Chapter 3

Connecting

Let me begin again with the snails in my garden. When I pick one up, its shell is thin and resilient rather than hard. Usually the snail just squirms, waving its horns at me rather than retreating within. Yet my eyes do not actually see the snail, nor my fingers feel it. Instead, they make possible the internal events I call "seeing" and "feeling." The sensations are mental reconstructions, re-presentations, of the external actions (all within the usual Kantian constraints).

The question is how the brain does this. There are actually at least three questions here. First, at the most general level, how does the brain create this seamless reenactment of flow that we experience as the mind? Secondly, how does the brain structure the sensory data for its internal syntheses: what is the logic of representation? Finally, what is the *process* of representation? How does the brain build its internal systems?

Much is at stake here, nothing less than the content of human experience. These questions focus on the route from neural structuring to those large-scale terms in which we describe our lives: sensation, perception, thought, intention, emotion, as well as trees, rocks, people, money, and poetry. What is the neural "content" of these large-scale terms. For example, what sort of "stuff" does the brain pull together in order for us to think about a rock? There are colors, textures, a sense of

heft, the letters R-O-C-K, the sound of the word, "rok," and all the associations that linger on the borders: "rock-paper-scissors," "upon this rock I will build my Church (Matt. 16.18)," Plymouth Rock, Rock of Ages, and on and on. The brain has found a strategy for not only representing all these facets but also for linking them together.[1] In itself, this feat may not seem especially remarkable, but it unobtrusively touches upon some weighty issues. Are rocks a "natural kind," a category in the world that would exist even without people? Moreover, are there aspects of the experience of "rock" that are built into the organization of the brain? This latter question may seem odd when asked about "rock," but what about "mother?"

How does our biology intervene in human language and the structure of human experience? Must "mother" be our biological mother? Must "mother" be female? This sort of question raises the specter of biological determinism, the thought that biological imperatives shape the surface content of experience as well as its deep structure. It becomes clear that those who reject biological determinism also have a stake, therefore, in the development of a neuroscientific model for the neuronal representation of experience. In the humanities, resistance to the reductive claims of the sciences has become very important and—in its more radical form—has taken the shape of a specific counterproposal: social constructionism.

---

[1] Terrence Deacon in *The Symbolic Species: the Co-evolution of Language and the Brain* (New York: W. W. Norton & Co., 1997) pursues the question of the strategies for linkage. Humans organize access to internal representations in a structurally superior manner through the increased role of the prefrontal cortex. The issues here are important, but I do not explore them in this book.

Construction, Difference, and Meaning

One of the great liberating moments in early twentieth century thought was Ferdinand de Saussure's realization that words don't mean *absolutely* but instead gain their meaning from their position within a large system of words. That is, the word "water" (signifier) refers to a thing WATER (signified) as a matter of learned convention.[2] Secondly, however, "water" is part of a system of sounds (the phonological structure) that speakers of English deem significant. If I said "whater" (adding aspiration), most speakers would ignore the "h" as meaningless. In contrast, some speakers of English who still preserve the distinction between "wail" and "whale" might consider my pronunciation odd indeed. The aspiration matters to them. This logic of meaning by difference is easy to see in the *sounds* of words. It is far more difficult to articulate in the meaning of words, yet the argument is the same. Rocks differ from pebbles, stones, and boulders. They differ from other materials: wood, plastic, metal, flesh. Yet petrified wood is a rock, while I would call a large clear tourmaline a crystal rather than a rock. Our notions of rocks are hemmed in—and indeed defined— by contrasts with non-rocks. We can develop a similar set of contrasts to define each of the non-rocks we listed, and on and on. For the spoken word, the set of contrasts used to define a language system derives from the physical structure of the mouth. What, however, is the nature of the web of mutual contrasts

[2] To use the examples from the Introduction, the Chinese learn "shui," the Japanese "mizu," the Spanish "agua," and so on.

that define the "things" signified by words? Are they somehow inherent in the human structure, the mental equivalent of teeth, tongue, and lips? Social constructionists argue that, no, the distinctions are socially determined. Moreover, the distinctions are not simply random but in fact shape hierarchies of value and corresponding distributions of power within the culture. Hence, the argument goes, if we can become aware of the systemic character of these distinctions, we are free to reflect on their implications and, if desired, reject or restructure them. This approach is the polar opposite of biological determinism.

