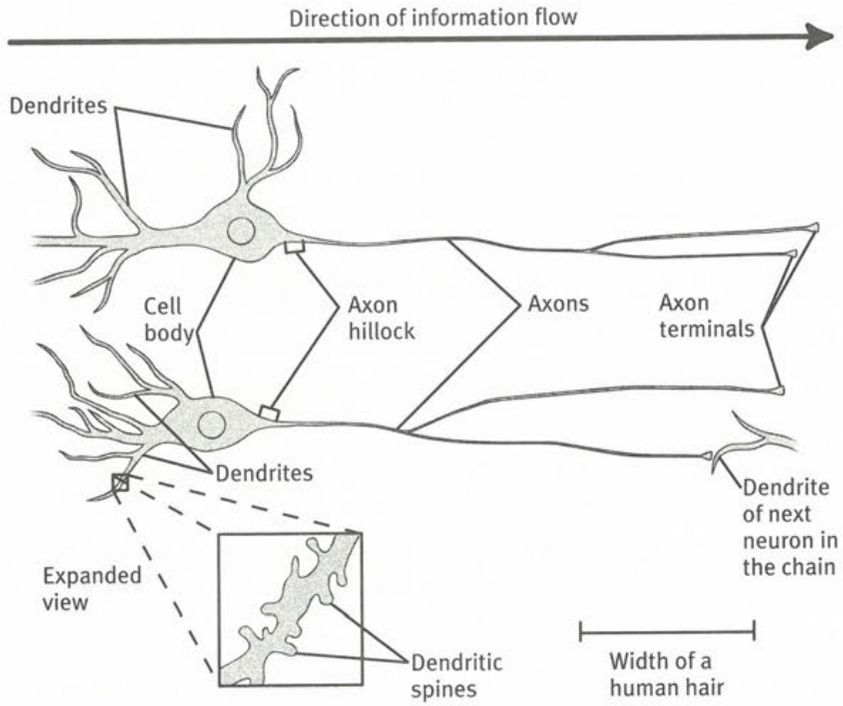


FIGURE 2.1. Two different neurons with their parts labeled. *Joan M. K. Tycko, illustrator.*

axon and is the information-sending side of the neuron. The axon, usually thinner than the dendrites, does not taper as it extends from the cell body. A single axon grows from the cell body, but it often subsequently branches, sometimes going to very different destinations. Axons can be remarkably long: some run all the way from the base of the spine to the toes (which makes the longest axons around 3 feet for average humans, and up to 12 feet long for a giraffe).

At specialized junctions called synapses, information passes from the axon of one neuron to the dendrite (or sometimes the cell body) of the next (Figure 2.2).

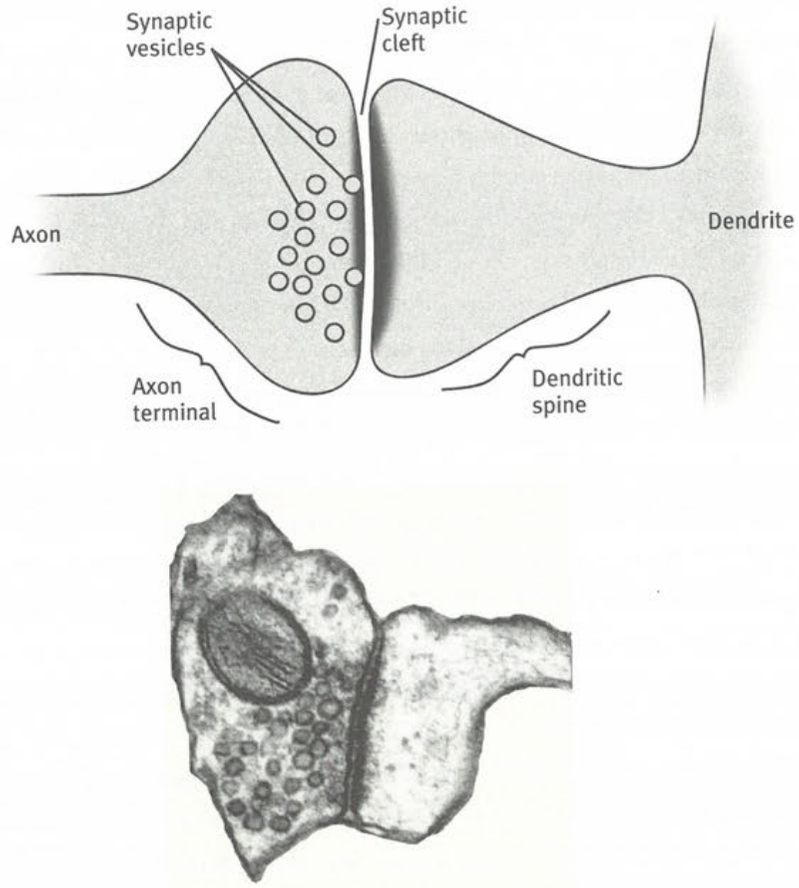


FIGURE 2.2. Parts of the synapse in a drawing (top) and in an actual electron microscope photo (bottom). *Joan M. K. Tycko illustrated the top panel. The bottom panel was kindly provided by Professor Kristen Harris of the Medical College of Georgia. Her website, synapses.mcg.edu, provides an excellent overview of the fine structure of synapses.*

At synapses, the ends of axons (called axon terminals) nearly, but not actually, touch the next neuron. Axon terminals contain many synaptic vesicles, tiny balls with a skin made of membrane. The most common type of synaptic vesicle in the brain is loaded with about 2,000 molecules of a specialized compound called a neurotransmitter. Between the axon terminal of one neuron and the dendrite of the next is a tiny saltwater-filled gap called the synaptic cleft. By tiny, I mean extremely tiny: about 5,000 synaptic clefts would fit in the width of a single human hair. The synaptic cleft is the location where synaptic vesicles release neurotransmitters to signal the next neuron in the chain.

Synapses are crucial to our story. They will come up repeatedly as I discuss everything from memory to emotion to sleep. We should therefore spend some time on them now. First, the number of synapses in the brain is staggering. On average, each neuron receives 5,000 synapses, locations where the axon terminals of other neurons make contact (the range is from 0 to 200,000 synapses). Most synapses contact the dendrites, some the cell body, and a few the axon. Multiplying 5,000 synapses per neuron by 100 billion neurons per brain, gives you an estimate of the astonishing number of synapses in the brain: 500 trillion, 500,000,000,000,000.

Synapses are the key switching points between the two forms of rapid signaling in the brain: chemical and electrical impulses. Electrical signaling uses a rapid blip, called a spike, as its fundamental unit of information. Spikes are brief electrical signals that originate at the axon hillock, the place where the cell body and the axon join. When spikes, having traveled down the axon, arrive at the axon terminals they trigger a series of chemical reactions that cause a dramatic structural change (see Figure 2.3). Synaptic vesicles fuse with the outer membrane of the axon terminal, dumping their contents, special neurotransmitter molecules, into the synaptic cleft. These neurotransmitter molecules then move across the synaptic cleft, where they contact specialized proteins

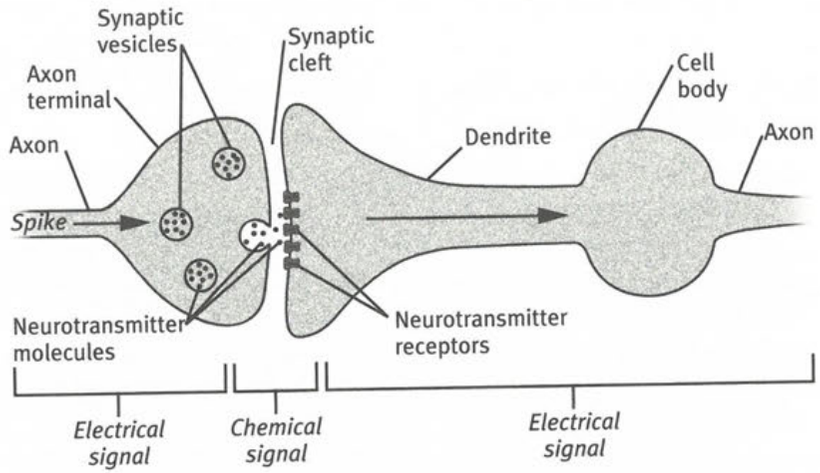


FIGURE 2.3. Synapses, the key sites in the brain for converting electrical signals to chemical signals and then back into electrical signals. Reading from left to right tells the story of synaptic signaling. *Joan M. K. Tycko, illustrator.*

called neurotransmitter receptors, embedded in the membrane of a neighboring neuron's dendrite. Receptors convert the neurotransmitter's chemical signal back into an electrical signal. Electrical signals from activated receptors all over the dendrite are funneled toward the cell body. If enough electrical signals arrive together, a new spike is triggered and the signal is passed farther along the chain of neurons.

That's the *Reader's Digest* version. Now, let's flesh that out with some real biology. At about 3 pounds, the brain constitutes about 2 percent of total body weight, and yet it uses about 20 percent of the body's energy. Clearly, the brain is an inefficient energy hog (the Hummer H2 of the body, if you will), but why is this so? The brain is naturally bathed in a special saltwater solution called cerebrospinal fluid that has a high concentration of sodium and a much lower

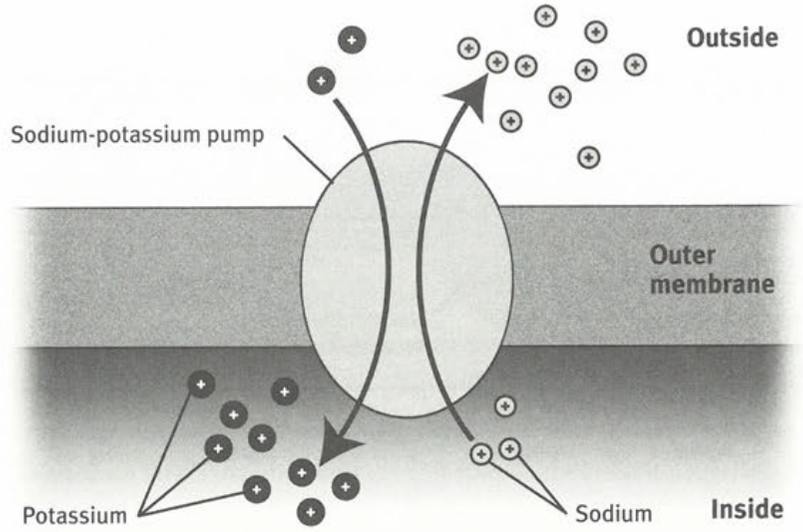


FIGURE 2.4. The sodium-potassium pump. Located in the outer membrane of neurons, it pumps sodium ions out and potassium in, thereby establishing the electrical gradient used by neurons to send information. *Joan M. K. Tycko, illustrator.*

concentration of potassium. These sodium and potassium atoms are in their charged state, called ions, in which they each have one unit of positive charge (+1). The brain's main energy expense involves continuously running a molecular machine that pumps sodium ions out of the cell and potassium ions in (see Figure 2.4). As a result of this pump's action, the concentration of sodium ions outside a neuron is about 10-fold higher than it is inside. For potassium, the concentration gradient runs in the other direction: the concentration of potassium ions is about 40-fold greater inside than outside. So neurons have saltwater solutions on both sides of their outer membranes (the skin of the cell), but very different saltwater solutions: the outside solution is high in sodium

and low in potassium; the inside solution is the opposite, low in sodium and high in potassium. That is the basis of electrical function in the brain. The differences in concentrations of sodium and potassium create potential energy, similar to that created by winding the spring on a child's toy, that can then be released in the appropriate circumstances to generate neural signals. Neurons rest with an electrical potential across their outer membranes: there is more negative charge inside the cell than outside.

Let's conduct an imaginary experiment that will help us understand neuronal electrical signaling. In our imagined lab, some neurons have been extracted from a rat's brain, placed in petri dishes, and grown in special solutions designed to mimic cerebrospinal fluid. This process is called neuronal cell culture and is a standard technique in brain research laboratories. In this experiment, illustrated in Figure 2.5, we insert recording electrodes into a neuron to measure the electrical signals across the outer membrane. Recording electrodes are hollow glass needles with very fine points, filled with a special saltwater solution that mimics the neuron's internal milieu (high potassium, low sodium). One electrode is in the dendrite, where a particular synapse is received, another is at the axon hillock, the place where the axon just starts to grow from the cell body, and a third electrode is way down in the axon terminal. Yet another electrode is used, not for recording, but rather for electrical stimulation of an axon terminal of another neuron that is contacting the dendrite of the first.

Before anything happens, we record the previously mentioned negative resting potential across the outer membrane of the information-receiving neuron. Measured in thousandths of a volt, or millivolts, our typical neuron's resting potential across its outer membrane is −70 millivolts, or about 1/20th the voltage of a single AA battery. Next, we electrically stimulate the adjacent axon terminal, causing it to release neurotransmitter molecules into the synaptic cleft. In our imaginary experiment, this neurotransmitter is the molecule glutamate.

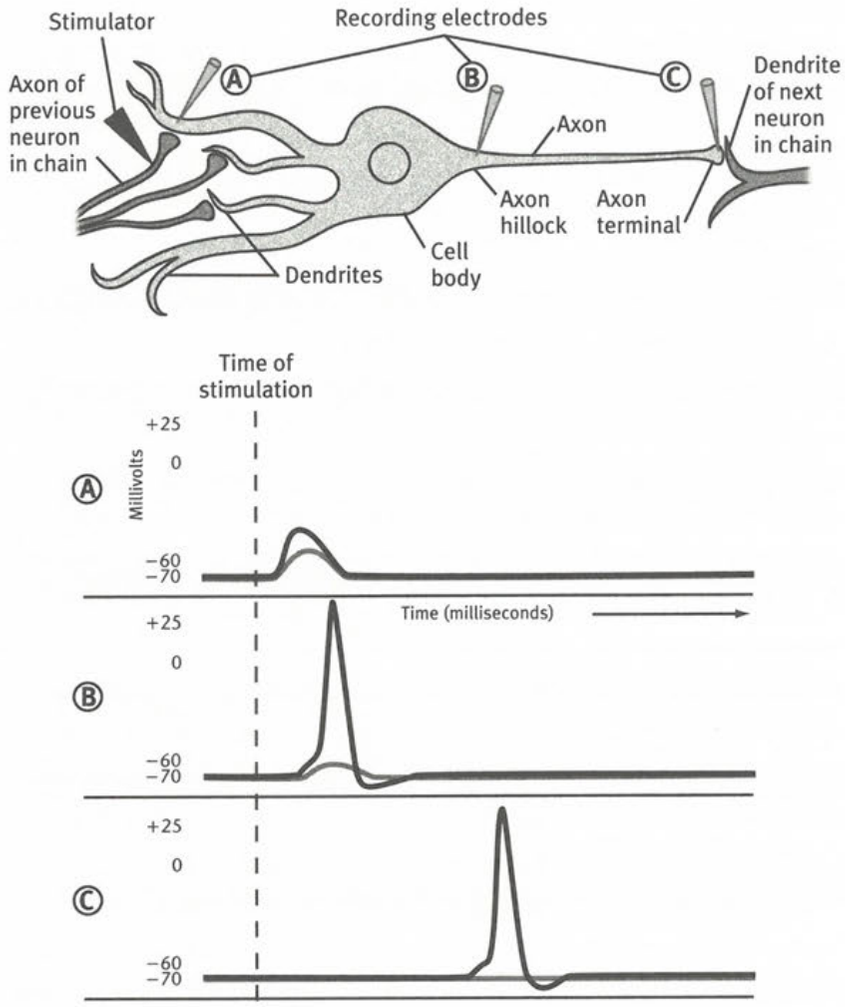


FIGURE 2.5. An imaginary experiment to investigate electrical signaling in neurons. Weak stimulation (of a few terminals) gives rise to the release of glutamate molecules, which diffuse across the synaptic cleft and bind glutamate receptors to evoke the responses indicated with gray lines in the chart at the bottom of the illustration. A small excitatory postsynaptic potential (EPSP) in the dendrites is even smaller in the axon hillock and fails to trigger a spike. Strong stimulation of terminals (responses indicated with black lines) causes a large EPSP in the dendrite. This EPSP is smaller in the axon hillock but is still big enough to cause a spike to be initiated here, and this spike then travels down the axon, where, after a delay, it is also recorded in the axon terminals. *Joan M. K. Tycko, illustrator.*

I have chosen glutamate as our example because it is by far the most common neurotransmitter molecule in the brain. When glutamate molecules are released at synapses, they diffuse across the narrow synaptic cleft separating two neurons. Glutamate molecules are not squirted across the synapse with force; they merely diffuse, like a single drop of red wine slowly mixing into a full glass of water. Because the synaptic cleft is so small, in only about 50 one-millionths of a second (5 microseconds) glutamate molecules released from the presynaptic axon terminal of one neuron cross to the other side, the postsynaptic membrane of the dendrite. Most of the glutamate molecules simply diffuse away and have no effect, but some will bind specialized glutamate receptor proteins that are embedded in the postsynaptic membrane. There are many different neurotransmitters in the brain, and though glutamate is the most common one, many others are important and will arise as I consider particular brain functions.

Glutamate receptor proteins are highly complex molecular machines. They are built of four similar parts that join together to form a doughnut-shaped structure around a central pore (Figure 2.6). In the resting state, this pore is shut tight, but when glutamate binds this receptor a gate that normally closes

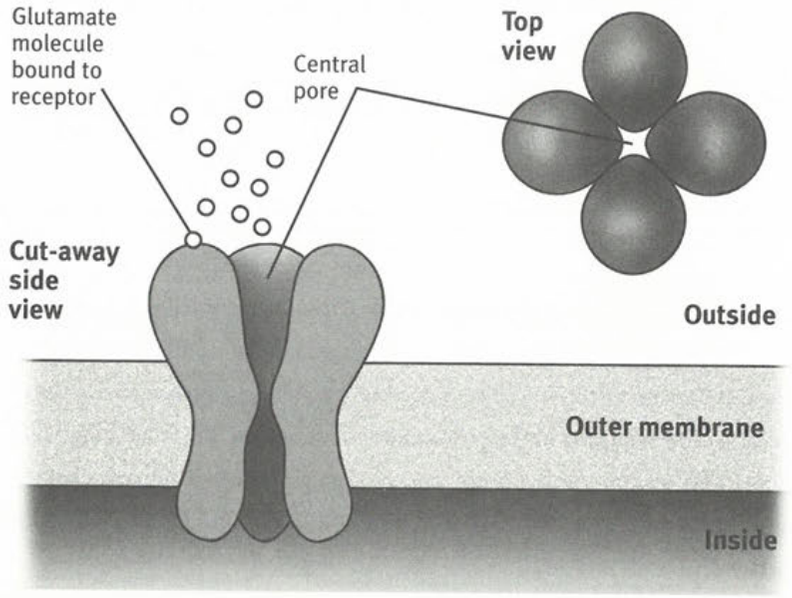


FIGURE 2.6. Schematic drawing of a glutamate receptor in the postsynaptic membrane. Glutamate binding to its receptor opens the central pore, the ion channel. *Joan M. K. Tycko, illustrator.*

off this central pore opens, thus allowing certain ions to flow in or out of the cell. The receptor's central pore is small and its particular chemical properties ensure that only particular ions can get through. Hence, the central pore has been given a special name, ion channel. In the case of the glutamate receptor, the ion channel allows passage of both sodium ions and potassium ions. When the pore opens, sodium ions from the outside (where sodium concentration is high) rush to the inside (where sodium concentration is low), and potassium ions flow in the opposite direction from inside (where concentration is high) to outside (where it is low). In this process, more sodium ions rush in than potassium ions flow out, so there is a net flow of positive charge into the cell, raising the voltage difference across the dendrite's outer membrane (the membrane potential) from its resting state of −70 millivolts to some more positive level, let's say −65 millivolts. As the glutamate molecules diffuse away from their receptors and the receptor-gated ion channel (central pore) closes again, the membrane potential returns to the resting state. This whole event, about 10 milliseconds in duration from start to finish, has been given a rather long and ponderous name, the excitatory postsynaptic potential, abbreviated EPSP.

In most neurons, a single EPSP produces a response like the one we have seen, a brief change in voltage, then nothing. This is a fairly typical mechanism that neurons have for ignoring very low levels of activity that are merely ongoing noise in the brain. Something very different happens if we activate a group of axon terminals to release glutamate all at the same time. We produce a larger EPSP at both the dendrite and the axon hillock, but when the strength of the signal at the axon hillock reaches a certain threshold level (say about −60 millivolts), an amazing thing happens. Rather than falling back down to rest, the membrane potential at the axon hillock explosively deflects upward and then rapidly returns. This explosive response is the spike, the fundamental unit of information in the brain.

Why is there a spike and why does it start at the axon hillock? The answer is in the structure of the outer membrane at this location. The axon hillock, but not the dendrite or cell body, has a high density of a different ion channel. These ion channels are not opened by binding glutamate, but rather have a built-in sensor of the local membrane voltage that allows them to be shut at rest (−70 millivolts) but open when the membrane voltage becomes more positive (to about −60 millivolts and beyond). When EPSPs from several different synapses add up at the axon hillock and move the membrane potential to −60 millivolts, then these voltage-sensitive ion channels begin to open. They are built

to allow only sodium ions through their central pore, and as this sodium rushes in, it moves the membrane to an even more positive potential. This, in turn, causes more voltage-sensitive sodium channel opening in a rapid positive feedback loop that underlies the explosive upstroke of the spike.

The spike typically peaks at about +50 millivolts and rapidly falls back to rest. There are two factors that contribute to this rapid peak-and-return behavior. First, voltage-sensitive sodium ion channels open rapidly but stay open only for about a millisecond before snapping closed again, which limits the spike's duration. Second, there is another type of voltage-sensitive ion channel involved. This one is also activated by positive-going changes in membrane potential, but it opens more slowly and when it opens, potassium ions rush out of the neuron. The loss of positively charged potassium ions from inside the cell makes the membrane potential more negative, causing the downstroke of the spike as the membrane potential returns to rest.

The axon hillock, where the spike originates, is the first stretch of a long highway to the axon terminal. Fortunately, the voltage-sensitive sodium channel's positive feedback loop allows the spike to travel along the axon. Sodium ions rushing in make the outer membrane more positive not just at the axon hillock, but also at the next bit of axon, farther from the cell body. Because the membrane in this next bit of axon also has voltage-gated sodium channels, they will open, sodium ions will rush in at that location and produce more positive charge in yet a further bit of axon membrane, and so on. In this manner, the spike travels down the axon like a flame racing along a fuse, each bit of axonal membrane "igniting" the next until the spike reaches the axon terminals.

The voltage-sensitive sodium channel that initiates neuronal spikes is a key target of neurotoxins generated by many plants and animals. Interfere with that channel and you block essentially all signaling in the brain (and the rest of the nervous system too). The most famous—or infamous—toxin is that of the

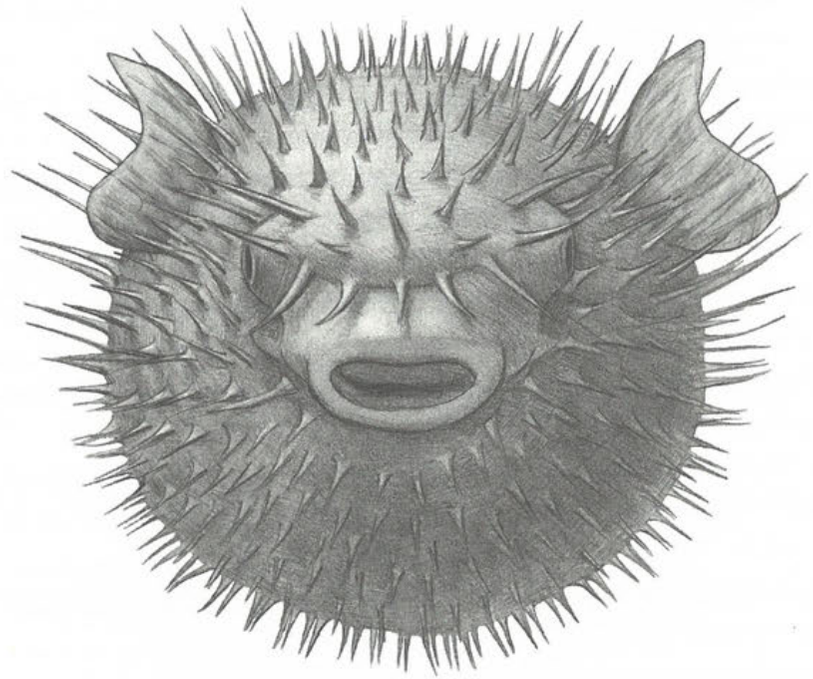


FIGURE 2.7. The pufferfish. *Joan M. K. Tycko, illustrator.*

fugu, the Japanese pufferfish (Figure 2.7). This toxin (called tetrodotoxin) is a tiny molecular plug that fits exactly into the outer portion of the sodium channel's central pore, thereby stopping it up. Tetrodotoxin is more than 1,000 times as powerful as cyanide and a single pufferfish has enough to kill 30 people. Considered a delicacy in Japan, pufferfish killed many people before preparation of fugu in restaurants was closely regulated by law to prevent people from ingesting the parts of the fish that have the highest concentrations of the toxin. Even today fugu is the one food the emperor and his family are prohibited from eating.

But let's return to movement along the axon when it is not interrupted by neurotoxins or other means. It is tempting to say that the axon is like an insulated copper electrical wire. But this obscures one of the fundamental inefficiencies of neurons. Copper wire need not do anything to keep electrical signals moving: it is totally passive, is a good conductor, and is well insulated against losing electrical charge to the outside. As a consequence, electrical signals in copper wires move at nearly the speed of light, about 669 million miles per hour. In contrast, the axon uses molecular machines with moving parts (voltage-sensitive ion channels snapping open and closed) to maintain the spike as it travels down its pathway. Comparatively, the axon is a quite poor conductor. The saltwater solution on the inside of the axon is not nearly as good a conductor as copper. Moreover, the outer membrane of the axon is a rather leaky insulator.

Perhaps the conduction of electrical signals along the axon is best understood through a hydraulic analogy. Insulated copper wire is like a steel water pipe (does not leak) that is 10 feet in diameter (great flow through its core), while the axon is like a "soaker" garden hose, 1 inch in diameter (poor flow through its core), that has been riddled with tiny holes along its length (leaks like hell) to allow you to irrigate a flower bed. This combination of poor core flow and leakiness makes water flow through a soaker hose slowly. Similarly, electrical current flow through an axon is also restricted by poor core flow and leakiness. As a consequence, electrical signals in axons typically travel slowly, at about 100 miles per hour. There is, however, quite a range, with the thinnest, uninsulated axons poking along at about 1 mile per hour and the very fastest (thick axons or those well insulated by neighboring glial cells) going at about 400 miles per hour. Nonetheless, even the very fastest axons, like those involved in reflexively withdrawing your finger from a hot stove, are conducting electrical signals at less than one-millionth the speed of copper wires.

Another way that our neurons differ from man-made devices, such as computers, to which they are often compared, involves the temporal range of their signals. The pattern of spike firing is the main way neurons encode and convey information, so timing limits on spike firing are particularly important. A desktop computer's central processing unit (circa 2006) may conduct 10 billion operations per second, but a typical neuron in a human brain is limited to around 400 spikes per second (though some special neurons, such as those in the auditory system that encode high-frequency sound, can fire up to 1,200 spikes per second). Furthermore, most neurons cannot sustain these highest rates for long (more than a few seconds) before they need a rest. With such constraints on speed and timing, it seems amazing that the brain can do what it does.

TO RETURN TO our neuronal story, we last left the spike racing down the axon highway to meet its fate. When the spike reaches the axon terminal it produces its characteristic explosive positive deflection in membrane potential. But, in the terminal, in addition to causing voltage-sensitive sodium channels to open, this voltage change also opens another class of ion channels that selectively pass calcium ions. Like sodium ions, calcium ions are positively charged (they have a charge of +2) and have a much higher concentration outside the cell than inside. So, like sodium ions, they too rush inside when a calcium channel is opened.

When calcium ions rush into the terminal they not only produce positive deflection in membrane potential, but also trigger unique biochemical events. Special sensor proteins for calcium ions are built into the neurotransmitter-containing synaptic vesicles. These sensors, upon binding calcium ions, set in motion a complex biochemical cascade that results in the presynaptic vesicle contacting a specialized patch of membrane called the release site and then fusing with it. Fusion of a vesicle causes the formation of a structure that resembles

the Greek capital letter omega (Ω), which allows the contents of the vesicle, the glutamate molecules, to diffuse into the synaptic cleft and ultimately bind postsynaptic receptors (see Figure 2.3). In this way, the cycle of neuronal signaling from EPSP to spike to glutamate release to EPSP is completed and information is conveyed from neuron to neuron.

ALBERT EINSTEIN, in an oft-quoted critique of Werner Heisenberg's Uncertainty Principle, said, "God does not play dice with the Universe." By the standards of modern physics, Einstein turned out to be wrong. If I were to make the related statement "Our brains do not play dice with our synapses," it would also be wrong. At most synapses in the brain, when a spike invades the presynaptic axon terminal and causes influx of calcium ions, this does not necessarily result in vesicle fusion and the release of neurotransmitter. It is, quite simply, a matter of chance. The probability of neurotransmitter release for a single spike might be 30 percent at an average synapse in the brain. Some synapses have release probabilities as low as 10 percent and a few release neurotransmitter every single time (100 percent probability), but these are the exceptions, not the rule. Most synapses in our brains do not function reliably: rather, they are probabilistic devices.

OUR IMAGINARY EXPERIMENT has now revealed the entire cycle of electrical signaling in neurons. This is a basic template that can be used to understand many brain phenomena. But the situation is a bit more complicated than shown by just this one example. Glutamate opens an ion channel that lets positive charge into the cell. This tends to move the membrane potential in a positive direction, close to the level where a spike will fire, referred to as excitation (as in *excitatory* postsynaptic potential, EPSP). There are other neurotransmitters that produce the opposite effect, inhibition, where the probability of the

postsynaptic cell's firing a spike is reduced. For example, the major inhibitory neurotransmitter in the brain is gamma-aminobutyric acid, abbreviated GABA. GABA binds a receptor that opens a channel that lets chloride ions flow into the postsynaptic neuron. Chloride ions have a negative charge (-1), and thus make the membrane potential more negative. This, not surprisingly, is called an inhibitory postsynaptic potential, or IPSP, and makes it even harder for the postsynaptic neuron to fire a spike.

In practice, whether or not a neuron fires a spike at any given moment is determined by the simultaneous action of *many* synapses, with excitatory and inhibitory actions summed to produce the total effect. Recall that the average neuron in the brain receives 5,000 synapses. Of these, about 4,500 will be excitatory and 500 will be inhibitory. Although only a small number are likely to be active at any one time, most neurons will not be driven to fire a spike from the brief action of a single excitatory synapse, but will require the simultaneous action of about 5 to 20 synapses (or even more in some neurons).

Glutamate and GABA are fast-acting neurotransmitters: when they bind their receptors, the electrical changes they produce occur within a few milliseconds. They are the dominant fast neurotransmitters in brain, but there are some other fast ones. Glycine is an inhibitory neurotransmitter that acts like GABA: it opens a receptor-associated ion channel to let chloride ions rush in and inhibit the postsynaptic neuron. The poison strychnine, which figures prominently in mystery novels, blocks glycine receptors and prevents their activation. Another example is acetylcholine, an excitatory neurotransmitter that, like glutamate, opens an ion channel that lets both sodium rush in and potassium out. This occurs in some parts of the brain, as well as at the synapses between neurons and muscles. The South American hunting arrow poison called curare blocks this receptor. Animals shot with a curare-tipped arrow become totally limp as commands from the nerves fail to activate muscular contraction.

In addition to the fast neurotransmitters, such as glutamate, GABA, glycine, and acetylcholine, there are also other neurotransmitters that act more slowly. These neurotransmitters bind a different class of receptors. Instead of opening ion channels, they activate biochemical processes inside the neurons. These biochemical events produce changes that are slow to start but that have a long duration: typically, from 200 milliseconds to 10 seconds. Many of these slow-acting neurotransmitters do not produce a direct electrical effect: the membrane potential does not change in either the positive or the negative direction after they bind their receptor. Rather, they change the electrical properties of the cell in ways that are only apparent when fast neurotransmitters also act. For example, the slow-acting neurotransmitter called noradrenaline can change the voltage at which a spike will be triggered from its normal level of −60 millivolts to −65 millivolts. In a neuron that is silent, there won't be any difference after noradrenaline release, but when that neuron receives fast synaptic input, there will be. If glutamate is released onto this neuron from synapses and this changes its membrane potential from the resting state of −70 millivolts to −65 millivolts, this will now result in a spike. This same action of glutamate in the absence of noradrenaline would fail to trigger a spike. In biochemical terms, we would say that noradrenaline has a modulatory action on spike firing: it doesn't directly cause spike firing but it changes the properties of spike firing produced by other neurotransmitters. The bottom line here is that fast neurotransmitters are suited to conveying a certain class of information that requires rapid signals, while slow neurotransmitters are better at setting the overall tone and range.

WHEN NEUROTRANSMITTERS are released into the synaptic cleft, they eventually diffuse away, achieving a low concentration. A while back, I invoked the image of a single drop of red wine released into a full water glass that, eventually, will turn the contents of the glass a very pale pink. This is fine if

neurotransmitters were released only once. But, over time, if neurotransmitter molecules are repeatedly released, there must be some mechanism to clear the neurotransmitter from the cerebrospinal fluid surrounding brain cells before it achieves dangerously high concentrations (continuous activation of neurotransmitter receptors can often kill neurons). In terms of our wine glass image, with repeated drops the wine glass would eventually turn a uniform shade of pink and then red.

Essentially, when it comes to cleaning up after neurotransmitter release, someone has to take out the trash. For some neurotransmitters, there is the quintessentially American solution: burn that junk in the front yard. For example, acetylcholine is destroyed in the synaptic cleft by an enzyme specifically built for that purpose. Most other neurotransmitters get the European treatment: they are recycled. Glutamate molecules, through the actions of specialized transporter proteins in the outer membrane, are taken up into glial cells, where they undergo some biochemical processing before being sent to neurons for re-use. Most of the slow-acting neurotransmitters, such as dopamine and noradrenaline, are taken up right back into axon terminals, where they can be repackaged into vesicles and used again. Interestingly, GABA seems to go both ways: it is taken up by both axon terminals and glial cells. Some neurotransmitter transporters make excellent targets for psychoactive drugs (such as the antidepressant Prozac and its relatives) because blocking them will cause neurotransmitters in the synapse to linger and achieve higher concentrations.

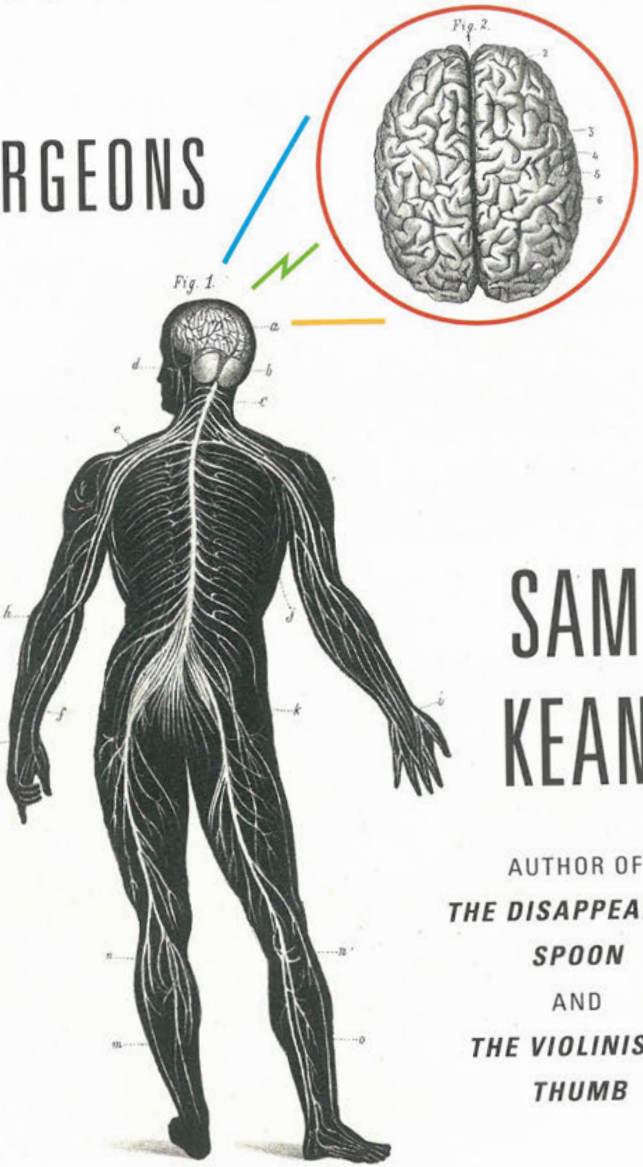
ALL THE INFORMATION in your brain, from the sensation of smelling a rose, to the commands moving your arm to shoot pool, to that dream about going to school naked, are encoded by spike firing in a sea of neurons, densely interconnected by synapses. Now that we have gained an overall understanding of electrical signaling in the brain, let's consider the challenges the brain must con-

front as it tries to create mental function using a collection of less-than-optimal parts. The first challenge is the limitation on the rate of spike firing caused by the time it takes for voltage-sensitive sodium and potassium ions to open and close. As a result, individual neurons are typically limited to a maximal firing rate of about 400 spikes/second (compared with 10 billion operations/second for a modern desktop computer). The second challenge is that axons are slow, leaky electrical conductors that typically propagate spikes at a relatively sedate 100 miles per hour (compared with electrical signals in a man-made electronic device moving at around 669 million miles per hour). The third challenge is that once spikes have made it to the synaptic terminal, there is a high probability (about 70 percent on average) that the whole trip will have been in vain, and no neurotransmitters will be released. What a bum deal! These constraints may have been tolerable for the simple problems solved by the nervous system of a worm or a jellyfish, but for the human brain, the constraints imposed by (ancient) neuronal electrical function are considerable.

How does the brain manage to create human mental function with neurons that are such crummy parts? More to the point, given the comparisons above, how is it that our brains can easily accomplish certain tasks that typically baffle electronic computers—for example, recognizing instantly that an image of a Rottweiler taken from the front and another of a teacup poodle taken from the rear should both be classified as "dog"? This is a deep question, central to neurobiology, for which a detailed answer is not at hand. Yet a more general explanation appears to be as follows. Individual neurons are horribly slow, unreliable, and inefficient processors. But the brain is an agglomeration of 100 billion of these suboptimal processors, massively interconnected by 500 trillion synapses. As a result, the brain can solve difficult problems by using the simultaneous processing and subsequent integration of large numbers of neurons. The brain is a kludge in which an enormous number of interconnected processors can function impressively even when each individual processor is severely limited.

In addition, while the overall wiring diagram of the brain is laid down in the genetic code, the fine-scale wiring of the brain is guided by patterns of activity, which allows the strength and pattern of synaptic connections to be molded by experience, a process called synaptic plasticity (which I will consider in Chapters 3 and 5). It is the massively interconnected parallel architecture of the brain combined with the capacity for subtle rewiring that allows the brain to build such an impressive device from such crummy parts.

THE TALE OF THE DUELING NEUROSURGEONS



THE HISTORY OF THE HUMAN BRAIN AS REVEALED BY TRUE STORIES OF TRAUMA, MADNESS, AND RECOVERY

SAM KEAN

AUTHOR OF
THE DISAPPEARING SPOON
AND
THE VIOLINIST'S THUMB

GUIDE INSIDE
READERS' PICK

"Entertaining....Kean proves an able guide, connecting each story with the science behind it, always with an air of enthusiastic curiosity." –BOSTON GLOBE

CHAPTER TEN

Honest Lying

Almost every structure we've examined so far contributes to forming and storing memories. Memory is therefore a wonderful way to see how different parts of the brain work together on a large scale.

Soldiers buried more than men in the graves of Southeast Asia. While conquering Singapore in February 1942, Japanese soldiers captured 100,000 mostly British POWs, more than they knew what to do with. The military worked thousands of them to death on the brutal Burma–Siam "Death Railway," a project that required hacking through 250 miles of mountainous jungles and constructing bridges over rivers like the Kwai. Most of the remaining captives, including many doctors, were crowded into the notorious Japanese prison camps. In fact, two British doctors incarcerated in the Changi camp, Bernard Lennox and Hugh Edward de Wardener, realized that their captors were essentially running a gruesome experiment: taking healthy men, depriving them of one nutrient, and watching their brains deteriorate.

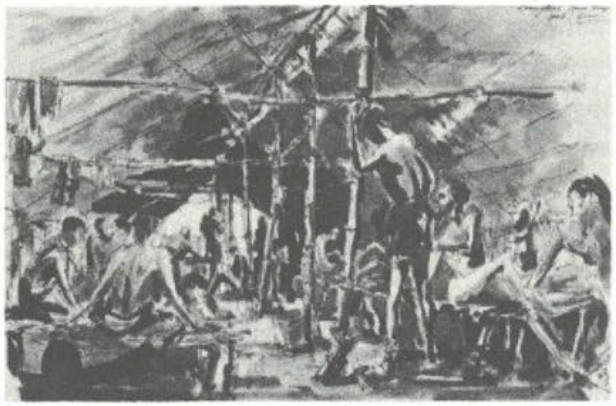
No matter his background, every doctor in the camps worked as a surgeon, dentist, psychiatrist, and coroner, and they suffered from the same ailments—dysentery, malaria, diphtheria—that ravaged the troops. They pared down bamboo shards for needles, unstitched parachutes for silk sutures, and drained human stomachs for acids. Monsoons tore through their "clinics"—often just tents draped over poles—and some doctors faced beatings and threats of being boiled in oil if they didn't cure enough soldiers to meet work quotas. Guards made things worse by restricting sick men to half rations, to "motivate" them to recover. But even among the healthy, the food—mostly plain rice—was never adequate, and led to beriberi disease.

For as long as people have eaten rice in Asia, doctors there have reported outbreaks of beriberi. Symptoms included heart trouble, anorexia, twitching eyes, and legs so swollen that the skin sometimes burst. Victims also walked with a shuffling, staggering gait that reminded locals of *beri,* sheep. When Europeans colonized Southeast Asia in the 1600s, their doctors began seeing cases as well; one early

report came from Dr. Nicolaes Tulp, the Dutchman later immortalized in Rembrandt's *The Anatomy Lesson*. But the number of cases exploded after the introduction into Asia, in the later 1800s, of steam-powered rice mills. The mills removed the outer husks from rice grains, producing so-called white rice. People back then called it polished rice, and cheap polished rice became a dietary staple—or, often, *the* diet—of peasants, soldiers, and prisoners. During the Russo-Japanese War alone, 200,000 Japanese troops fell victim to beriberi.

Scientists eventually began to suspect that beriberi was a nutritional deficiency—probably a lack of vitamin B_1 (a.k.a. thiamine). In shucking off the nutritious rice husks, the mills stripped out almost all the B_1, and many people didn't get enough thiamine from eating vegetables, beans, or meat. Our bodies use B_1 to harvest energy from glucose, the end result of digesting carbohydrates. Brain cells especially rely on glucose for energy, since other sugars cannot cross the blood-brain barrier. The brain also needs thiamine to make myelin sheaths and to build certain neurotransmitters.

The first cases of beriberi appeared two weeks after the Changi camp opened, among a few alcoholics cut off cold turkey. Many more cases appeared after another month. Doctors tended to the ailing as best they could and sometimes kept their spirits up by lying about the



A hospital in a Japanese POW camp in Singapore.

progress of Allied armies. When all else failed, some doctors ordered men to live or face court-martial (a threat reminiscent of those old medieval laws that made suicide illegal). Nevertheless, by June 1942 there were a thousand beriberi cases in Changi alone. Helpless to stop the epidemic, de Wardener and Lennox started doing autopsies in secret and collecting tissues from the brains of beriberi victims, to study the pathology of the disease.

Although considered contraband, these tissues and autopsy records were mostly safe inside Changi. But in 1943 Lennox and de Wardener were herded off to different camps near the Death Railway in Siam and had to split their medical stash. Wary of confiscation, Lennox arranged to smuggle the brain tissues out of his camp, only to have them perish in a train wreck. De Wardener guarded the all-important paper records, a four-inch sheaf. But as the war turned sour for Japan in early 1945, de Wardener realized that Japanese leaders wouldn't look kindly on hard evidence of starving POWs. So when he received another transfer order—and saw guards frisking his fellow-transferees and searching their belongings—he made a hasty decision. He had a metallurgist friend seal his papers inside a four-gallon petrol tin. He then wrapped the tin in a cape and buried the bundle three feet deep in a fresh grave, leaving only the dead soldier as a sentinel. To remember which grave it was—there were so many—he and some friends took compass bearings on a few enormous trees nearby. As he departed camp, de Wardener could only pray that the heat, rot, and miasma of Siam wouldn't eat through the bundle before he returned. If he returned.

The records were precious because they resolved a half-century-long dispute about the brain, B_1, and memory. In 1887 a Russian neuroscientist named Sergei Korsakoff described a peculiar ailment among alcoholics. Symptoms included emaciation, staggering, a lack of the patellar kick reflex, and urine "as red as the strongest tea." But the outstanding symptom was memory loss. Korsakoff's patients could play chess, banter, make wisecracks, and reason properly—but couldn't remember the previous day, even the previous hour. During

conversations they repeated the same anecdotes over and over, verbatim. And if Korsakoff left the room for a spell, they repeated the same anecdotes over and over, verbatim, when he returned. Other brain diseases cause memory loss, of course, but Korsakoff noticed something distinctive about these cases. If asked a question they can't answer, most people with memory loss admit they don't know. Korsakoff's patients never did—they always lied instead.

Today, Korsakoff's syndrome—the tendency to lie compulsively due to brain damage—is a well-recognized ailment. And truth be told, it can be quite entertaining, in a gallows-humor way. When asked why Marie Curie was famous, one Korsakoff victim declared, "Because of her hairstyle." Another claimed to know Charlemagne's favorite meal ("maize porridge") and what color horse King Arthur rode ("black"). Victims lie especially often about their personal lives. One man claimed to remember, thirty years later, what he wore the first day of summer in 1979. Another told his doctor, in consecutive sentences, that he'd been married for four months and that he'd sired four children with his wife. After a quick calculation, he marveled at his sexual prowess: "Not bad."

Beyond the occasional Münchhausenian whopper, most Korsakoff victims tell plausible, even mundane lies: unless you knew their life histories, you'd never peg them as bullshit artists. Unlike most of us, they don't lie to make themselves look good, or to get an edge, or to conceal something. And unlike people suffering from delusions, they don't defend themselves ferociously if called out; many just shrug. But no matter how many times someone catches them, they keep lying. This fibbing for no obvious or underhanded reason is known as confabulation.

Korsakoff focused on the psychology of confabulation, but other scientists extended his work in the early 1900s and started linking these psychological symptoms to specific brain damage. In particular, they discovered tiny hemorrhages in the brains of victims, as well as patches of dead neurons. Pathologists also linked Korsakoff's syndrome to another, related disease called Wernicke's syndrome. In fact, because Wernicke's syndrome often turns into Korsakoff's, the two were eventually yoked together as Wernicke-Korsakoff syndrome.

The underlying cause of Wernicke-Korsakoff syndrome took longer to suss out, but by the later 1930s a few scientists had linked it to a lack of B_1. As doctors now know, alcohol prevents the intestines from absorbing the thiamine in food. This shortage then causes changes inside the brain, especially to glial cells. Among other jobs, glial cells sponge up excess neurotransmitters from the synapses between neurons. And without thiamine, the glia cannot sop up glutamate, which stimulates neurons. As a result of this excess, neurons get overstimulated and eventually exhaust themselves, dying of excitotoxicity.

Because they seemed to share a common root—B_1 deficiency—beriberi and Wernicke-Korsakoff syndrome should have caused similar symptoms and similar destruction inside the brain. But through the 1940s no one had any hard evidence to link them. This was partly because Wernicke-Korsakoff remained rare and associated primarily with alcoholics, and partly because doctors who studied beriberi focused on nerve and heart damage, not brain damage. The net result was confusion: was this one disease or two? More important, it highlighted a growing concern over efforts to link physiology and psychology: many doctors frankly doubted that the lack of a simple vitamin—a molecular problem—could leap up so many levels of scale and cause complex mental troubles like confabulation.

Changi proved it could. Among the thousand-plus beriberi victims there, several dozen also came down with symptoms of Wernicke-Korsakoff syndrome, including confabulation. As an example, de Wardener asked one far-gone man, just to test his mental state, "Do you remember when we met in Brighton? I was riding a white horse and you a black horse, and we rode on the beach." This was bunk, but the man answered that of course he remembered, and filled in the details. Such imaginings often became the patients' reality, sadly, and a few men died in this state—their last "memories" nothing but

vapors and fabrications. Medically, the fact that beriberi always preceded Wernicke-Korsakoff, and that those with the worst beriberi got the worst Wernicke-Korsakoff, implied a common cause. Autopsies then cemented the link: even without a microscope, Lennox, a trained pathologist, could see the characteristic hemorrhages and patches of dead neurons in the brains of victims. Beriberi and Wernicke-Korsakoff seemed to be two stages, chronic and acute, of the same underlying disease.

As further evidence, treating victims with pure thiamine (some doctors had tiny stashes) usually relieved the symptoms of both Wernicke-Korsakoff and beriberi, sometimes within hours: de Wardener remembers a few men roaring to life and consuming whole mountains of rice to combat their sudden hunger. (Mental symptoms such as confabulation might take several weeks to dissipate.) For less acute cases doctors might add Marmite to meals (however unappetizing, this yeast-based extract is lousy with B_1) or ferment rice and potatoes to cultivate wild yeast, also chock-full of B_1. Some doctors sent men to gather thiamine-rich hibiscus leaves as well. The smarter doctors lied to the men and claimed that hibiscus would pump up their libidos for when they got back home to their gals. After that, troops no doubt couldn't consume enough hibiscus.

In tandem, the fact that consuming too little thiamine provoked Wernicke-Korsakoff, and that restoring thiamine to the diet relieved it, convinced Lennox and de Wardener that the lack of a simple nutrient could indeed destroy something as profound as our memories, even our sense of truth. But the duo still had to make their case to the medical world—which meant not only surviving the camps but preserving their autopsy files. This wasn't easy in a war zone, and as de Wardener discovered, about the only way to conceal such things was to bury them, and pray to God they survived.

After V-J Day, de Wardener received mysterious orders to report to Bangkok. Although anxious to start searching for his files, he remembers enjoying the journey: "I took a victorious ride across Siam

in a Jeep, with all the Nips bowing... which was very satisfying." To his surprise, he found his records waiting for him at Bangkok HQ. Apparently a friend had returned to Changi with a shovel not long before, scrabbled through the dirt above the dead sentinel's body, and liberated the bundle. It was a close thing: the cape had rotted away and the solder sealing up the tin had disintegrated. But the papers had survived, perhaps by a matter of days. Lennox and de Wardener finally published this, well, groundbreaking work in 1947.*

~

Since World War II, neuroscientists have continued to mine confabulation for insight into how memory works, and it has proved a rich vein indeed. For example, confabulations reveal that each memory seems to have a distinct time stamp, like a computer file. And just like computer files, that time stamp can be corrupted. Most confabulators tell plausible lies; in fact, many of their false "memories" did happen to them at some point. But confabulators often mistake *when* the memory happened: the scenes in their lives have been shuffled wrong. So while they claim they ate truffled duck last night, in truth they did that thirty years ago while honeymooning in Paris. In some sense, then, confabulation is a breakdown in the ability to tell a coherent story about our lives.

The fact that virtually all confabulators have frontal lobe damage also tells us something. The frontal lobes help coordinate multistep processes, and despite how effortless memory seems, remembering something specific (say, the worst Christmas present you ever got) is complicated. The brain has a fraction of a second to search for the memory, retrieve it, replay it, and summon up the proper sensations and emotions—and that's assuming you recorded the memory accurately in the first place. If the frontal lobes suffer damage, any one of those steps can go awry. Perhaps confabulators simply retrieve the wrong memory each time they "recall" something, and don't recognize their error.

Some scientists trace confabulations to shame and a need to cover up deficiencies. Confabulators don't generally blurt things out unprovoked; you have to ask questions to elicit the lie. And according to this theory, admitting they don't know something upsets and embarrasses people, so they pretend. For example, most doctors ask at intake how many children someone has. Having to admit "I don't know" could be catastrophic to a person's well-being, since what kind of monster doesn't remember his own children? In short, confabulations could be a defense mechanism, a way for people to hide their brain damage, even from themselves.

As another defense mechanism, some confabulators invent fictional characters and foist their personal failings onto them. One alcoholic confabulator raved to his doctor about imps who kept breaking into his apartment, even after he changed the locks, and stealing things like his remote control. He eventually heaved the imps outdoors on a brutal January night. But, feeling guilty, he braved the weather and draped clothes over them later that night, then called an ambulance. In reality medical workers had discovered *him* outside that winter, stone drunk and mostly naked. In telling the story he was basically confabulating an allegory on the fly. That's a remarkable deed for someone with brain damage, and the ruse allowed him to ponder his own flaws more objectively, without implicating himself.

As that last case shows, it's not always clear whether confabulators understand that they're lying. Most seem blithely unaware, and many neuroscientists insist that Korsakoff patients don't realize what's happening. But is that possible? Covering up a memory gap, even subconsciously, implies that they know on some level that the gap exists. Which means they know and don't know at the same time. It's a doozy of a conundrum, and it raises all sorts of stoner questions about whether you can truly deceive yourself, and more broadly about the nature of truth and falsehood. Consider asking a confabulator what she ate for breakfast. If she hasn't the foggiest, she might blurt, "Leftover pizza." But of course it's possible she did have cold pizza for

breakfast, in which case she would be telling the truth—even though her brain tried, consciously or not, to put one over on you. What on earth would you call that? Neither *lying* nor *telling the truth* quite encompasses it. It's slipperier, and some neuroscientists have taken to calling it "honest lying."

Philosophical conundrums aside, work on confabulation helped make memory a proper object of neuroscientific study last century, since scientists could finally link memory to the brain and its biology. That said, the biggest breakthrough in memory research in the past hundred years didn't spring from the minds of confabulators. Indeed, most memory work until the 1950s relied on a flawed assumption—that all parts of the brain contribute equally to forming and storing memories. It's an idea that took something drastic, a botched operation by a lobotomist, to overturn.

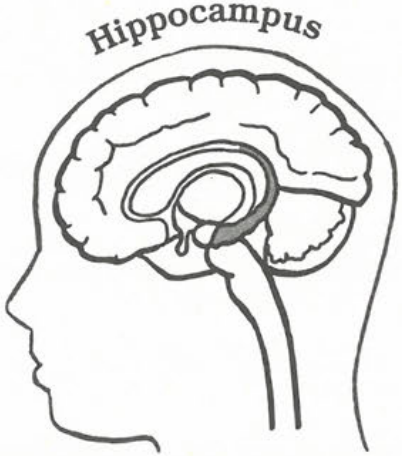
~

In the early 1930s a bicyclist in Connecticut struck a small boy, who tumbled and cracked his skull. No one knows whether the accident alone caused his epilepsy—three cousins had it, so he might have been predisposed—but the blow probably precipitated it, and at age ten he started having seizures. Each lasted around forty seconds, during which time his mouth flopped open, his eyes slipped shut, and his arms and legs crossed and uncrossed as if curled by an invisible puppeteer. He suffered his first grand mal on, of all days, his fifteenth birthday, while riding in the car with his parents. More followed, in class and at home and while shopping—up to ten seizures a day, with at least one major episode per week. So at an age when most people are struggling to find an identity, he was saddled with one he didn't want: the kid who shook, who bit his tongue, who slumped over and blacked out and pissed himself. The mockery got so bad he dropped out of high school, and he earned his diploma only at twenty-one, from a different school. He ended up living at home and working in a motor shop.

Finally the desperate young man—soon immortalized as H.M.—decided to try surgery. When younger, H.M. had dreamed of practicing neurosurgery himself and studying how the brain works. But while H.M. did end up contributing, profoundly, to neuroscience, his affliction ensured that he would never grasp his own importance.

H.M. started seeing Dr. William Scoville around 1943. A noted daredevil—before a medical conference in Spain once, he'd stripped off his jacket and mixed it up with the *toros* in the bullring—Scoville liked risky surgeries, too, and had jumped onto the American lobotomy bandwagon* early. But he disliked the drastic changes in his patients' personalities, so he began experimenting with "fractional" lobotomies, which destroyed less tissue. Over the years he basically worked his way around the brain, carving out this piece or that and checking the results, until he finally reached the hippocampus.

Because it was part of the limbic system, scientists at the time believed that the hippocampus helped process emotions, but its exact function remained unknown. Rabies often destroyed it, and James Papez had singled it out for attention. (A poetaster, Papez even penned a ditty to his wife that read: "It's Pearl, my girl on Broad Street / that I



miss... My hippocampus tells me this.") Scoville was less enamored: he'd seen the mental turmoil that hippocampus damage could cause. So in the early 1950s he started removing the hippocampi (you have one in each hemisphere) from a few psychotics. Although it was hard to be sure in people with such disturbed minds, they seemed to suffer no ill effects, and two women in particular showed a marked reduction in seizures. Unfortunately Scoville neglected to do careful follow-up tests until November 1953—after he'd convinced H.M. to try the surgery.

H.M.'s operation took place in Hartford, Connecticut, on September 1, 1953. Scoville peeled back his patient's scalp, then used a hand crank and one-dollar drill saw from a local hardware store to remove a bottle cap's worth of bone from above each eye. As cerebrospinal fluid drained away, the brain settled down in its cavity, giving Scoville more room to work. With what looked like an elongated shoehorn, he nudged aside H.M.'s frontal and temporal lobes and peered inside.

The hippocampus sits at ear level and has the rough shape and diameter of a curled thumb. Hoping to remove as little tissue as possible, Scoville first sparked each hippocampus with wires to find the origin of H.M.'s seizures. No luck, so he grabbed a long metal tube and began cutting and sucking out tissue gram by gram; he eventually removed three inches' worth of hippocampus on each side. (Two nubs of hippocampal tissue remained behind, but because Scoville also removed the connections between those nubs and other parts of the brain, the nubs were useless, like unplugged computers.) For good measure, Scoville removed H.M.'s amygdalae and other nearby structures as well. Given how deeply all these structures are embedded in the brain, only a neurosurgeon could have destroyed them with such precision.

Post-op, H.M. remained drowsy for a few days, but he could recognize his family and carry on a seemingly normal conversation. And by many measures, the operation succeeded. His personality never

changed; the seizures all but disappeared (two attacks per year at most); and when the fog of epilepsy lifted, his IQ jumped from 104 to 117. Just one problem: his memory was shot. Aside from a few small islands of recollection—like the fact that Dr. Scoville had operated on him—an entire decade's worth of memories from before the surgery had vanished. Equally terrible, he couldn't form new memories. Names escaped him now, as did the day of the week. He repeated the same comments over and over, verbatim, and while he might remember directions to the bathroom long enough to get there, he always had to ask again later. He'd even consume multiple lunches or breakfasts if no one stopped him, as if his appetite had no memory, either. His mind had become a sieve.

In light of modern knowledge, H.M.'s deficit makes sense. Memory formation involves several steps. First, neurons in the cortex jot down what our sensory neurons see and feel and hear. This ability to record first impressions still worked in H.M. But like messages scrawled on the beach, these impressions erode quickly. It's the next step, involving neurons in the hippocampus, that makes memories last. These neurons produce special proteins that encourage axon bulbs to swell in size. As a result, the axons can stream more neurotransmitter bubbles toward their neighbors. This in turn strengthens the synapse connections between those neurons before the memory decays. Over months and years—provided the first impression was strong enough, or we think about the event from time to time—the hippocampus then transfers the memory to the cortex for permanent storage. In short, the hippocampus orchestrates both the recording and the storage of memories, and without it, this "memory consolidation" cannot occur.

Scoville couldn't have known all this, but he'd clearly sabotaged H.M.'s memory, and he didn't know what to do. So a few months later, when he saw that Wilder Penfield was about to publish a report on hippocampus damage, Scoville called the renowned surgeon and confessed.

Penfield had recently operated on two patients with hippocampal epilepsy. To be safe, he'd removed the structure on just one side, but unbeknownst to him, the seizures had already destroyed the other hippocampus in each person. So removing the one left both patients without a working hippocampus, and they developed the purest amnesia Penfield had ever seen. Although he was still puzzling through the cases, a graduate student was going to present them at a scientific meeting in Chicago in 1954.

When Scoville called, Penfield reportedly flipped out, berating him for his recklessness. After calming down, though, the scientist in Penfield realized (much as the beriberi doctors had) that Scoville had actually performed an invaluable experiment: here was a chance to determine what the hippocampus did. As part of its mission Penfield's clinic in Montreal tracked the psychological changes that patients experienced after psychosurgery. So Penfield dispatched a Ph.D. student from the Neuro, Brenda Milner, down to Connecticut to investigate the hippocampusless H.M.

After his memory vanished, H.M. lost his job and had no choice but to keep living with his parents. He spoke in a monotone now and had no interest in sex, but otherwise seemed normal. To the neighbors, it probably just looked like he was loafing his life away. He took a part-time job packing rubber balloons into plastic bags, and did odd chores around the house. (Although his parents had to remind him where they kept the lawn mower every single time, he could actually mow just fine, since he could see what grass he hadn't cut.) His temper did flare up occasionally: his mother tended to nag, and he cuffed her a few times and kicked her shins. Another time, when an uncle removed a few choice rifles from the family's gun collection, he flew into a rage. (Despite his amnesia he retained a lifelong love of guns, and always remembered to renew his NRA membership.) But he whiled away most days peacefully, either doing crossword puzzles—working through the clues methodically, in order—or flopping in front of the television and watching either Sunday Mass or the old

movies that, to him, would never become classics. It was like early retirement, except for the days Milner arrived to test him.

Milner would take the night train down from Montreal to Hartford, arriving at 3 a.m. and spending the next few days with H.M. Her battery of tests confirmed Scoville's basic observations pretty quickly: H.M. had little memory of the past and no ability to form new memories going forward. This was already a big advance—proof that some parts of the brain, namely the hippocampus, contribute more to forming and storing memories than other parts. And what Milner discovered next redefined what "memory" even meant.

Rather than keep asking him questions he couldn't answer, she started testing H.M.'s motor skills. Most important, she gave him a piece of paper with two five-pointed stars on it, one nested inside the other: ☆. The outer star was about six inches wide, and there was a half-inch or so gap between them. The test required H.M. to trace a third star between the two with a pencil. The catch was, he couldn't see the stars directly: Milner had shielded the diagram, and he had to look at them in a mirror instead. Left was right, right was left, and every natural instinct about where to move his pencil was wrong. Anyone taking this mirror test for the first time makes a mess—the pencil line looks like an EKG—and H.M. proved no exception. Somehow, though, H.M. got better. He didn't remember any of the thirty training sessions Milner ran him through. But his unconscious motor centers did remember, and after three days he could trace the star in the mirror fluently. He even commented near the end, "This is funny...I would have thought it would be rather difficult, but it seems I've done pretty well."

Milner remembers the star test as a eureka. Before this, neuroscientists thought of memory as monolithic: the brain stored memories all over, and all memory was essentially the same. But Milner had now teased apart two distinct types of memory. There's declarative memory, which allows people to remember names, dates, facts; this is what most of us mean by "memory." But there's also procedural

memory—unconscious memories of how to pedal a bicycle or sign your name. Tracing the stars proved that H.M., despite his amnesia, could form new procedural memories. Procedural memories must therefore rely on distinct structures within the brain.

This distinction between procedural and declarative memories (sometimes called "knowing how" versus "knowing that") now undergirds all memory research. It also sheds light on basic mental development. Infants develop procedural memory early, which explains why they can walk and talk fairly quickly. Declarative memory develops later, and its initial weakness prevents us from remembering much from early childhood.

Another distinct type of memory emerged from Milner's tests as well. One day Milner asked H.M. to remember a random number, 584, for as long as possible. She then left him alone for fifteen minutes while she had a cup of coffee. Contrary to her expectation, he still knew the number when she returned. How? He'd been repeating it under his breath, over and over. Similarly, H.M. could remember the words "nail" and "salad" for several minutes by imagining a nail piercing some salad greens and reminding himself over and over not to eat the impaled leaves. Any distraction during those minutes would have ejected the words clean out of H.M.'s mind, and five minutes after the test ended, even the memory of having to remember something had vanished. Nevertheless, as long as H.M. concentrated and kept refreshing his memory, he could hold on. This was the first clue that short-term memory exists; moreover, it showed that short-term memory (which H.M. had) and long-term memory (which he lacked) must utilize different brain structures.

After Milner's discoveries, H.M. became a scientific celebrity, and other neuroscientists began clamoring to explore his unique mind. He did not disappoint. In April 1958, five years after the operation, H.M. and his parents moved into a small Hartford bungalow. In 1966 a few American neuroscientists asked him to draw the home's floor plan from memory. He succeeded. He didn't know the bungalow's address,

but walking through its six rooms over and over had tattooed the layout into his brain. This proved that our spatial memory systems, while normally reliant on the hippocampus, can circumvent it if need be (probably via the parahippocampus, a nearby navigation center).

Scientists also discovered that time worked differently for H.M. Up to about twenty seconds, he reckoned time as accurately as any normal person. After that, things veered wildly. Five minutes lasted, subjectively, just forty seconds for him; one hour lasted three minutes; one day fifteen minutes. This implies that the brain uses two different timekeepers—one for the short term and one for everything beyond twenty seconds, with only the latter suffering damage in H.M. Once again, H.M. allowed scientists to break a complex mental function down into different components and to link those components to structures in the brain. Eventually more than one hundred neuroscientists examined H.M., making his probably the most studied mind in history.

All the while H.M. got older, at least physically. Mentally, he remained stuck in the 1940s. He remembered not a single birthday or funeral after that time; the Cold War and sexual revolution never registered; new words such as *granola* and *Jacuzzi* remained forever undefined. Worse, a vague sense of uneasiness often bubbled up inside him, and he could never quite shake it. The feeling, Milner reported, was "like that fraction of a second in the morning, when you are in a strange hotel room, before it all falls in[to] place." Only for H.M. it never did.

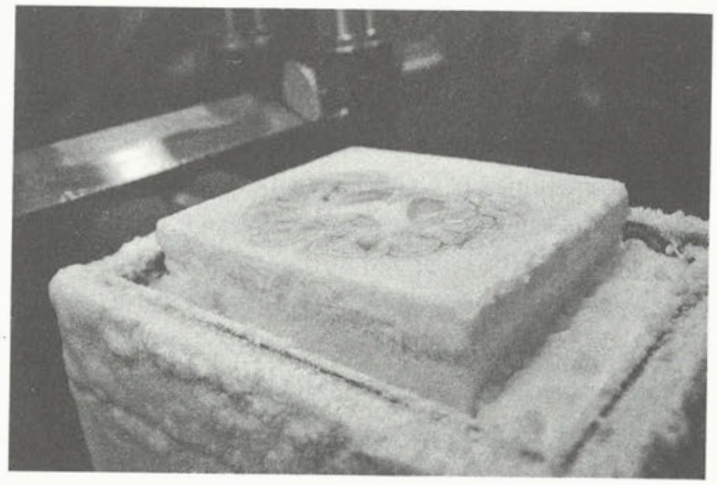
In 1980, after H.M.'s father died and his mother got too sick to care for him, he moved into a nursing home. He walked a little gimpily by that point: years of taking heavy-duty epilepsy drugs had withered his cerebellum, and his wide, shuffling gait resembled that of kuru victims. He also got pretty portly after too many forgotten second helpings of cake and pudding. But overall he was a fairly normal patient and lived a (mostly) placid life. He loafed through the nontesting days reading poems or gun magazines, watching trains rumble by,

and petting the dogs, cats, and rabbits the facility owned. He learned how to use a walker, thanks to his intact motor memories, and he even attended his thirty-fifth high school reunion in 1982. (Although he recognized no one there, other attendees reported the same problem.) When he dreamed at night, he often dreamed of hills—not of struggling up them, but cresting them and being at the top.

Still, the old, volatile H.M. did flare up now and again. He sometimes refused to take his meds—at which point his nurses scolded him, warning him that Dr. Scoville would get angry if he disobeyed. (That Scoville had died in a car crash didn't matter. H.M. always fell for it.) He got into fights with other residents as well. One harpy at the nursing home would erase his bingo card midgame and taunt him. H.M. sometimes responded by running to his room and either banging his head on the wall or grabbing his bed and shaking it like a gorilla would its cage. One fit got so violent that his nurses called the police. These were moments of pure animal frustration—and yet in some ways they seem like his most human moments. For a few seconds a real person broke through the dull, bovine exterior. He was reacting the way we'd all want to if dealt his fate: he raged.

As soon as a nurse distracted H.M., he forgot his torment, of course. And aside from those flare-ups he lived a quiet life, albeit in declining health. He finally died in 2008, aged eighty-two, of respiratory failure—at which point scientists revealed him to the world as Henry Gustav Molaison.

The world of neuroscience mourned Molaison: his death led to numerous tributes about his patience and kindness, as well as scores of puns about his being unforgettable. And his brain is still providing insight today. Before his death, his nursing home had started stockpiling ice packs in preparation; when he passed, employees ringed his skull with them to keep his brain cool. Doctors soon arrived to claim the body, and that night they scanned his brain in situ and then liberated it. After two months hardening in formalin, it was flown cross-country in a cooler (which got the window seat) to a brain institute in



The brain of H.M., the unforgettable amnesiac, being sliced in preparation for future study. (Courtesy Jacopo Annese, the Brain Observatory, San Diego)

San Diego. Scientists there soaked it in sugar solutions to draw out excess water, then froze it to solidify it. Finally, they used the medical equivalent of a deli slicer to shave Molaison's brain into 2,401 slices, each of which they mounted on a glass plate and photographed at 20x magnification, to form a digital, zoomable map down to the level of individual neurons. The slicing process was broadcast live online, and 400,000 people tuned in to say goodbye to H.M.

~

Although H.M. dominated the scientific literature and popular imagination, plenty of other amnesiacs have contributed to our understanding of memory. Take K.C., an amnesiac in suburban Toronto. During a wild and extended adolescence, K.C. jammed in rock bands, partied at Mardi Gras, played cards till all hours, and got into fights in bars; he was also knocked unconscious twice, once in a dune buggy accident, once when a bale of hay konked him. Finally, in October 1981, at age thirty, he skidded off an exit ramp on his motorcycle. He

spent a month in intensive care and lost, among other structures, both hippocampi.

After the accident a neuroscientist named Endel Tulving determined that K.C. could remember certain things just fine. But everything he remembered fell within one restricted category: it was all stuff you could look up in reference books, like the difference between stalactites and stalagmites or between spares and strikes in bowling. Tulving called these bare facts "semantic memories," memories devoid of all context and emotion.

At the same time K.C. had zero "episodic memory"—no memories of things he'd personally done or felt or seen. For instance, in 1979 K.C. surprised his family the night before his brother's wedding by getting a perm. To this day he knows his brother got married and can recognize family members in the wedding album (the facts), but he doesn't remember being at the wedding and has no idea how his family reacted to his curly hair (the personal experiences). The little that K.C. did retain about his preaccident life sounds like something he looked up in a particularly dry biography of himself. Even pivotal moments have been reduced to bulleted points in an index. He knows his family had to abandon his childhood home because a train derailed and spilled toxic chemicals nearby; he knows a beloved brother died two years before his own accident. But these events have no emotional import anymore. They're just stuff that happened.

These details, along with scans of K.C.'s brain, provided strong evidence that our episodic and semantic memories rely on different brain circuits. The hippocampus helps record both types of memories initially, and it helps retain them for the medium term. The hippocampus probably also helps us access old *personal* memories in long-term storage. But to access old semantic memories, the brain seems to use the parahippocampus, an extension of the hippocampus on the brain's southernmost surface. K.C., whose parahippocampi survived, could therefore remember to sink the eight ball last in pool (semantic

knowledge), even though every last memory of playing pool with his buddies had disappeared (personal knowledge).*

What's more, while a healthy hippocampus will usually take responsibility for recording new semantic memories, the parahippocampus can—albeit excruciatingly slowly—absorb new facts if it has to. For instance, after years of shelving books as a volunteer at a local library, K.C.'s parahippocampus learned the Dewey decimal system, even though he had no idea why he knew it. Similarly, H.M.'s healthy parahippocampus picked up a few choice facts after his 1953 surgery. After seeing the crossword clue a thousand times he dimly recalled that "Salk vaccine target" equaled P-O-L-I-O. And through incessant references, he retained a sliver of information about the 1969 moon landing and 1963 Kennedy assassination. Contra the cliché, he couldn't recall where he was when he learned those things—that's episodic memory. And his knowledge of the events remained weak and fragmentary, since the parahippocampus cannot learn very well. He nevertheless absorbed that they'd happened.

Along these same lines, K.C. helped neuroscience come to grips with another important distinction in memory research, between recollection and familiarity. Colloquially, recollection means *I specifically remember this,* while familiarity means *this sounds familiar, even if the details are fuzzy.* And sure enough, the brain makes the same distinction. In one test K.C.'s doctors compiled a list of words (El Niño, posse) that entered the common parlance after his accident in 1981. They then sprinkled those words into a list of pseudo-words—strings of letters that looked like plausible words but that meant nothing. Time and again K.C. picked out the real word, and did so with confidence. But when asked to define the word, he shrugged. From a list of common names he picked out the people who'd become famous after 1981 (e.g., Bill Clinton). But he had no inkling what Clinton had done. In other words, K.C. found these terms familiar, even though specific recollection eluded him. This indicates that recollection once

again requires the hippocampus, while a feeling of familiarity requires only certain patches of cortex.

A final type of memory that amnesiacs have helped illuminate is emotional memory—which makes sense, given that the hippocampus belongs to the limbic system. Possibly because he had no amygdalae, H.M. was always pretty affable around the scientists who visited him, despite never recognizing them. (Not even Milner, who worked with him for a half century.) Other amnesiacs lacked his easygoing manner, though, and a few got outright snarly. In 1992 herpes simplex—the same bug that knocked out people's ability to recognize fruits, animals, and tools—hollowed out the hippocampi and other structures inside the brain of a seventy-year-old San Diego man named E.P. He started repeating the same anecdotes over and over, verbatim, and eating up to three breakfasts each day. And despite being a former sailor who lived less than two miles from the coast, he suddenly couldn't remember even the general direction of the Pacific Ocean.

Doctors arranged to test E.P., but he grew suspicious of the "strangers"—really the same woman each time—invading his home. Every visit, he dug in his heels, and every visit, his wife had to talk him into playing nice and drag him to the kitchen table to start testing. Eventually, though, after more than a hundred visits, E.P. let his guard down. He started greeting the tester warmly, despite maintaining that he'd never seen her; he even started moving toward the kitchen table on his own to start testing. Somehow, even though his mind was telling him otherwise, his emotions remembered to trust his tester. Amnesiacs can retain negative emotional memories, too. When H.M. learned that his father had died, his conscious brain of course forgot that fact within minutes. But his emotional brain remembered, and took the news so hard that he plunged into a months-long funk, even though he couldn't explain why he felt so low. In another example, from around 1911, a Swiss doctor named

Édouard Claparède concealed a pin between his fingers before greeting a middle-aged amnesic woman; when they shook hands, he pricked her. Although she remembered nothing of this, she always withdrew her hand, and eyed him, on subsequent meetings.

Taken as a whole, this alphabetic soup of amnesiacs (q.v., e.g., H.M., K.C., E.P.) helped scientists sort out how the brain divides up responsibility* for memories. Nondeclarative memories (like motor memories) rely on the cerebellum and on certain internal clusters of gray matter such as the striatum. Episodic (personal) memories lean heavily on the hippocampus, while semantic (factual) memories utilize the parahippocampus to a much larger degree, especially for retrieval. The frontal lobes contribute as well, both in searching for memories and in double-checking that the brain has grabbed the right memory from long-term storage in the cortex. Sensory and limbic circuits also kick on to reanimate the moment in our minds. Meanwhile, the parietal and frontal lobes whisper to us that we're reviewing old information, so we don't get terrified or amorous all over again. Each step works independently, and each one can malfunction without affecting other mental faculties in the slightest.

That's the theory, at least. In reality it seems impossible to tear out any one aspect of memory—especially our episodic memories, memories of holidays and lovers and times we fell short—without tearing out so much more. K.C. knows how to play solitaire and change a tire, but he can never recall a moment of contentment, peace, loneliness, or lust. And however paradoxical it might seem, losing his past wiped out his future as well. The ultimate biological purpose of memory isn't to recall the past per se, but to prepare for the future by giving us clues about how to act in certain situations. As a result, when K.C. lost his past self, his future self died along with it. He cannot tell you what he'll do over the next hour, the next day, the next year; he cannot even imagine these things. This loss of his future self doesn't pain K.C.; he doesn't suffer or rue his fate. But in some ways that lack

of suffering seems sad in and of itself. However unfair, it's hard not to see him as reduced, diminished.

In our minds, we more or less equate our identities with our memories; our very selves seem the sum total of all we've done and felt and seen. That's why we cling to our memories so hard, even to our detriment, and that's why diseases like Alzheimer's, which rob us of memories, seem so cruel. Indeed, most of us wish that we could cling to our memories more securely—they seem the only bulwark against the erosion of the self that K.C. and H.M. experienced. That's why it's such a shock to realize that the opposite burden—a hoarding, avaricious memory that *cannot* forget—can crush people's identities in the selfsame way.

~

Each morning when Moscow reporter Solomon Shereshevsky got to work, his editor assigned him and the other reporters their daily stories, telling them where to go, what to look for, and whom to interview. Despite the intricacy of the instructions, Shereshevsky never took notes, and according to some accounts he never took notes during interviews, either. He just remembered. Still, Shereshevsky wasn't a great reporter, and at one morning meeting in the mid-1920s his editor's fuse went off when he saw Shereshevsky blithely nodding at him, no pencil in hand. He called Shereshevsky out, challenging him to repeat his instructions. Shereshevsky did, verbatim—and then repeated every other word the editor had said that morning, too. When his fellow reporters stared, Shereshevsky's brow knit in confusion. Didn't everyone have complete recall? Half amazed, half creeped out, the editor sent Shereshevsky to a local neuroscientist, Aleksandr Luria.

Although a young man then, Luria had already started down the path that would make him one of the most celebrated neuroscientists of the twentieth century. He championed the romantic side of

neuroscience, neuroscience that encompassed more than just cells and circuits. He wanted to capture how people actually experienced life, even the messy bits. In doing so, he swam against the current of modern science, which tends to dismiss anecdotal accounts (the plural of anecdote, after all...). But individual case studies have always been crucial to neuroscience: as with the best fiction, it's the particulars of people's lives that unveil the universal truths. Indeed, Luria's book-length case reports have been called "neurological novels," and he wrote one of his finest on Shereshevsky.

In all their years of collaboration, Luria found "no distinct limits" to Shereshevsky's memory.* The man could recite lists of thirty, fifty, seventy random words or numbers, in order, forward or backward, after hearing or reading them just once. All he needed was three seconds in between each item, to fix it in his hippocampus; after that, it was lapidary. Even more impressive, whatever he memorized stuck with him for years. In one test Luria read the opening stanzas of Dante's *Inferno* in Italian, a language Shereshevsky didn't speak. Fifteen years later, with no rehearsals in between, Shereshevsky recited the lines from memory, with all the proper accents and poetic stress. *Nel mezzo del cammin di nostra vita...*

You'd think Shereshevsky would have his pick of six-figure jobs, but like many so-called mnemonists, he drifted somewhat loserishly between careers, spending time as a musician, reporter, efficiency consultant, and vaudeville actor (memorizing lines was a snap). Unfit for anything else, he finally landed a job in what was essentially a neurological freak show, touring the country and regurgitating numbers and nonsense words to audiences. The gap between his obvious talents and his lowly status gnawed at Shereshevsky, but to Luria the discrepancy made sense. That's because Luria traced both his mnemonic prowess and his employment woes to the same root cause—excessive synesthesia.

In Shereshevsky's mind no real boundary existed between the senses. "Every sound he heard," Luria reported, "immediately produced an experience of light and color and... taste and touch." And unlike "normal" synesthetes, whose extra sensations are pretty vanilla (simple odors, single tones), Shereshevsky experienced full-on scenes, full mental stage productions. This became handy when memorizing items. Instead of a violet 2 or chartreuse 6, 2 became "a high-spirited woman," 6 "a man with a swollen foot." The number 87 became a stout woman cozying up to a fellow twirling his mustache. The vividness of each item made recalling it later trivial.

To then remember the *order* of such items, as in a list, Shereshevsky used a trick. He imagined walking along a road in Moscow or in his hometown (whose layout he knew by heart, needless to say) and "depositing" each image at a landmark. Each syllable of the Dante, for instance, summoned up a ballerina or goat or screaming woman, which he'd then plunk down near whatever fence, stone, or tree he happened to be passing at that moment on his mental stroll. To recall the list later, he simply retraced his route, and "picked up" the images he'd left behind. (Professional mnemonists still use this trick today.) The technique backfired only when Shereshevsky, who was rather rigid, did something foolish, like deposit images in dark alleys. In these cases he couldn't make the image out, and he'd skip the corresponding item on the list. To an outsider this seemed like a lapse, a chink in Shereshevsky's memory. Luria realized that this was actually less a failure of memory than of perception—Shereshevsky simply couldn't see the image, nothing more.

Shereshevsky's memory played other tricks as well. He could increase his pulse rate and even make himself sweat simply by remembering a time when he'd chased down a departing train. He could also (and Luria confirmed this with thermometers) raise the temperature of his right hand by remembering a time he'd held it next to a stove, while simultaneously lowering the temperature of his left hand by remembering what ice felt like. (Shereshevsky could even mentally block out pain in the dentist's chair.) Somehow his memory could override the "this is just a recollection, it's not actually happening"

signal from the frontal and parietal lobes that should have quelled these somatic reactions.

Unfortunately, Shereshevsky couldn't always corral his imagination or confine it to turning mnemonic tricks. When reading a book, synesthetic images would start multiplying inside his head, crowding out the text. A few words into a story, he'd be overwhelmed. Conversations took wrong turns, too. He once asked a gal in an ice cream parlor what flavors they had. The (probably innocent) tone in which she responded "Fruit ice cream," he said, caused "whole piles of coals, of black cinders, to come bursting out of her mouth. I couldn't bring myself to buy any." He sounds insane, or like Hunter S. Thompson at his druggiest. If menus were printed sloppily, Shereshevsky's meal seemed contaminated by association. He couldn't eat mayonnaise because a certain sound *(zh)* in the Russian word for it nauseated him. No wonder he struggled to hold a job—simple instructions would mutate inside his imagination and stagger him.

Even the traveling mnemonist gig eventually became oppressive. After too many years of doing the show, Shereshevsky felt old lists of numbers and words haunting him, cacophonizing inside his skull, elbowing newer memories aside. To rid himself of them, he more or less resorted to voodoo, writing out the lists on paper and burning them. (No luck—the exorcism failed.) Relief came only from suppressing such memories, by training his mind to not acknowledge them. Only dumbing his memory down took the edge off.

Most people who met Shereshevsky considered him dim and timid, a bumbling Prufrock. Indeed, he considered himself pathetic, someone who'd wasted his talent in sideshows. But what else could he have done? With so many memories crowded into his skull—his memory actually stretched back to before his first birthday—his mind became what one observer called "a junk heap of impressions." As a result he lived in a veritable haze, nearly as befuddled and helpless as H.M. or K.C. A memory that's too good is just as broken as one that's no good at all.

To be useful, to enrich our lives, memory cannot simply record the world around us. It needs to filter, to discriminate. In fact, while we joke about a poor memory as a sieve, that's actually the wrong way around. Sieves let water leak through, but they catch substantial things—they catch what we want to preserve. In the same way, a mind functions best when we let some things, like traumatic memories, go. All normal brains are sieves, and thank goodness for that.

~

However useful, the sieve metaphor isn't perfect. Human memory doesn't just filter things. Our memories actually sculpt and rework and—with surprising regularity and slyness—distort what remains behind.

Even neuroscientists, who should know better, fall prey to distortions. Otto Loewi, whose dream about frog hearts helped prove the soup theory of neurotransmission, claimed to have had the dream over Easter weekend in 1920. But the journal in which he published his results, according to its records, received his initial submission a week before Easter that year. A few killjoy historians also think that Loewi didn't rush from his bed to the lab at 3 a.m., but instead merely wrote out the details of the experiment, step by step, then resumed snoozing. Perhaps Loewi—who loved telling tales—let the demands of narrative drama mold his memory. Similarly, William Sharpe, who harvested the glands of the giant while the family stewed in the front parlor, couldn't have done so (as he claimed) on New Year's Day, since the giant died in mid-January. Furthermore, a colleague of Sharpe's later claimed to have accompanied him on his clandestine errand—and also claimed that they picked through the giant's innards not right before the funeral but the night before, around 2 a.m. Both men cannot be correct.

Why does this happen? Why do memories get twisted like metal girders in a fire and harden into the wrong shape? Neuroscientists disagree on the answer. But one theory gaining momentum says that the

very act of remembering something—which you'd think would solidify the details—is what allows mistakes to infiltrate.

When capturing a memory, neurons jury-rig a connection for the short term. They then solder those connections together with special proteins, a process called consolidation. But the brain may use those proteins for more than just capturing memories; the proteins may help retrieve and replay memories, too. Consider: If you play a beep, then shock a mouse, it sure as hell remembers this. Play the tone again, and it freezes in terror, anticipating another shock. Scientists have found, however, that they can make the mouse forget that terror. They do so by injecting a drug into the mouse's brain just before the second beep, a drug that suppresses the memory-capturing proteins. Shockingly, the *next* time the tone plays, the mouse keeps on doing mousey things. Without those proteins the memory apparently unravels, and the mouse never fears the beep again. This implies that our brains, when recalling a memory, probably don't just replay a pristine "master copy" each time. Instead, they might have to re-create and re-record the memory each time through. And if that recording gets disrupted, as it did in the mouse, the memory vanishes. This theory, called *re*consolidation, argues that there's little inherent difference between recording first mnemonic impressions and recalling them later.

Now, mice aren't little humans: humans have richer, fuller memories, and our memories work differently. But not that differently, especially on a molecular level. And if reconsolidation happens in humans—and there's evidence it does—then having to rerecord a memory each time through probably makes it labile and therefore corruptible. To be sure, we humans don't often forget events completely, like the mice did. But we do garble details,* especially personal details, all the time. As a troubling corollary, the memories that most define us—our tenderest moments, our traumas—could be most prone to distortion, since we reminisce about them most often.

So why do distortions creep in at all? Because we're human. Subsequent knowledge can always taint a memory: you can never remember your first date quite as fondly if that son of a bitch cheated on you later. So you retroactively retouch things and convince yourself that he mistreated you from the start. We also don't store memories the way computer hardware does, with each datum in a well-defined location. Human memories live in overlapping neuron circuits that can bleed together over time. (Some observers have compared this to Wikipedia editing, with each neuron able to tweak the master copy.) Perhaps most important, we feel the need to save face or goose our reputations, either by gliding over inconvenient facts or misrepresenting them. Indeed, some scientists argue that the unconscious mind confabulates—makes up plausible stories to mask our true motivations—far more often than we care to admit. Unlike victims of Korsakoff's syndrome, normal folk don't confabulate because of memory gaps. But we do tint what we recall and suppress what's convenient to suppress—until we "remember" what we want to, and can believe that a life-changing dream really did occur on Easter. Memories are memoirs, not autobiographies. And the memories we cherish most may make honest liars of us all.

Chapter Five

Learning, Memory, and Human Individuality

WHAT'S A BRAIN good for? We've seen that the lower portions of our brains have essential control circuits that govern basic body functions: key reflexes, an automatic thermostat, a regulated appetite for food and drink, and wakefulness/sleepiness. The lower brain also has regions for coordinating our movements and modifying our perception to direct our attention to the outside world. This is the basic stuff that we share with frogs and fish, the "bottom scoop of the ice cream cone." The top two scoops, the limbic system and the neocortex, are where things get really interesting. Many complex functions such as language and social reasoning emerge in the cortex, but I contend that there are two key brain functions that are the basis upon which these higher capacities are built. These are memory and emotion—and the interaction between the two.

Consider this analogy: the brain does for the individual what the genome does for the species. The genome, the sequence of information encoded in the DNA, undergoes random mutation and sometimes a mutation (or a collection of mutations) confers an advantage on an individual that allows him or her to have more and/or healthier offspring. The genome, through the Darwinian process of natural selection, is the book in which the story of evolution is written: the experience of the species ultimately modifies the genome and thereby the genetic traits of the species, sometimes rendering it better adapted to the environment. The limitation of evolution through natural selection is that it is not a rapid process. Species adapt to their experiences (environments) slowly, over many generations.

The brain, by storing memories, performs a related function for the individual. It is the book in which individual experience is written. Because memory storage is rapid, it allows an individual to adapt to new experiences and situations. This is a much more flexible and powerful solution than relying solely upon mutation and selection acting upon the genome.

But how does emotion come into it? In our lives, we have a lot of experiences and many of these we will remember until we die. We have many mechanisms for determining which experiences are stored (where were you on 9/11?) and which are discarded (what did you have for dinner exactly 1 month ago?). Some memories will fade with time and some will be distorted by generalization (can you distinctly remember your seventeenth haircut?) We need a signal to say, "This is an important memory. Write this down and underline it." That signal is emotion. When you have feelings of fear or joy or love or anger or sadness, these mark your experiences as being particularly meaningful. These are the memories you most need to store and keep safe. These are the ones that are most likely to be relevant in future situations. These are the building blocks that form logic, reasoning, social cognition, and decision making. These are the

memories that confer your individuality. And that function, memory indexed by emotion, more than anything else, is what a brain is good for.

IN CASUAL CONVERSATION we may say that a certain person has a good memory or another person has a bad memory. In truth, however, we know from our everyday experience that things are not so simple. Memory is not a unitary phenomenon. You may have a great ability for matching names to faces while you struggle to memorize music for a piano recital. Your brother might remember everything he ever reads but progresses slowly with motor memory tasks such as learning to improve his golf swing.

Brain researchers have worked for many years to develop a taxonomy of memory, a means of classifying types of memory that has its roots in clinical observation (see Figure 5.1). Much of this work relies on the analysis of human amnesiacs who have sustained damage to various parts of their brains through infections, stroke, trauma, chronic abuse of drugs or alcohol, or, as in the case of the patient called H.M. (Chapter 1), surgery to treat otherwise incurable seizures. Other insights have come from studying more temporary forms of disruption, such as transiently acting drugs and electroconvulsive shock (used to relieve depression that fails to respond to other therapies).

In the 1950s it was generally thought that patients like H.M. and others who sustained damage to the hippocampus and surrounding cortical tissue were unable to form any new memories at all. But detailed study of these patients revealed that although they could no longer form new memories of facts or events, so-called declarative memories, they could lay down memory traces for a number of other tasks. One of these is mirror reading: learning to read words in English that have been printed with left-right reversal (Figure 5.2). This is a task that both normals and hippocampal amnesiacs such as H.M. can learn with daily practice. It's also a nice task for illustrating different types of mem-

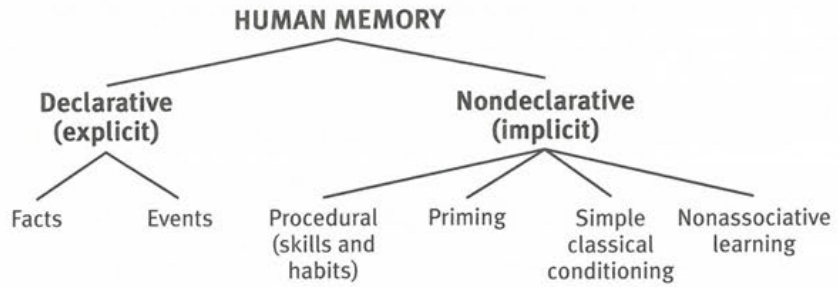


FIGURE 5.1. A taxonomy of human memory. Adapted with permission from Elsevier from B. Milner, L. R. Squire, and E. R. Kandel, Cognitive neuroscience and the study of memory, *Neuron* 20:445–468 (1998). *Joan M. K. Tycko, illustrator.*

ory: although both the amnesiacs and the normals showed daily improvement in the mirror-reading task (as indicated by progressively faster reading times), only the normals could recall some of the words used in the test the previous day—the amnesiacs had no memory of these words whatsoever (indeed, they also had no memory that the previous day's training session had even occurred).

Further experiments with hippocampal amnesiacs have revealed a large group of memory tasks that are retained. Amnesiacs still have memory for motor coordination—they can improve at sports with practice. Both mirror reading and motor coordination learning fall into the larger category of "skills and habits" shown in Figure 5.1. Amnesiacs also retain the ability to learn simple, subconscious associations, through a process called classical conditioning. For example, your heart rate will reflexively accelerate if you receive a mild shock to your arm, but it will not do so in response to a more neutral stimulus such as the sight of a dim red light briefly appearing in your field of view. But if the light is paired with the shock repeatedly, after a while your brain will begin to learn that the light predicts the shock and your heart rate will accelerate in response to the

FIGURE 5.2. Mirror reading, a skill that both hippocampal amnesiacs and normals can acquire and retain with practice. The memories of the particular words read, however, will be retained only by the normals.

light alone. Hippocampal amnesiacs trained in this task for several days will have no memory of the previous day's training, but their heart rate will accelerate in response to the light nonetheless.

Perhaps the most interesting form of memory retained in amnesiacs is achieved by a method called priming. In this task, initially devised by Elizabeth Warrington and Larry Weiskrantz of Oxford University, amnesiacs are asked to recall a list of words they have seen the previous day. Not surprisingly, if you simply ask them to list the words from the earlier session they have no memory of them at all. But if you give them the first few letters of a word, they will often be able to correctly produce the complete word even if it feels to them as if they

are guessing randomly. For example, if a word on the list was "crust," the stem "cru____" would probably evoke the correct answer rather than other possibilities such as "crumb" or "crud" or "cruller." What's interesting about priming is that, unlike many of the other forms of memory retained in hippocampal amnesia, it is a cognitive rather than a motor task.

All of these memory tasks that are retained in amnesiacs (priming, skill and habit learning, classical conditioning, as well as some others I didn't discuss) fall into a category called nondeclarative, or implicit, memory: they are forms of memory that do not involve conscious retrieval. These memories are not recalled, but rather are manifest as a specific change in behavior. Nondeclarative memory is *not* what we usually think of when we talk about memory in casual conversation—it is not memory for facts and events—such as what you had for breakfast yesterday morning or the name of the British prime minister. Nonetheless, nondeclarative memory is central to our experience.

OBSERVATIONS OF HUMAN amnesiacs clearly suggest that storage of new declarative memories requires an intact hippocampal system (the hippocampus proper and some adjacent cortical structures). This brings up a central question: are particular memories stored in specific locations in the brain or are they stored in a distributed fashion, spread over many brain regions? One early indication of the answer came from the work of the Montreal Neurological Institute neurosurgeon Wilder Penfield, who, starting in the 1930s, stimulated the brains of patients undergoing surgery for epilepsy.

This was not just an academic exercise. It allowed him to map more carefully than previously the location of the area that triggered the seizure and thereby minimize damage to nearby parts of the brain. Because brain tissue has no pain-sensing system itself, neurosurgery can be performed on conscious people while they are under a local anesthetic to block pain from their scalp and

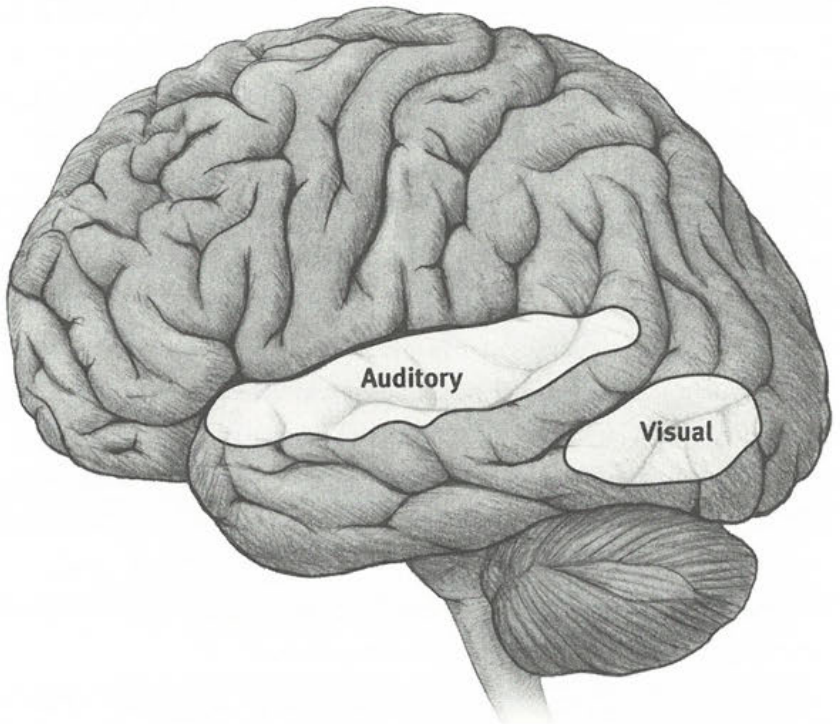


FIGURE 5.3. Reminiscence evoked by brain stimulation during neurosurgery. The Canadian neurosurgeon Wilder Penfield inserted electrodes to stimulate the cortical surface of awake patients in the course of neurosurgery. This figure shows two types of memory-like experiences evoked by stimulation in various regions. *Joan M. K. Tycko, illustrator.*

skull. Penfield's stimulation was restricted to the cortical surface and was performed on over 1,000 surgical patients (Figure 5.3). In a small fraction of cases, stimulation of the cortical surface would evoke a coherent perception: a snatch of music, a human voice, a vision of a pet or loved one. Were these electrical

stimuli evoking memories? Well, yes and no. In some cases it does seem that particular, real past events were recalled, at least in fragmentary form. More often, however, the stimulation evoked sensations that were dreamlike, with typical elements of fantasy and violations of physical laws. Often, the area that evoked a "memory" was itself the epileptic focus. In these cases, destruction of that cortical tissue did not obliterate the memory of that particular stored experience. So, the Penfield experiments, while titillating, did not directly address the question of memory localization.

If formation of new nondeclarative memories can proceed when the hippocampal system is destroyed, then where are the critical locations for these forms of memory? Some information about this can be derived from studies of humans with damage to other brain regions. Damage to the amygdala, for example, seems to be associated with memory storage for classical conditioning of emotional responses, particularly fear conditioning. Damage to the cerebellum has similar effects for classical conditioning of emotionally neutral stimuli (see Figure 5.4).

So, is memory storage localized or not? The answer is not so simple. It's also a bit different for nondeclarative versus declarative memory. Nondeclarative memory is not consciously recollected. Rather, it is evoked by a specific stimulus or set of stimuli and is manifested as a change in behavior. As a result, nondeclarative memories can often be localized, not just to a brain region, but to a certain subregion or even class of neuron. Declarative memory is a different story. Such memories are consciously recollected. They are useful in large part because we can access them using stimuli that are very different from the ones that created them initially. For example, when you read "Imagine your mother's face," the sensation of reading that line is nothing like the stimuli that laid down the memory of your mother's face, yet you can probably re-

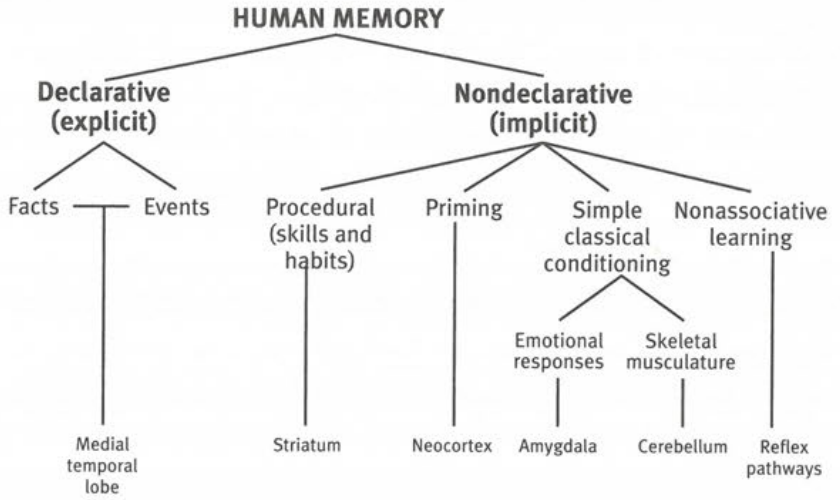


FIGURE 5.4. A taxonomy of human memory, now elaborated to show some crucial brain regions involved in different tasks. Here, the medial temporal lobe means the hippocampus and some associated regions of the cerebral cortex. It should be cautioned that, in real-world behavior, most experiences will simultaneously lay down memory traces of several types. For example, if you take lessons to improve your tennis game, you will probably recall particular things that may have happened during the lesson (declarative memory for events), but what you are really trying to achieve is an unconscious improvement of your motor performance as you play (nondeclarative skill memory). Adapted with the permission of Elsevier from B. Milner, L. R. Squire, and E. R. Kandel, Cognitive neuroscience and the study of memory, *Neuron* 20:445–468 (1998). *Joan M. K. Tycko, illustrator.*

call her face with ease after reading that line. This imposes an important constraint on declarative memories: nondeclarative memories can simply be accessed in a subconscious fashion through specific stimuli, but declarative memories must be embedded in a much richer informational system, which

makes it less likely that they will be localized to the same degree as nondeclarative memories.

ALTHOUGH DAMAGE TO the hippocampal system produces anterograde amnesia, an impairment in the storage of new memories for facts and events, it does not erase a lifetime's worth of declarative memories. Rather, there is typically a "hole in declarative memory," or retrograde amnesia, which stretches back 1 or 2 years before the infliction of hippocampal damage. Thus H.M. and others like him have lost a part of their past forever, but older memories have been spared. The explanation for this seems to be that declarative memories are initially stored in the hippocampus and some adjacent regions, but, gradually, over months to years, the storage site changes to other locations in the cerebral cortex. The dominant theory today is that the final locations for declarative memories are distributed in the cortex, not in a random fashion, but rather in those parts of the cortex initially involved in their perception. In this fashion, memories for sounds are stored in the auditory cortex (indeed, memories for words appear to be stored in a particular subregion of the auditory cortex), memories for scenes in the visual cortex, and so on. What this means for any real experience, involving multiple senses, is that your memory for, say, your first trip to the beach is stored in a number of cortical locations, each corresponding to a particular sensory modality or submodality. There does not appear to be a single dedicated site for the permanent storage of declarative memories. This underlies, at least in part, the observation that memory is not a unitary system. This may be why your Aunt Matilda can remember every word to every song Elvis Presley recorded, but can't keep track of your birthday.

MEMORIES MAY BE classified not only by type but also by duration. There is evidence for separate neural processes underlying at least three stages of memory. The first and most transient of these is known as working memory. Anyone who grew up with an annoying sibling knows certain aspects of working memory well: You've just read a phone number from your address book and you're repeating it to yourself, trying to keep it in your memory long enough to dial the phone while your sib is trying equally hard to interfere by shouting random numbers in your ear. Working memory is a temporary "scratchpad" for holding information just long enough to complete a task (to dial a phone number or to remember the first part of a heard sentence long enough to match it with the ending). You can hold information in this scratchpad for a somewhat longer time through rehearsal or by employing mental imagery, but otherwise it will quickly fade away. Working memory is a form of declarative memory that is crucial to understanding lengthy experiences as they unfold over time. It is the glue that holds our perceptual and cognitive lives together.

Working memory is preserved in hippocampal amnesiacs. Although we don't have a complete understanding of its neural basis, there is now a generally accepted model that holds that working memory requires the ongoing firing of particular sets of neurons. This has been tested in monkeys by using a working-memory task called delayed matching to sample (Figure 5.5). In this task, a colored light briefly flashes, and then after a delay of a few seconds the monkey must correctly choose the previous color from a display of two or more to get a food reward. Investigators found that some neurons in the higher regions of the visual "what" pathway (in an area called TE) fired continually during the working-memory interval. These higher visual areas send a lot of axons to the prefrontal cortex, and neurons in this area also fire in this fashion. Similar activity can also be recorded with scalp electrodes in the prefrontal cortex of human subjects performing working-memory tasks. So, a current model is that there are separate working-memory systems for different areas in the brain, each located at some point in the appropriate region of cortex (auditory, visual,

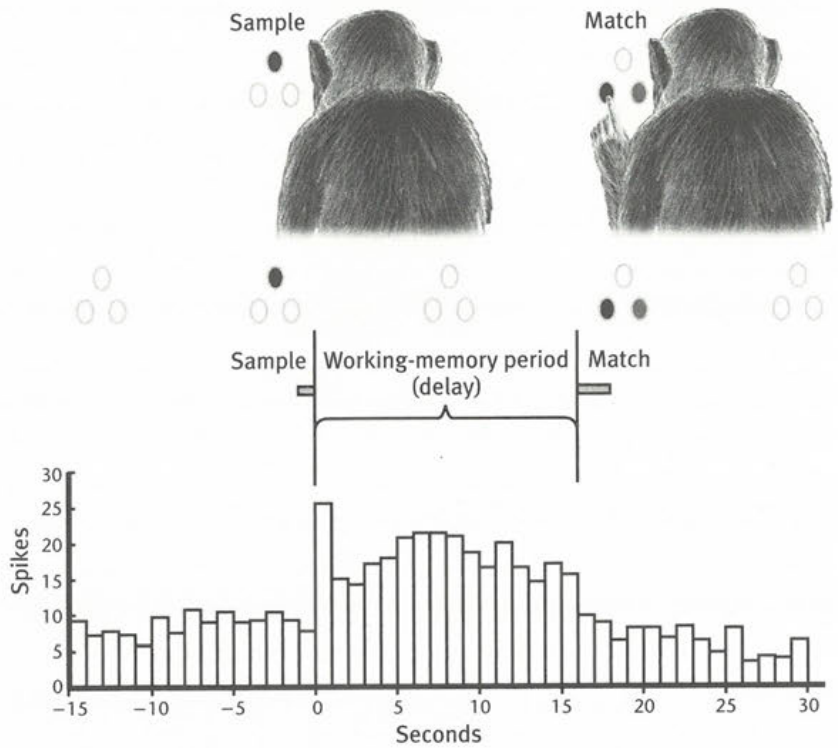


FIGURE 5.5. Persistent neuronal activity as a substrate of working memory. In this delayed matching to sample task, a monkey sees a given color and then, after a delay of a few seconds, must pick this color from two choices. The lower panel shows a recording made from a neuron in the higher visual region called area TE, illustrating that the firing rate of this neuron was elevated when the sample was presented and remained so through the 15-second-long working-memory interval. Adapted from L. R. Squire and E. R. Kandel, *Memory: From Mind to Molecules* (Scientific American Library, New York, 1999); © 1999 by Scientific American Library; used by permission of Henry Holt and Company, LLC. *Joan M. K. Tycko, illustrator.*

and so on). These regions all seem to project to the prefrontal cortex, which, at least to some degree, integrates working memory across sensory modalities. This model is further supported by the findings that damage to the prefrontal cortex in both humans and monkeys results in impairment on working-memory tasks.

More subtly, if the prefrontal cortex of monkeys is artificially electrically stimulated during the working-memory interval, performance can be disrupted. A similar effect can be produced by injecting drugs into the prefrontal cortex that either block or overactivate receptors for the modulatory neurotransmitter dopamine. Dopamine functions to tune the amount of spike firing in the prefrontal cortex that is triggered by information flowing from other cortical regions such as the auditory and visual systems. This may explain why schizophrenics and patients with Parkinson's disease, people whose ailments are associated with defects in dopamine signaling, perform poorly on tests of working memory.

IF YOU QUERY middle-aged people on general knowledge (news, popular culture), you typically find that they have better recollection of recent events than more distant past ones. This predictable result is called the forgetting curve. Yet distant memories that do survive normal forgetting are unusually resistant to disruption. As the record of a particular experience moves from working memory through short-term memory and into long-term memory, the memory trace, or engram (those changes in the brain that encode memory), gradually changes from being fragile and easily disrupted to being more stable. This transformation process takes time and has been given the name consolidation. The evidence for this comes from both human and animal studies. If you repeat the experiments mentioned above in which you query general cultural knowledge among people who have received bilateral electroconvulsive shock treat-

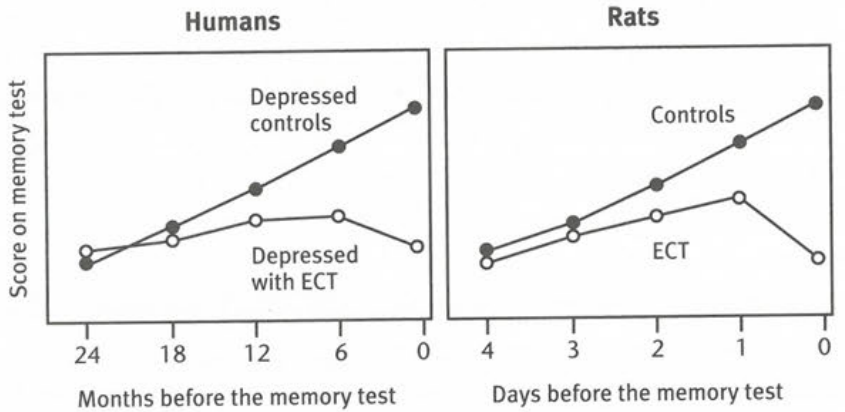


FIGURE 5.6. The persistence of old memories and the fragility of new memories. Both humans and rats were tested for their recollection of past events. Controls showed some degree of forgetting of information in the more distant past. Both humans and rats receiving ECT had severely impaired memory for events occurring immediately before ECT, but had normal rates of forgetting for information in the more distant past. Note the different time scales for the human and rat data. *Joan M. K. Tycko, Illustrator.*

ment (ECT) to relieve drug-resistant depression, you will find that superimposed upon the forgetting curve is an additional disruption of memory that is strongest for events that occurred immediately before the treatments and that gradually trails off as you query events further in the past (Figure 5.6). Of course, in this type of experiment, it is important that the control group be others with severe depression and not just the general population.

Similar studies can also be done using laboratory animals such as rats. Obviously, in this case, you can't quiz them about general knowledge, so instead you train them in a particular task (such as navigating a maze to get food) and then wait for various intervals before giving ECT. When you test them for their memory of the maze task the next day, you will find a result that parallels that

seen in humans: recent memory is easily disrupted by ECT, but memory in the more distant past is sufficiently consolidated to withstand this treatment (Figure 5.6). This basic strategy can be applied to many types of experimental amnesia, including those caused by the administration of various drugs. One class of drug that has been particularly well studied in terms of disrupting memory consolidation is protein synthesis inhibitors. These are compounds that interfere with any one of several biochemical steps by which genes ultimately direct the synthesis of new proteins. Thus a popular hypothesis is that synthesis of new proteins is one important step in the consolidation of short-term into long-term memory, thereby rendering the memory less vulnerable to erasure. Although these examples use tests of declarative memory, several lines of evidence indicate that nondeclarative memories also undergo consolidation and that this consolidation requires new protein synthesis.

ON THE EVENING of October 6, 1996, I was watching TV. I know this, because so were about 46 million others in the United States: It was the night of the first Presidential debate between incumbent Bill Clinton and challenger Bob Dole. When questions were posed to Clinton, he had a habit of pausing for about 3 seconds with his eyes rolled back in his head and then launching into a carefully constructed and detailed answer. However you might have felt about his policies, you had to admire his command of information. After several questions like this, with the characteristic 3-second pause, my wife said, "Look, he's rewinding the tape!" We laughed because it really did seem as if something mechanical was happening in the President's brain that night.

Although we might imagine that our memories for facts and events are stored on a tape we can rewind or a set of photographs we can browse in an album, this does not seem to be true. As discussed previously, one of the biggest challenges of declarative memory (memory for facts and events) is to store information so

that it can be retrieved by diverse stimuli. The key point here is that retrieval of memory is an active process. It is not like browsing through an album of photographs, even an album of fading photographs. Rather, it is a bit like searching the Internet with Google. A question such as "Who was with us on that day trip to the beach last summer?" provides a few search terms that will yield a large number of memory fragments associated with key terms such as "beach" and "last summer." But the question "Who was with us on that day trip to the beach last summer—you know—when we got caught in the thunderstorm and then you threw up in the car on the way home?"—with its greater number of search terms—not only makes it more likely that you will recall the memory of those events, but also makes it more likely that you will recall more aspects of the events. Of course, unlike a Google search, retrieval of declarative memory is not fundamentally text based.

Memory retrieval is an active and dynamic process. But this dynamic recollection and rewriting of memory is a two-edged sword. In some ways it's very useful for subsequent experience and recollection to modify memory traces of events in the past, but this can also lead to errors. Memories of recurring commonplace events are often rendered generic. This is something we all know from our own personal experience. As a child growing up in Santa Monica, California, I probably ate dinner with my father at Zucky's Delicatessen hundreds of times. Although I have many memory fragments associated with those times—the smell of matzo-ball soup, my father's secret-agent like insistence upon sitting with a view of the door, the weird mechanical sound of the cigarette machine, the unnatural colors of the glossy marzipan fruits in the bakery case—these are mostly not related to specific incidents. I can't really say that I remember a particular meal that I ate any night when I was 12 years old. I can, however, remember almost everything about the particular night at Zucky's in 1974 when my father told me that he would be having triple bypass heart surgery. It scared the hell out of me and, as a consequence, that particular meal is etched into my memory forever.

Everyone knows how emotion-laden events can be written into long-term memory with unusual strength. One would be tempted to think that this could be entirely explained by activation of emotional systems at the time of the event. Indeed, that's part of the story, but it's not the whole story. It is now clear that consolidation of long-term memory is also reinforced by subsequent conversation—when you repeatedly tell the story of where you were on 9/11, this repetitive narration reinforces consolidation. Furthermore, the emotions in both you and your listeners that are evoked by the retelling will subtly influence the memory trace itself—the event and the retelling will begin to blend in your mind.

This dynamic reconsolidation of memory is all well and good in some ways: memory of commonplace events is probably of more use to us when rendered generic by the passage of time and subsequent experience. This also has the useful result that emotionally important events stand out more clearly in our memories. But this dynamic process renders our memories particularly subject to certain forms of error above and beyond the slow, gradual fading of long-term memory over time. In his splendid book *The Seven Sins of Memory,* the Harvard University psychologist Daniel Schacter speaks of three of these "sins of commission" in declarative memory retrieval: misattribution, suggestibility, and bias.

Misattribution is very common form of error in which some aspects of a memory are correct but others are not. It can happen in many domains. Take, for example, source misattribution: I may correctly recall a joke I heard which begins "Ted Kennedy walks into a bar . . ." but I will swear I heard it from my sister-in-law when I really heard Jay Leno tell it on TV. Sometimes misattribution can cause you to think that you've created something original, whereas re-

ally you had heard it from another source and attributed it to your own internal processes. I went around for over 30 years humming a snippet of tune I thought I had composed myself, only to hear it years later when I bought my children a DVD containing Bugs Bunny cartoons from the 1940s.

Misattribution is at the heart of one of the most famous cases in music copyright law, which involved the 1970 number-one pop hit by George Harrison called "My Sweet Lord." Although the lyrics and instrumentation differ, the tune of "My Sweet Lord" very strongly resembles that of a previous number-one hit recorded by the Chiffons in 1963 called "He's So Fine." The judge in the case ruled that though Harrison had no intent to plagiarize, he had almost certainly misattributed his memory of the tune of "He's So Fine" (which Harrison admitted he had heard before), thereby imagining that he had composed it de novo. The company which held the copyright to "He's So Fine" was ultimately awarded millions of dollars in damages from Harrison.

These examples are forms of source misattribution. A variant of this is misattribution of time or place. A common experimental design is to give subjects a list of words to study. When they return the next day they are given a new list of words and are asked to indicate which ones they had seen the day before. People in these experiments will often misattribute new words to the previous list. Their propensity for doing this can be manipulated by experimental context. For example, if a word appearing on the new list for the first time is more familiar to the subject or is thematically related to several words on the first list, it is more likely to be misattributed. If the first list contained "needle," "sewing," "pins," and "stitch," then the chance of misattributing the word "thread" to the first list will be high. It may be that we have an evaluative system in our brains that says "If I recognize this word rapidly then it's likely that I've seen it before," and this is the basis of some forms of misattribution.

Suggestibility and bias are additional forms of memory error in which the act

of recollection involves the incorporation of misleading information. Suggestibility is the term used when this information comes from external sources (other people, films, books, media), while bias is warping one's recollection to fit present circumstances: "I always knew that the Red Sox could win the World Series." It turns out to be surprisingly easy to alter people's recollections. For example, a number of studies have sought to simulate police line-ups: a group of experimental subjects will watch a video of a (simulated) convenience store robbery and will then see a line-up of six suspects, none of whom was the robber in the video. When subjects are presented with the suspects one by one and asked to make a yes-or-no decision, almost all will correctly respond "no" to all six. But if the six are presented all at once and the subjects are asked, "Is any of these the robber?" then about 40 percent of people will pick a suspect (usually the one who resembles the perpetrator most closely). If the subject is told by the experimenters in advance that several others have already identified suspect X and they need them to confirm or deny, then about 70 percent of people can be manipulated into false recollection. These results not only highlight the suggestibility of memory recall but also have obvious implications for police procedures and our legal system.

The problem of suggestibility is even greater in children, particularly preschool-age children. In a typical study, a group of preschoolers had a bald man visit their room, read a story, play briefly, and leave. The next day, these children were asked nonleading questions such as "What happened when the visitor came?" and they related a series of memories that, though not complete, were quite accurate. But when leading questions were used, such as "What color hair did the man have?" then a large number of children made up a color. Even those children who initially responded that the man had no hair would typically, after having the question repeated several times in different sessions, begin to confabulate and even extend the false recollection—"He had red hair.

And a mustache too!" Initially, such studies were done using rather innocuous questions like the one above. The consensus at that time was that though children were suggestible about trivial details, they could not easily be made to confabulate entire events, particularly events that would be emotionally traumatic.

A series of high-profile accusations of child abuse in the 1980s prompted several teams of researchers to reexamine this point. What they found was startling. Both preschoolers and, to a lesser extent, elementary school children could easily be made to completely manufacture allegations of abusive behavior (such as yelling, hitting, or taking off their clothes) against an adult in a laboratory setting. All it took was some social incentives: leading questions, reinforcement of particular answers, and a lot of repetition. These are exactly the techniques that were used by many therapists and police officers in developing evidence to accuse preschool teachers in the 1980s. Most (but not all) of these cases were ultimately dropped or overturned on appeal. Let's be clear about what this means: Abuse of children happens and spontaneous reports of abuse volunteered by children are often true and warrant careful examination. But extreme care must be taken in questioning children in cases where abuse is suspected. It is extraordinarily easy for caring professionals with the best intentions to distort a child's recollection or even implant memories that are completely false. The neural basis for the increased suggestibility of small children is unknown but is likely to reflect the fact that brain regions required for retaining memory of events and evaluating confidence in the accuracy of one's own recollections, particularly the frontal lobes, are still undergoing rapid growth and reorganization in the preschool years and slower growth from age 5 to age 20.

WHAT CHANGES OCCUR in brain tissue to store long-term memory? Let's begin our consideration of this key question by stepping back a bit and playing engineer. In building neural memory storage there are a lot of difficult design goals

we'll have to meet. First, the capacity for memory storage must be large. Even though we forget things, we still have to store a huge amount of information over many years and do so with reasonable fidelity. Second, memory must be durable. Some memories will last for an entire lifetime. Third, memories must be stored in a way such that they are retrieved readily, but not too readily. For declarative memories, this means that they must be recollected by using fragmentary cues that can be very different from those which laid them down ("Imagine your mother's face"). Nondeclarative memories are optimally triggered by an appropriate range of stimuli—if you've been trained to blink to a 400 hertz tone, then you probably also would want to blink to a 410 hertz tone but not a 10,000 hertz tone. Fourth, memories must be malleable, based upon subsequent experience in order to place them in a useful context and absorb them into the totality of the conscious self. All in all, this is a rather tall order. Memory must be accurate, but it must also be useful in supporting generalization. It must be permanent, but also subject to modification by subsequent experience. Given these competing requirements, it is not surprising that our memories for facts and events are often subject to misattribution, suggestibility, and bias.

On a smaller scale, what we need to build are systems through which particular patterns of experience-driven neuronal activity will create enduring changes in the brain. What are the general classes of change that could be used to store memories? We know that the fundamental unit of neuronal information is the spike. The probability of spike firing is driven by the integrated activity of many of the excitatory and inhibitory synapses, which add together to produce changes in the voltage across the membrane at the axon hillock, where the spike originates. So if a particular pattern of neuronal activity results in a lasting modification of, say, voltage-sensitive sodium channels located at the axon hillock, such that the threshold for firing a spike was moved closer to the resting

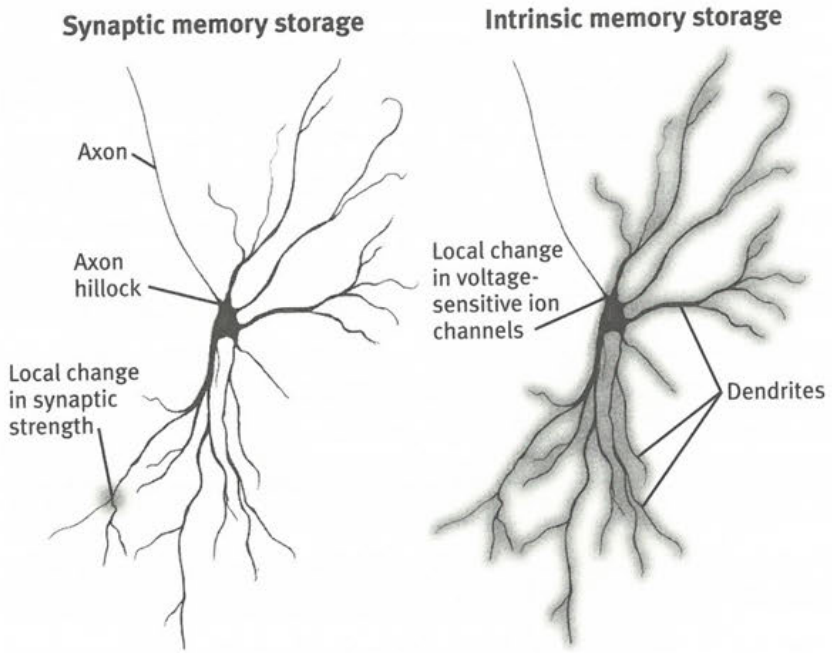


FIGURE 5.7. Synaptic versus intrinsic modulation in memory storage. Long-term modulation of synaptic strength (left) results in changes in throughput that are confined to the activated synapses (shaded area). Changes in intrinsic excitability through modification of voltage-sensitive channels in the axon hillock (right) will change throughput from synapses received throughout the dendritic arbor (shaded area). As a consequence, intrinsic changes have the advantage of producing useful generalization but the disadvantage of having a much smaller capacity to store memory. Adapted from W. Zhang and D. J. Linden, The other side of the engram: experience-driven changes in neuronal intrinsic excitability, *Nature Reviews Neuroscience* 4:885–900 (2003). *Joan M. K. Tycko, illustrator.*

potential, then this could produce a lasting change in the firing properties of that neuron, thereby contributing to an engram. This is only one of many possible changes that would affect neuronal spiking. For example, modifying the voltage-sensitive potassium channels that underlie the downstroke of the spike could change their average time to open. This would result in alterations to the rate and number of spikes fired in response to synaptic drive. Indeed, changes in voltage-sensitive ion channels can persistently alter the intrinsic excitability of neurons and, in animal experiments, these changes can be triggered by learning.

Although changes in intrinsic excitability are likely to contribute to some aspects of memory storage, it's unlikely that they are the whole story. Computationally, this mode of memory storage doesn't make the most efficient use of the brain's resources. Recall that there are about 5,000 synapses received by the average neuron. When you change ion channels underlying spike firing, you are changing the probability of firing a spike in response to synaptic input for all 5,000 of those synapses at the same time. One can imagine that this generalizing property might be useful for certain aspects of memory storage, but an engram solely built upon modifying neuronal intrinsic excitability would, by its nature, have a much smaller capacity than one that allowed individual synapses to change.

Experience-dependent modification of synaptic function is a general mechanism that is thought by most brain researchers to underlie a large part of memory storage. There are many steps in synaptic transmission and several of these are subject to long-term modulation. As a sort of shorthand, people speak of "synaptic strength" as a parameter that can be changed. If, as a test, you stimulate 10 excitatory axons to fire spikes and they all converge on the same postsynaptic neuron and you then measure the resultant deflection in membrane voltage (the EPSP), you might find that this produces a depolarization of 5 mil-

livolts. If, after a certain period of conditioning stimulation (a particular pattern of activation designed to mimic the results of sensory experience), this same test stimulation produced a depolarization of only 3 millivolts, this would be called synaptic depression. An increase in the response to 10 millivolts would be called synaptic potentiation. If these changes were long-lasting in nature, they could contribute to the storage of memory. Because there are about 500 trillion synapses in your brain, this mechanism, experience-driven persistent changes in synaptic strength, has a very high capacity for information storage.

There are two general ways to modify the strength of existing synapses. On the presynaptic side, you could potentiate or depress the amount (or probability) of neurotransmitter release following arrival of an action potential. Or, on the postsynaptic side, you could potentiate or depress the electrical effect produced by a constant amount of released neurotransmitter. In molecular terms, each of these forms of modification can come about in several different ways. For example, if one modifies voltage-sensitive calcium channels in the presynaptic terminal so that they pass fewer calcium ions into the cell when an action potential invades, this will depress neurotransmitter release. A similar effect may be produced by modifying the proteins that control the fusion of neurotransmitter-laden synaptic vesicles with the presynaptic membrane. In this case, for a constant spike-evoked presynaptic calcium signal the probability of a vesicle being released would become lower. On the postsynaptic side, you can depress the effect of released transmitter by reducing the number of neurotransmitter receptors in the postsynaptic membrane. Alternatively, a similar result could be achieved by keeping the number of receptors constant but modifying them so that they pass fewer positively charged ions when they open. The point here is that almost every function on both sides of the synapse is subject to modulation and is therefore a candidate for a memory mechanism. In practice, these different molecular mechanisms are not mutually exclusive, and in most synapses several can be working at the same time.

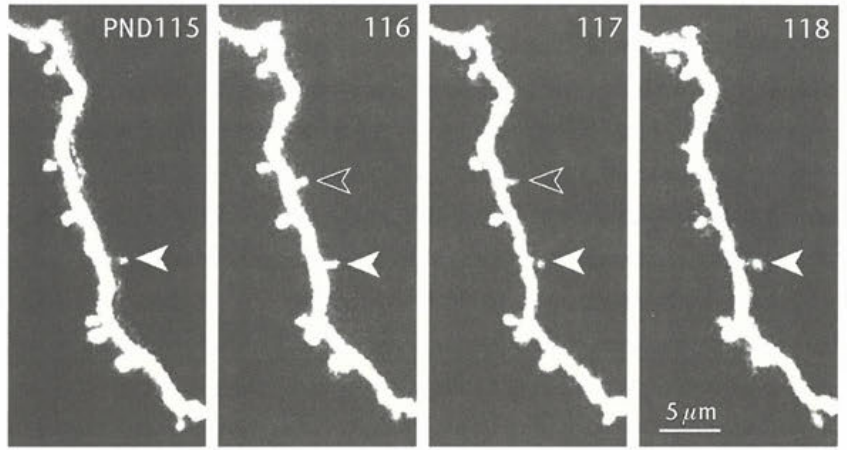


FIGURE 5.8. Changes and stability in the fine structure of dendrites in adult cerebral cortex. These images of a segment of neuronal dendrite from living mouse visual cortex were taken every day from day 115 to day 118 (PND = postnatal day). The filled arrowhead shows one of several stable dendritic spines, while the open arrowhead shows a transient one. This mouse was genetically engineered to express a fluorescent protein in some of its cortical neurons. Reproduced with the permission of Elsevier from A. J. Holtmaat, J. T. Trachtenberg, L. Wilbrecht, G. M. Shepherd, X. Zhang, G. W. Knott, and K. Svoboda, Transient and persistent dendritic spines in the neocortex in vivo, *Neuron* 45:275–291 (2005).

Modifying synaptic function is not the only way to create long-term memories. Such memories may also be encoded through changes in synaptic structure. Although the overall wiring plan of the brain is largely fixed in adult brains, the same cannot be said of individual axons, dendrites, and synapses. Short-term memory is likely to involve changes in the function and structure

of existing synapses, but long-term memory can involve the creation of new branches of dendrites and axons. The tiny spines that cover dendrites are structures that are particularly subject to experience-dependent rearrangement. One recent study by Karel Svoboda and his coworkers at Cold Spring Harbor Laboratory used a novel form of microscopy to repeatedly examine dendritic structure in the cerebral cortex of living adult mice (Figure 5.8). They found that over a period of 30 days, approximately 25 percent of dendritic spines disappeared or were newly formed. At a microscopic level, the synapses of the brain are not static. They grow, shrink, morph, die off, and are newly born, and this structural dynamism is likely to be central to memory storage.

I'VE NOW PRESENTED a theoretical overview of some cellular mechanisms by which memory could be stored in the brain. How do we then go about testing whether any of these mechanisms really operate in behaving animals? There are two general approaches. One is to alter the brain function (with drugs, lesions, genetic manipulation, electrical stimulation, and so on) and observe the resultant effects on behavior. This is an interventional strategy (in animals at least; in humans we usually let nature make the lesions). The other is a correlational strategy, where we measure physiological properties of the brain (electrical activity, microscopic structure, biochemistry, gene expression, and so on) to try to determine how they change as a result of experience.

To get a sense of the current state of the struggle, let's examine a form of declarative memory for which scientists have made substantial progress in illuminating the cellular and molecular substrates of the engram. After reports of the amnesiac patient H.M. became known in the 1950s, there was a determined effort to reproduce his deficit, complete anterograde amnesia for facts and events, in an animal model (preferably an inexpensive animal like a rat). It wasn't until the 1970s that this really started to pay dividends. It's not hard to use surgical

techniques to damage the hippocampus of a rat. The challenge was to find appropriate declarative memory tasks for this animal. The best ones turned out to be tests of spatial learning.

There are several ways to test spatial learning but the most widely adopted have had animals learn to navigate a maze in order to escape from a situation they find stressful. One clever maze was developed by Richard Morris and his colleagues at the University of Edinburgh. It's not what we normally think of when we hear the word "maze." This maze has no passageways. Rather it consists of a circular swimming pool 1.2 meters in diameter with a wall at the edge to prevent escape and dry milk powder added to make the water opaque. The pool is housed in a room with prominent and unique visual landmarks on the walls to aid in navigation. A rat (or mouse) is placed at a random location at the edge of the water and is then allowed to explore by swimming. Eventually, it will find that there is an escape platform, the top surface of which is just a centimeter or so below the opaque surface of the water. When the rat reaches the platform, it is allowed to stand there for a moment before being gently returned to its cage. The task is to remember where this escape platform is located so that on subsequent trials the rat can swim to it directly and make a quick exit. Not surprisingly, rats that have had their hippocampus surgically destroyed on both sides of the brain cannot learn the Morris water-maze task. Even after many trials, they behave as if they are experiencing the maze for the first time. This appears to be a specific deficit in spatial memory because they can easily learn to swim rapidly to a platform marked with a flag, which indicates that they do not merely have a problem with swimming or vision but rather have a genuine and specific memory deficit.

Also in the 1970s, a report on hippocampal physiology fired the imagination of brain researchers around the world. Terje Lomo of the University of Oslo and Tim Bliss of the National Institute of Medical Research in the United

Kingdom reported that if they briefly stimulated a population of glutamate-using excitatory synapses in the hippocampus of anesthetized rabbits at high frequency (100 to 400 stimuli per second for 1 or 2 seconds), this produced an increase in synaptic strength that could last for days. This phenomenon was named long-term synaptic potentiation (commonly abbreviated LTP). You can see why people got so excited. LTP was an experience-dependent, long-lasting change in neuronal function that occurred in a location in the brain already known to be crucial for memory. Furthermore, high-frequency bursts of the kind known to trigger LTP occur naturally in rats (and rabbits and monkeys). The hypothesis that LTP might underlie memory storage for facts and events in the hippocampus rapidly became one of the most exciting and controversial ideas in brain research.

In the years that followed, thousands of papers were published about LTP. One of the most interesting things scientists learned is that, although LTP was initially found in the hippocampus, it is actually a phenomenon that occurs throughout the brain. It is found in the spinal cord and in the cerebral cortex and almost everywhere in between. Although it is most commonly studied at excitatory synapses that use glutamate as a neurotransmitter, it is present in other types of synapse as well. Another important finding is that there is a complementary process: a persistent use-dependent weakening of synapses called long-term synaptic depression, or LTD. The exact parameters for evoking LTP and LTD vary from synapse to synapse, but at most locations LTP is produced by brief, high-frequency activation (100 stimuli per second for 1 second is typical) while LTD is produced by more sustained activation at moderate frequencies (say, 2 stimuli per second for 5 minutes). So far, all synapses that have LTP also have a form of LTD and vice versa.

Given that random, low-frequency spiking of neurons goes on all of the time, how does the synapse undergo LTP when there's a burst of high-frequency

stimulation, but not in the presence of ongoing background activity? This is a problem the brain has solved using several different molecular strategies. Here, I'll consider the most commonly used solution, which involves a special receptor for the neurotransmitter glutamate.

Previously, I've talked about glutamate receptors that rest with a closed ion channel and then open this channel when glutamate binds, allowing the sodium ions to flow in and potassium ions to flow out. This type of glutamate receptor is called an AMPA-type glutamate receptor (named after a synthetic drug that activates it strongly). These receptors cannot differentiate low-level background activity from high-frequency bursts. They are activated by both stimuli. The receptor that can make this differentiation (also named after a potent synthetic drug) is the NMDA-type glutamate receptor (Figure 5.9). The reason that the NMDA receptor can perform this trick is that, at the resting potential of −70 millivolts, its ion channel is blocked by a magnesium ion from the outside (magnesium ions float freely in the saltwater solution that surrounds neurons). This blockade remains until the membrane potential becomes more positive than about −50 millivolts.

So, neither glutamate binding alone nor depolarization of the membrane alone will open the NMDA-type receptor's ion channel. Background activity will produce the former but not the latter, but bursts of high-frequency spikes will produce both glutamate binding and depolarization and the ion channel will open. This ion channel is also unique in that it allows the influx of calcium ions together with sodium ions, while most AMPA-type receptors allow only sodium influx. This means that strong calcium influx through NMDA-type glutamate receptors is a unique consequence of high-frequency bursts. Or, stated another way, the NMDA receptor is a coincidence detector: it opens and fluxes calcium ions when both glutamate is released *and* the postsynaptic membrane is depolarized, but neither of these events alone is sufficient.

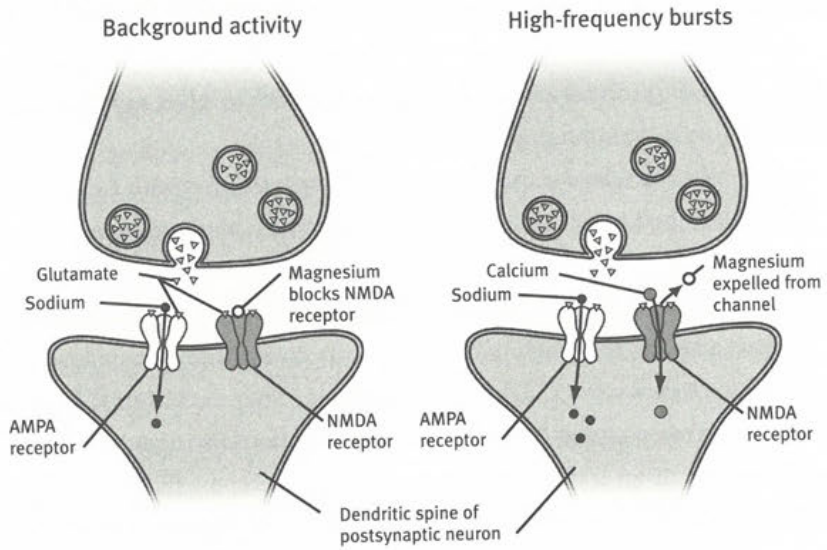


FIGURE 5.9. NMDA-type glutamate receptors and AMPA-type glutamate receptors. The NMDA receptors are activated by high-frequency bursts but not by background activity, because the voltage-dependent blockade of the NMDA receptor's ion channel by magnesium ions (Mg^{2+}) is only relieved when the postsynaptic membrane is depolarized to a level positive to −50 millivolts. The AMPA receptors are activated by both background activity and high-frequencey bursts. Adapted from L. R. Squire and E. R. Kandel, *Memory: From Mind to Molecules* (Scientific American Library, New York, 1999). *Joan M. K. Tycko, illustrator.*

If this process is the trigger for LTP, then drugs that block the NMDA receptor should also block LTP. This is indeed what happens in most hippocampal synapses. Furthermore, if one injects neurons with drugs that rapidly bind calcium ions as soon as they enter the cell, thereby preventing them from interacting with other molecules, this will also prevent LTP. Calcium ions entering

through NMDA receptors can activate lots of different calcium-sensitive enzymes in the neuronal dendrite. Rapid, large calcium transients can activate an enzyme called calcium/calmodulin protein kinase II alpha, typically abbreviated CaMKII. This enzyme transfers chemical phosphate groups onto proteins to change their function. Although the substrates of CaMKII action relevant for LTP are not known, one popular hypothesis is that this process ultimately results in the insertion of new AMPA-type receptors into the postsynaptic membrane, thereby strengthening the synapse. It should be mentioned that though this NMDA receptor → CaMKII → AMPA receptor insertion cascade is the most common form of LTP, it is not the only one. There are others that can use different biochemical steps and produce LTP through different means (such as increased glutamate release or increased conductance of existing AMPA receptors).

What about LTD? How does sustained synaptic activation at moderate frequencies result in persistent synaptic weakening? Interestingly, in its most common form, LTD also uses the NMDA receptor. In this case, moderate frequency stimulation results in partial relief of the magnesium ion blockade of the NMDA receptor. This produces a calcium flux, but one that is small and sustained rather than large and brief. Small, sustained calcium signals are insufficient to activate CaMKII and therefore don't produce LTP. Instead, they activate an enzyme that does the opposite job: protein phosphatase 1 (PP1) removes phosphate groups. Activation of PP1, not surprisingly, ultimately results in the removal of AMPA receptors from the postsynaptic membrane, thereby depressing synaptic strength in a way that is the functional opposite of LTP. This LTD cascade involving NMDA receptor → PP1 → AMPA receptor internalization is a dominant form of LTD in the hippocampus, but it is only one of several mechanisms for producing persistent depression of synaptic strength.

Thus both LTP and LTD can be produced in several different ways. In reality, some individual synapses are able to express multiple forms of both LTP and LTD.

So, the model being developed here is that somehow, memory for facts and events, including memory for the location of the escape platform in the Morris water maze, is encoded by producing LTP and LTD in an array of hippocampal synapses, and these forms of LTP and LTD are critically dependent upon triggering by NMDA receptors. One central test of this hypothesis was to inject rats with NMDA receptor–blocking drugs to see if they could learn the Morris water-maze task in conditions where LTP and LTD were mostly blocked. This experiment, which has now been repeated several times in different labs, showed that spatial memory was indeed severely impaired under these conditions. Later, a similar result was obtained by using mutant mice that failed to express functional NMDA receptors in a crucial subregion of the hippocampus (called area CA1; see Figure 5.10). In all of these cases, the general sensory and motor functions of these mice were largely intact—the failure in the maze task appeared to be a genuine memory deficit and not a trivial defect in seeing or swimming or stress level.

Would it be possible to train rats in the Morris water maze task and then analyze their hippocampal tissue? Many attempts have been made over the years to look for the electrical, biochemical, or structural correlates of learning in the hippocampus. There have been intermittent claims, but in truth not much has come of these efforts. Here's the problem. Spatial learning is likely to produce changes in a very small fraction of spatially distributed hippocampal synapses, and we don't have a good way to know where these synapses are. So, whether you're recording synaptic strength electrically or making biochemical or structural measurements, there's a big "needle-in-a-haystack" problem: it's almost

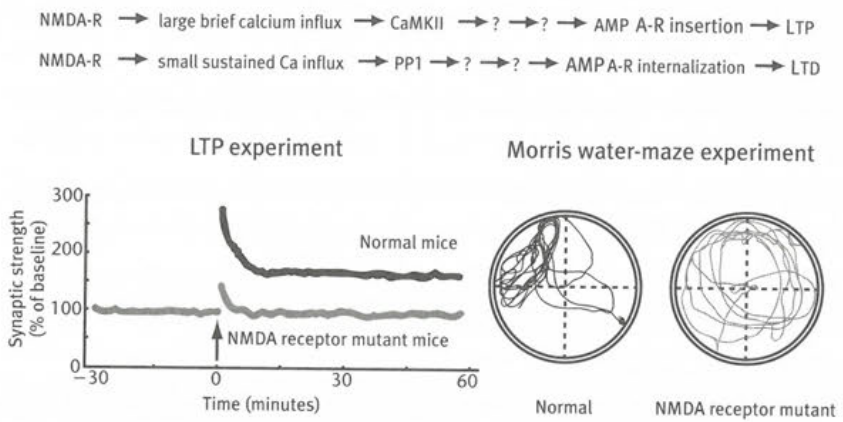


FIGURE 5.10. An experiment showing that mutant mice lacking functional NMDA-type glutamate receptors in a crucial region of the hippocampus have impaired LTP, LTD, and spatial learning. The top panel shows the signaling cascades triggered by the NMDA receptor to induce both LTP and LTD. The question marks indicate that there are multiple steps leading to AMPA receptor insertion and internalization that we still do not understand. The lower left panel is a plot of synaptic strength as a function of time in an LTP experiment. LTP was induced by applying high-frequency bursts to the presynaptic axons at the point indicated by the upward arrow. The lower right panel shows the path of well-trained mice in a Morris water maze. These are the results of a probe trial in which the platform is removed to see where the mouse will hunt for it. The normal mouse has a well-established memory for the correct platform location in the upper left quadrant while the LTP/LTD-lacking mutant mouse has little memory for the location and therefore searches widely in the water maze. Adapted from J. Z. Tsien, P. T. Huerta, and S. Tonegawa, The essential role of hippocampal CA1 NMDA receptor-dependent synaptic plasticity in spatial memory, *Cell* 87:1327–1338 (1996). *Joan M. K. Tycko, illustrator.*

impossible to measure the relevant changes when they are diluted in a sea of other synapses that are not a part of the memory trace.

The results showing that treatments that interfere with hippocampal NMDA receptor function block spatial learning in rats and mice *suggest* that our working model is correct: the engram for declarative memory in the hippocampus requires LTP and LTD. Do these findings *prove* this hypothesis? Unfortunately, no. Although hippocampal NMDA receptor manipulations interfere with spatial learning, attempts to interfere with LTP and LTD by targeting biochemical signals that follow NMDA receptor activation have met with mixed success. One can block most forms of LTP or LTD by interfering with CaMKII or PP1 or certain types of AMPA receptors, but this will not always produce a deficit in spatial learning tasks. In addition, it's very likely that these manipulations affect many processes in addition to LTP and LTD. The calcium flux through the NMDA receptor activates many enzymes, not just PP1 and CaMKII. There is further divergence of the signaling cascade as we move along: CaMKII, for example, transfers phosphate groups to hundreds of proteins in hippocampal neurons, not just those involved in LTP. As a consequence, one cannot be completely confident that the blockade of spatial learning produced by these drugs or molecular genetic tricks is really due to an LTP/LTD deficit as opposed to some side effect.

To summarize, we know that destroying the hippocampus will prevent spatial learning in rats and mice, and there is suggestive, but not conclusive, evidence that memory for locations in space is stored in the hippocampus by changing the strength of synapses through LTP and LTD. How does making certain synapses weaker or stronger in the hippocampus give rise to the behavioral memory that allows an animal to learn the Morris water maze or another spatial task? The short answer is that we don't know. The hippocampus is

not anatomically or functionally organized in a way that makes this obvious. The slightly longer answer is that even through we don't know, there is an interesting hint that may be relevant to this difficult problem.

John O'Keefe, Lynn Nadel, and their coworkers at University College, London, made recordings from neurons in the hippocampus of rats as they explored an artificial environment in the lab. What they found was that about 30 percent of one class of cells in the hippocampus (called pyramidal cells) seemed to encode the animals' position in space. When a rat is placed in a new environment and has a chance to explore, recordings will reveal that, after a few minutes, one cell fires only when the animal is in a particular location (say, the upper left edge of a large circular cage (see Figure 5.11). This particular cell, referred to as a "place cell," will once again fire in this fashion even if the rat is removed from this environment and returned days to weeks later. Place cells have been found in mice as well as rats. Recording from additional cells reveals that there are place cells that fire specifically for all different parts of the explored environment. Some are fairly sharply tuned for place (Figure 5.11) while others fire over a broader area. When place cells are recorded in mutant mice in which the hippocampal pyramidal cells have been engineered to have a form of CaMKII that is always on (they can't have more LTP because it is already turned up to maximum levels), interesting properties emerge. Place cells do form characteristic firing patterns when the mouse explores an environment, but then, when the animal returns to the environment, the tuning of these cells tends to change (Figure 5.11). Because these mutant mice are also impaired in spatial learning tasks, it has been suggested that LTP is required to maintain the tuning of place cells and that these place cells form a cognitive map of space that allows the animal to store spatial memory.

The problem is that the physical details of this cognitive map in the hippo-

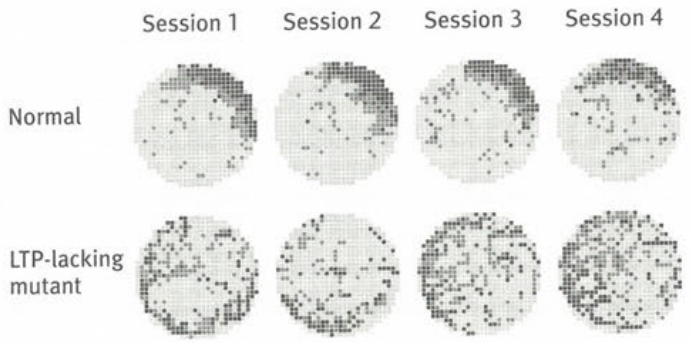


FIGURE 5.11. Place cells in the hippocampus of mice. These figures show the firing rate of individual pyramidal cells in the hippocampus of a mouse exploring a circular environment. Black pixels indicate high firing rates, and light gray low firing rates, for that particular location. Place cells from normal mice may be either sharply tuned (as shown here in the example from the normal mouse) or broadly tuned, but the representation of space is stable over many training sessions. Recordings made from a mutant mouse that lacks LTP (it has been engineered to have a form of CaMKII that is always turned on in these neurons) show that place fields are not stable over repeated training sessions. This correlates with a failure of these mutant mice to learn certain spatial tasks. Adapted from A. Rotenberg, M. Mayford, R. D. Hawkins, E. R. Kandel, and R. U. Muller, Mice expressing activated CaMKII lack low frequency LTP and do not form stable place cells in the CA1 region of the hippocampus, *Cell* 87:1351–1361 (1996). *Joan M. K. Tycko, illustrator.*

campus are not obvious. Sensory systems have maps that represent the external world anatomically in the brain: adjacent cells in the primary visual cortex will be activated by light coming from adjacent points in the visual field. Likewise, adjacent cells in the primary somatosensory cortex will be stimulated by touch on adjacent points on the body surface. But although different hippocampal neurons that code for the same location in space tend to fire together, they are

not physically organized in any coherent fashion. One cell that codes for the upper left quadrant of the environment may be located at the opposite end of the hippocampus from another cell that codes for the same region, and cells in the intervening tissue are not organized in any fashion to represent the spatial world. So, while we are beginning to gain an understanding of the molecular processes that represent experience as changes in neuronal function (LTP, LTD, and changes in intrinsic excitability) and structure, and there is some evidence emerging to link these processes with specific forms of learning, we are still far, far away from a complete "molecules-to-behavior" explanation of declarative memory.

WE HAVE SEEN that the brain does not use a single cellular process or a single brain region to store memory. Rather, memory storage involves multiple brain locations and several broad classes of mechanism (synaptic plasticity, intrinsic plasticity), each of which can be produced by a number of different molecular strategies. Crucially, the cellular and molecular mechanisms of memory storage are not unique. In true, kludgy evolutionary fashion, the mechanisms for storing memory have been largely adapted from those designed to wire up the brain in response to experience during the later stages of development (during late pregnancy and early childhood).

Let's put this back in a historical context. The design of the brain has been limited as it has evolved by three main considerations.

1. During the course of evolution, the brain has never been redesigned from the ground up. It can only add new systems onto existing ones.
2. The brain has a very limited capacity for turning off control systems, even when these systems are counterproductive in a given situation.

3. Neurons, the basic processors of the brain, are slow and unreliable, and they have a rather limited signaling range.

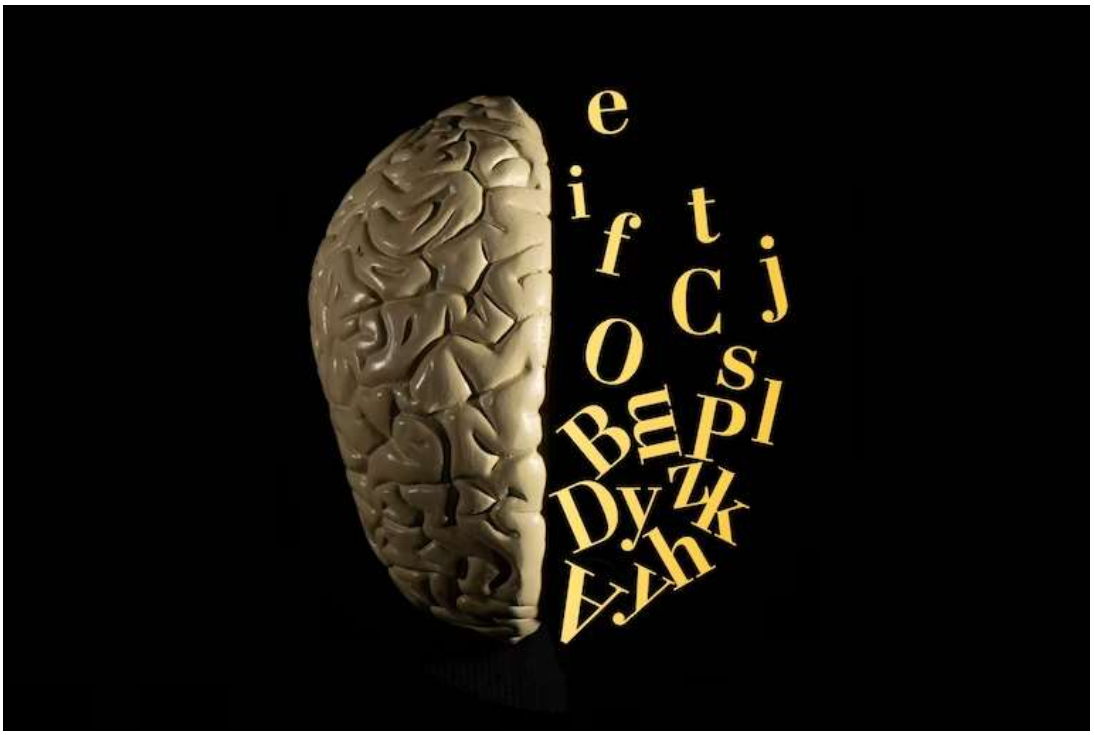
These considerations have driven the brain's solution to the problem of building computational complexity: a brain that has a huge number of neurons and in which these neurons are highly interconnected. This big complex brain creates two problems. How do you get a large head through the birth canal? And how do you specify the wiring diagram for 500 trillion synapses genetically? The solutions, as previously discussed, have been to only roughly specify the wiring diagram of the brain genetically and to reserve significant brain growth and synapse formation until after birth. This design allows a head that can pass through the birth canal. It also allows sensory experience to guide the fine-scale wiring of the brain. In order to do that, there had to be mechanisms by which particular patterns of sensory experience could drive changes in synaptic strength (LTP and LTD), intrinsic excitability, and the growth and retraction of axonal and dendritic branches as well as synapses. These, of course, are the same cellular and molecular mechanisms that, with slight elaboration, are retained in the mature brain to store memory.

This is the ultimate example of "when life gives you lemons, make lemonade." Our memory, which is the substrate of our consciousness and individuality, is nothing more than the accidental product of a work-around solution to a set of early evolutionary constraints. Put another way, our very humanness is the product of accidental design, constrained by evolution.

When the words won't come. This is my life with aphasia.

The disorder results from damage to the brain that affects speech and language comprehension, and it's far more common than many realize

By Judith Hannah Weiss
The Washington Post
February 25, 2023 at 7:45 a.m. EST



(Illustration by Elizabeth von Oehsen/The Washington Post; Marvin Joseph/The Washington Post)

Imagine that you're trying to talk, but you can't get the words out — and then, if you finally do, no one understands what you're saying. And you don't understand what others are saying to you. That's what it's like to live with aphasia.

Aphasia results from damage to the brain that affects speech and language comprehension. Frequently, aphasia follows a stroke, but it can also result from a traumatic brain injury; in my case, I suffered a "coup contrecoup injury with diffuse axonal shearing of the brain" — and, consequently, aphasia — when a drunk driver plowed into a parked car that I was sitting in one Tuesday morning in 2006.
I'm sharing my story not because I think it is exceptional, but because I know it is not. If anything, it's the telling that makes it unusual because so few of us with aphasia can speak about the challenges we face.

At least 180,000 Americans are diagnosed with aphasia every year, and it's estimated that some 2 million Americans have it; it's more prevalent than Parkinson's disease, cerebral palsy, multiple sclerosis, muscular dystrophy and Lou Gehrig's disease combined. Yet the condition remains largely in the shadows, maybe in part because so few of us with it can tell others about the challenges we face. Actor Bruce Willis and former congresswoman Gabrielle Giffords are perhaps the most famous people to have publicly acknowledged their aphasia. (Willis's diagnosis, it was recently announced, has progressed to frontotemporal dementia.)

In research from the National Institutes of Health, aphasia had the largest negative impact on quality of life of any of 60 measured conditions, more than cancer and Alzheimer's disease.
I'm sharing my experience to give hope to others with aphasia and to their families.

A brain stuck on static

Within days of my injury, I could unstick my tongue from the roof of my mouth and create an odd sound every now and then, but I couldn't communicate in any traditional sense. I felt like a human radio pumping out static — with sporadic bursts of clarity.

When I was asked to point to a picture of a teapot, an apple, an elephant, my adrenaline kicked in, my breathing got faster, my heart rate got faster, and I started to sweat. Sometimes I just pointed to my head. The odds of a sinkhole opening within me were approximately equal to the odds that I'd find the right word at the right time, something I'd done with ease before the accident as a professional freelance writer.

I couldn't navigate the smallest space or the smallest thing. None of the tools I had used before made any sense. Not words or places or names or directions or signs on bathroom doors. It's hard to navigate when you can't decipher anything on your desktop or phone and can't tell anyone that you can't.

I pointed to a chair because I couldn't say "chair." I mimed drinking from a bottle because I couldn't find the word "bottle" or "water" or "thirsty" or "drink." I spoke, if I spoke at all, with an urgency bordering on panic. In the first year after the accident, once I began to put words together, I said things like "white stuff sky," which meant snow, or "cow thing pants," which meant belt.

My condition still interferes with my life, although not the way it did in those early days. But it flares up when I have to be somewhere I can't find or do something I can't do or say something I can't say.
More than 16 years after the accident, I can have one conversation every two or three days. Then I wilt. I still want to hide my mind, or at least the damaged part.

Recovery and coping

Most of my recovery grew from my own self-devised therapies. About a year after the accident, I began listening to audiobooks, one sentence at a time. I would pause, replay, pause, replay and try to repeat the words. Although I'm no longer doing that, I can still only "read" audiobooks. I also long ago began writing anything I could recall on any surface I could find. I stuffed scraps in brown paper shopping bags, then began to build a book.

People living with aphasia are the real experts on this often overlooked condition, but we rarely hear their voices. It's hard to speak up when you can't speak.

Some things I've learned about aphasia

Although some sources give the number of Americans with aphasia as at least 2 million, Swathi Kiran, the director of the Aphasia Research Laboratory at Boston University, says that number represents only aphasia caused by stroke.

The nonprofit organization AphasiaAccess says that, every approximately four minutes, someone in the United States acquires aphasia, and that as many as 4 million Americans may have the condition.

Tests say they quantify cognition, but I would say they quantify only what we can say, which is not the same as what we know or feel. We need words for that, plus neural functions, synapses and soul.

The boundaries between "able" and "disabled" are fluid, not fixed. Sometimes you're able, sometimes you're not, depending on what's going on inside you and around you at any given time. It's hard to learn your limitations when they change every few moments.

Aphasia cannot be described or prescribed with one set of secrets or tactics or remedies. The damage from aphasia is variable and involves different deficits unique to each person.

The one constant is how hard those of us with aphasia work to achieve even small results.

Judith Hannah Weiss is a freelance writer and the author of an as-yet unpublished memoir titled "Away With Words."

Quantamagazine

Light-Triggered Genes Reveal the Hidden Workings of Memory

By Elizabeth Svoboda

December 14, 2017

Nobel laureate Susumu Tonegawa's lab is overturning old assumptions about how memories form, how recall works and whether lost memories might be restored from "silent engrams."

Neuroscientists gained several surprising insights into memory this year, including the discovery that the brain creates multiple copies of memories at once — even though it hides the long-term copy from our awareness at first.

Eero Lampinen for Quanta Magazine

Susumu Tonegawa's presence announces itself as soon as you walk through the door of the Massachusetts Institute of Technology's Picower Institute for Learning and Memory. A three-foot-high framed photograph of

Tonegawa stands front and center in the high-ceilinged lobby, flanked by a screen playing a looping rainbow-hued clip of recent research highlights.

The man in the portrait, however, is anything but a spotlight-seeker. Most days, he's ensconced in the impenetrable warren of labs and offices that make up Picower's fifth floor. His hair, thick and dark in the photo, is now a subdued silver, and today, a loosely draped blue cardigan replaces the impeccable suit jacket. His accommodating, soft-spoken manner belies his reputation as a smasher of established dogma, or at least as a poker of deep and abiding holes.

Along with his MIT neuroscientist colleague Dheeraj Roy and others, Tonegawa is upending basic assumptions in brain science. Early this year, he reported that memory storage and retrieval happen on two different brain circuits, not on the same one as was long thought. His team also showed that memories of an event form at the same time in the brain's short-term and long-term storage areas, rather than moving to long-term storage later on. Most recently (and tantalizingly), his lab demonstrated what could someday be a way to bring currently irretrievable memories back into conscious awareness.

Tonegawa, now MIT's Picower Professor of Biology and Neuroscience, first carved out his maverick identity back in the 1980s. While at the Basel Institute for Immunology in Switzerland, he published a theory — first seen as heretical, then brilliant — that immune cells reshuffle their DNA to create millions of different antibodies from a small number of genes. His discovery won him the Nobel Prize in 1987, which explains the oversized lobby portrait. Most researchers would have stayed in the field and basked in the attention, but Tonegawa left immunology behind entirely. He spent the next couple of decades reinventing himself as a master of memory's workings at the cellular level.

Despite his professional stature, Tonegawa is no TED-circuit regular or fount of startup concepts. Instead of selling his ideas or his persona, he prefers to let his data speak for themselves. And they do, perhaps more loudly than some of his colleagues would like. “The way he continues to disrupt and innovate is really striking,” said Sheena Josselyn, a neuroscientist at Toronto’s Hospital for Sick Children who also studies memory formation. “He tackles the tough questions. He doesn’t do something that is easy and expected.”

Tracking Memories Cell by Cell

Upon meeting Tonegawa, I sensed that he considers his fame a slightly cumbersome side effect of his vocation. The day I visited his office, he was immersed in research banter with a colleague, breaking away only reluctantly to revisit his own journey. The whole immunology sideline, he told me, was something of an accident — his real love has always been molecular biology, and immunology was a fascinating expression of that. He ended up at Basel mostly because his U.S. work permit had run out.

“Immunology was a transient interest for me,” he said. “I wanted to do something new.”

After making Nobel Prize-winning contributions to immunology, Susumu Tonegawa, now a professor of biology and neuroscience at the Massachusetts Institute of Technology, focused his passion for molecular biology on the brain.

Tonegawa Lab

That “something” turned out to be neuroscience, which Francis Crick and other well-known biologists were touting as the wave of the future. In the late 1980s and early ’90s, researchers knew relatively little about how the cellular and molecular workings of the brain underpin its capabilities, and nothing excited Tonegawa more than mapping unexplored territory.

Tonegawa’s venture into brain science wasn’t a complete turnabout, though, because he brought some of his investigative techniques with him. He had been using transgenic (genetically modified) mice in his immunology studies, knocking out particular genes and observing the

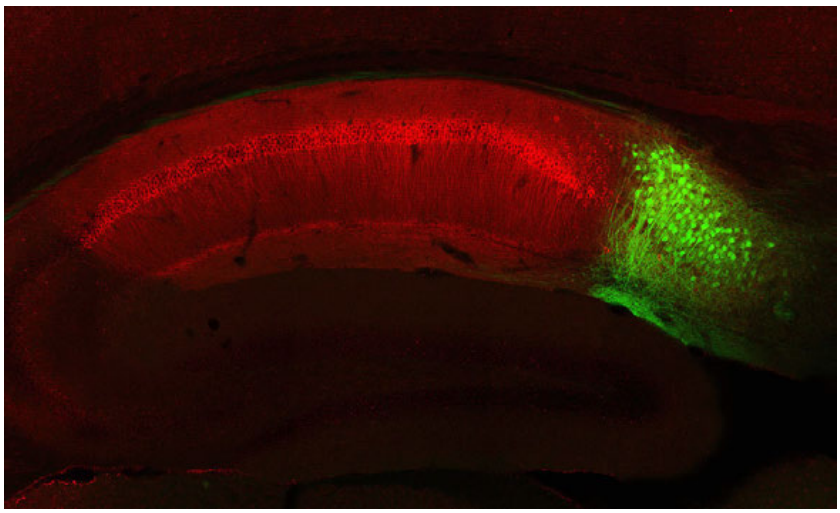
physical effects, and he used a similar approach to uncover the biological basis of learning and memory. In an early MIT study, he bred mice that did not produce a particular enzyme thought to be important in cementing long-term memories. Although the behavior of the mutant mice seemed mostly normal, further testing showed that they had deficiencies in spatial learning, confirming the enzyme's key role in that process.

With that high-profile result, Tonegawa was off and running. About 10 years ago, he was able to take his work to a new level of precision in part by employing a technique called optogenetics. Developed by the Stanford University bioengineer Karl Deisseroth and others, the technique involves modifying the genes of lab animals so that their cells express a light-sensitive protein called channelrhodopsin, derived from green algae. Researchers can then activate these cells by shining light on them through optical fibers. Tonegawa and his colleagues use optogenetics to generate neural activity on command in specified regions of the brain.

This method has allowed Tonegawa to show that existing theories about memory formation and storage are wrong, or at least incomplete. This past summer, along with Roy and other colleagues, he reported that — contrary to neuroscience dogma — the neural circuit in the brain structure called the hippocampus that makes a particular memory is not the same circuit that recalls the memory later. Instead, retrieving a memory requires what the scientists call a "detour circuit" in the hippocampus's subiculum, located just off the main memory-formation circuit.

To illustrate the discovery for me, Roy called up an image of a magnified brain slice in the lab. "What you're looking at is the hippocampus section of a mouse," he said. He gestured to a dense cloud of glowing green neurons in the upper right — the subiculum itself — and explained that his team had genetically engineered the mouse to produce channelrhodopsin only in the

subiculum's neurons. He and his team could then activate or deactivate these subiculum neurons with piped-in laser light, leaving the surrounding neurons unaffected.



Studies have shown that the hippocampus (red) is essential for creating new memories. But short-term recall of those memories depends on a "detour circuit" involving a specialized area called the subiculum (green).

Dheeraj Roy/Tonegawa Lab, MIT

Armed with this biological switch, the researchers turned the subiculum neurons on and off to see what would happen. To their surprise, they saw that mice trained to be afraid when inside a certain cage stopped showing that fear when the subiculum neurons were turned off. The mice were unable to dredge up the fearful memory, which meant that the subiculum was needed for recall. But if the researchers turned off the subiculum neurons only while teaching the fearful association, the mice later recalled the memory with ease. A separate part of the hippocampus must therefore have encoded the memory. Similarly, when the team turned the main hippocampal circuit on and off, they found that it was responsible for memory formation, but not for recall.

To explain why the brain would form and recall memories using different circuits, Roy framed it in part as a matter of expediency. “We think these parallel circuits help us quickly update memories,” he said. If the same hippocampal circuit were used for both storage and retrieval, encoding a new memory would take hundreds of milliseconds. But if one circuit adds new information while the detour circuit simultaneously calls up similar memories, it’s possible to apply past knowledge to your current situation much more quickly. “Now you can update on the order of tens of milliseconds,” Roy said.

That difference might prove crucial to creatures in danger, for whom a few hundred milliseconds could mean the difference between getting away from a predator scot-free and becoming its dinner. The parallel circuits may also help us integrate present information with older memories just as speedily: Memories of a new conversation with your friend Shannon, for instance, can be added seamlessly to your existing memories of Shannon.

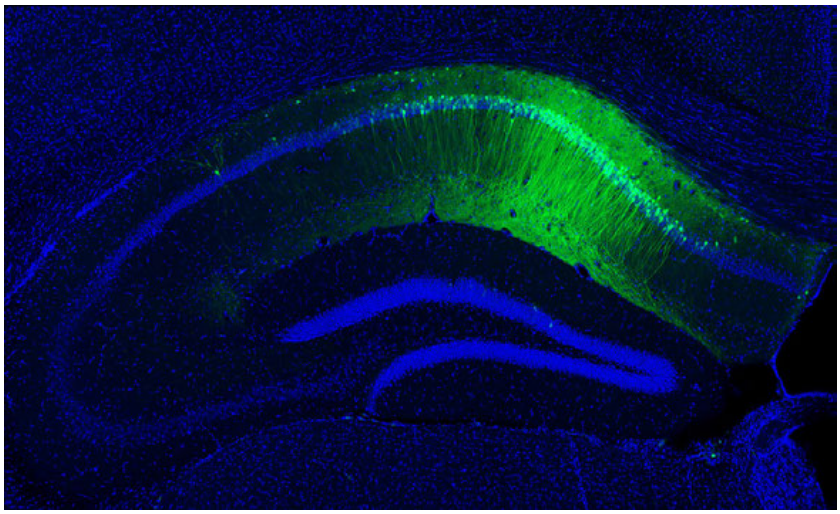
Reassessing How Memories Form

In addition to revealing that different mechanisms control memory formation and recall, Tonegawa, Roy and their colleague Takashi Kitamura (who recently moved from MIT to the University of Texas Southwestern Medical Center) have shown that memory formation itself is unexpectedly complex. Their work concerned the brain changes involved in the transformation of short-term memories to long-term memories. (In mouse experiments, short-term memory refers to recollections of events from within the past few days — what is sometimes called recent memory to distinguish it from more transient neural impressions that flicker out after only minutes or hours. Long-term memory holds events that happened on the order of two weeks or more ago.)

For decades in neuroscience, the most widely accepted model posited that short-term memories form rapidly in the hippocampus and are later transferred to the prefrontal cortex near the brain's surface for long-term storage. But Tonegawa's team recently reported in *Science* that new memories form at both locations at the same time.

The road to that discovery started back in 2012, when Tonegawa's lab came up with a way to highlight brain cells known as engram cells, which hold a unique memory. He knew that when mice take in new surroundings, certain genes activate in their brains. His team therefore linked the expression of these "experiential-learning" genes in the mice to a channelrhodopsin gene, so that the precise cells that activated during a learning event would glow. "You can demonstrate those are the cells really

holding this memory," Tonegawa said, "because if you reactivate only those neurons with laser light, the animal behaves as if recalling that memory."



In this magnified slice of brain tissue enhanced with an optogenetic protein, the green glow shows which engram cells in the hippocampus stored a short-term memory.

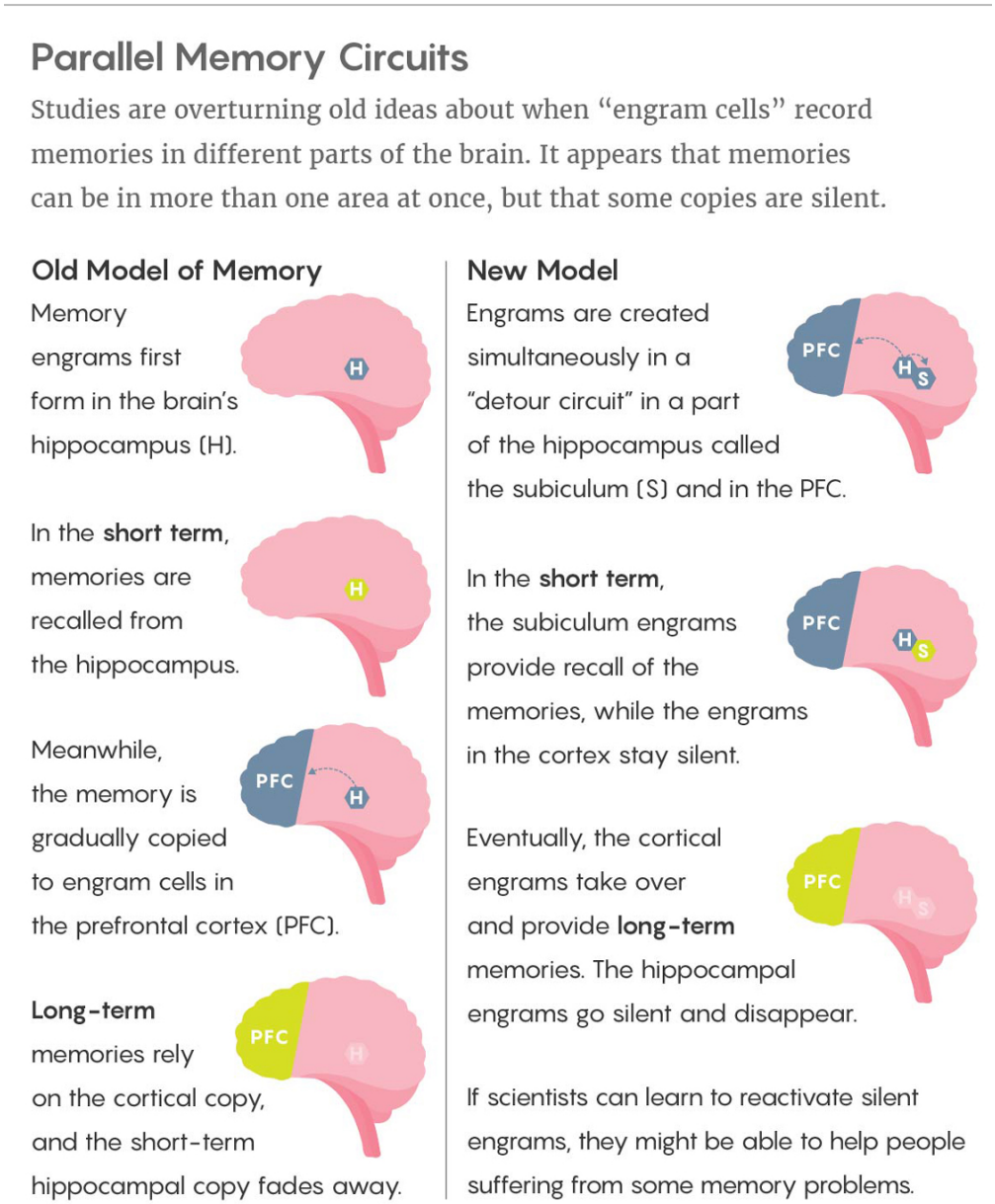
Dheeraj Roy, Tonegawa Lab/MIT

In the new *Science* study, the team used this technique to create mice whose learning cells would respond to light. They herded each mouse into a special cage and delivered a mild electric shock to its foot, leading the mouse to form a fearful memory of the cage. A day later, they returned each mouse to the cage and illuminated its brain to activate the brain cells storing the memory.

As expected, hippocampal cells involved in short-term memory responded to the laser light. But surprisingly, a handful of cells in the prefrontal cortex responded as well. Cortical cells had formed memories of the foot shock almost right away, well ahead of the anticipated schedule.

Yet the researchers noticed that even though the cortical cells could be activated early on with laser light, they did not fire spontaneously when the mice returned to the cage where the foot shock happened. The researchers called these cortical cells “silent engrams” because they contained the memory but did not respond to a natural recall cue. Over the next couple of weeks, however, these cells seemingly matured and became integral for recalling the memory.

“The dynamic is, the hippocampal engram is active [at first] and goes down, and the prefrontal-cortex engram is silent at the beginning and slowly becomes active,” Tonegawa said. This detailed understanding of how memories are laid down and stored could inform the development of drugs that aid formation of new memories.



Lucy Ikkanda-Reading/Quanta Magazine

Some in the neuroscience community, however, think it's prudent to be cautious in interpreting the significance of findings like these. Last year, Tonegawa's MIT colleagues Andrii Rudenko and Li-Huei Tsai emphasized that engram science is still so new that we don't know exactly how engram cells might work together, nor which cells contain which parts of memories. "In these early days of functional memory engram investigation," they

wrote, “we still do not have satisfactory answers to many important questions.”

Tonegawa has asserted that brains contain silent engrams that could potentially be externally activated — an idea that strikes a few neuroscientists as overblown even as it excites others, according to Josselyn. “It really forces the scientific community to either update our thinking or try experiments to challenge that,” she said.

Bringing Silent Memories to Life

Despite the uncertainty that surrounds it, the silent-engram concept offers us the fascinating prospect of gaining access to hidden memories — a prospect that Roy, in particular, continues to explore. In October, he published a paper with Tonegawa that generated a flurry of excited emails from scientists and nonscientists alike. One of the paper’s blockbuster

findings was that, at least in mice, it was possible to awaken silent engrams without using a laser light or optical fibers.

Dheeraj Roy, a postdoctoral associate at MIT, has collaborated with Tonegawa on several recent studies that have overturned old ideas about how memory works.

Vicky Roy

The question the team asked themselves, Roy said, was whether they could make hidden memories permanently active with a noninvasive treatment. A cellular protein called PAK1 stimulates the growth of dendritic spines, or protrusions, that allow communication between neurons, and Roy had a hunch that this protein — when transported into brain cells — might help bring silent engrams back into direct awareness. “Can we artificially put [in] more of one gene that would make more protrusions?” he asked, excitedly noting that this approach might be simpler than optogenetics.

To test this possibility, the researchers first gave mild shocks to mice in a cage while also suppressing their ability to make the proteins that normally cement long-term memories. When these mice returned to the same cage later on, they showed no fear, indicating that they did not naturally recall the shock in response to a cue. Yet laser light could still switch on the mice's fearful response, which meant the memory was still there in silent-engram form.

When the team injected these mice with the PAK1 gene to make them overproduce the protein, the animals froze up spontaneously when entering the dreaded cage. They were recalling the memory of the cage all on their own: The silent engram was coming to life. When PAK1 is administered, "you just wait four days, [and] they recover it with natural cues," Roy said. In the future, he added, a therapeutic injection of PAK1 molecules that enter the brain's memory cells could awaken people's silent memories as well.

"So it would just be an injected protein?" I asked.

"That's right — one molecular transporter that has one protein. People already have ways to put proteins into brain cells. I don't think we're that far [away] anymore."

It's amazing to think that all of our minds hold hundreds or thousands of silent memories that are just waiting for the right activation to re-emerge into conscious awareness. If Roy's findings hold true in humans, the retrieval of hidden memories might someday be as easy to initiate as getting a flu shot. "What would happen if you did that to a normal person? What would come flooding back?" I asked. "What would that experience be like?"

“Very sci-fi, even for me,” Roy said. “My family says, ‘Is this all real?’ I say, ‘Yeah, I’m not lying to you!’”

A few minutes later, back in Tonegawa’s office, I posed more or less the same question to him. Reactivating silent engrams could allow people with memory issues — like Alzheimer’s sufferers, soldiers who have survived explosive blasts and concussed athletes in contact sports — to regain memories that have become inaccessible. (To be sure, these people would often need to get such treatments early, before their conditions progressed and too many brain cells died.) Roy and Tonegawa’s past research suggests that people with cognitive difficulties have many stored memories that they simply can’t recall. But what about the rest of us who just want to mine our memories, to excavate what’s buried deep within?

Tonegawa paused to consider. “It could be these silent memories could come out,” he said. “If you artificially increase the spine density, inject enzymes which promote spine formation, then the silent engram can be converted to active engram.”

When I pressed him further, though, he exuded caution. It was as if he was used to hearing people like me run away with the possibilities and wanted to tamp down my expectations. Even though his lab successfully reactivated mice’s silent engrams after a few days, that’s no guarantee that silent engrams last very long, he said. And once the cells that encode particular memories die off from old age or dementia, it might be game over, no matter what kind of proteins you inject. Tonegawa pointed to Roy, who was sitting across from him. “I won’t remember his name.”

His patience seemed to be running out. The contrarian in him, I could tell, wanted to assert that he was a student of the essential nature of things, not a pursuer of drug patents or quick cures or even the ideal of perfect recall.

"I know a joke," he said cryptically. "Not injecting protein or genes, but I keep an external brain. I hold the information in that brain." He pointed to Roy again — the person he counts on to remember things he can't. "The only thing I have to do is have a relationship with that person," he explained. It's comforting, in a way, to know that the wizard of tracing and unlocking memories also believes that no brain is an island. "It's better," he said, "not to memorize everything."

This article was reprinted on Wired.com.