The perspective derived from developments in neuroscience is, I believe, somewhere in between. In the end, however, the choice between biological determinism or social constructionism dissolves as simply too crude to be very useful. What DNA gives us is protein synthesis that in turn shapes neural structures and the physiology of response to neurotransmitters, neurotrophins, and neuromodulators. DNA gives us the sequences for development and the patterns of brain anatomy and large-scale connectivity. These factors, when confronted with the regularity of patterns in the body and external world, largely guide the remarkable process of the self-assembling of the brain and human experience. Tracing these patterns can, I think, fundamentally transform one's perspective and make it more closely attentive to the contours of human experience. This shift in turn will alter the terms of debate about the structuring of meaning.[3]

---

[3] This revision of the debate is the goal of *Rethinking Innateness: A Connectionist Perspective on Development* (Cambridge, MA: MIT Press, 1996), jointly authored by Jeffrey

Wetware

In a fetus, neurons build an initial cortical structure by migrating from a "proliferative zone" along fibers left by glial cells. They are smarter than artificial neurons, with greater (genetic) programming. Being biological, real neurons (within the usual constraints) are vastly more complex than the artificial models. Nerve cells have to metabolize; they have cell membranes with ion pumps and neuropeptide receptors. The systemic logic and emergent properties of neural nets as formal structures are extremely important to our understanding of the mind—indeed, a major breakthrough—but the biological character of the actual system is equally important. This chapter explores the biological embodiment, the structures by which the brain both processes peripheral sensory input and generates (hopefully appropriate) responses. Formal issues—how neural nets are trained—will remain the focus, but I shall also stress that the particularity of the biological system matters profoundly.[4]

In this account of the neuronal structures that underlie our day-to-day experience of the world, I begin with the fundamental unit, the neuron. I then move to more complex networks of neurons that make the higher-order ways of organizing

---

L. Elman, Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Dominico Parisi, and Kim Plunkett. They argue precisely that connectionism allows us to rethink *what* is innate in the human mind, and what processes produce these innate traits.

[4] I shall not attempt to give a complete portrait of the current state of our understanding of the physiology and structure of the brain. There are many fine resources that I shall not duplicate. For example, in the section that follows, I shall introduce the larger structural features of the brain via a developmental schema that I borrow from *Rethinking Innateness*.

our encounter with the world possible. Using vision as a model perceptual system, I race the flow of sensory information from the retina through the layers of neurons in the thalamus and primary visual cortex that learn to extract the basic feature data. As Chapter Two stressed, however, even simple perception with seemingly no cognitive content is in fact a layered series of two-way negotiations. The primary visual cortex shares its simple feature information with the higher visual cortices. These networks extract information about shape, color, location, and movement, which they share with visual object-memory. The object system in turn engages higher order semantic and episodic memory structures. All these networks join in conversation just to make sense of a ball on the floor. However, what makes those networks and those conversations possible are the basic properties of neurons whose patterns of firing shape the organization of synaptic connections and network architecture.

Neurons: the Basics of Connectivity and Communication

Neurons create the electrochemical signaling system not only within the brain but also between the brain and the body. They tell the muscles to contract and report information from the sensors on the surface; they process, store, and retrieve information as well as direct the actions they innervate at the periphery.

The signals of this system—the basic units of information—are waves of depolarization (discussed below) that travel along the cell membrane of the neuron

---

Almost all accounts of connectionism and/or neuroscience written for a broad audience have

and stimulate it to release chemical messages at the synaptic junction formed by the tips of the axons and the dendrites of adjacent neurons. Polarization here is the build up of charge, an imbalance between electrons (or anions, negatively charged ions) and protons (or cations, positively charged ions). In neurons specifically, the imbalance is a net negative charge inside the cell and a positive charge outside. The neuron creates this polarization by constantly running an "ion pump" on the cell membrane that moves sodium (Na^+) ions *out* of the cell and moves potassium (K^+) *into* the cell. Since Na^+ and K^+ both are positive ions with a charge of +1, the exchange would have no net effect, *except* that K^+ passes through the cell membrane far more easily. Because the concentration of K^+ is significantly higher inside the cell than outside, K^+ ions diffuse outward. Each ion that leaves contributes to creating a net negative charge inside and net positive charge outside. As the diffusion continues, the ion pump's action and the electrical attraction due to the growing polarization come to exactly balance the rate of diffusion due to the high K^+ concentration. The level of polarization at this balancing point determines the *resting potential*—a measure of the difference in electrical charge—of the neuron's cell membrane.

The two major ways to activate *depolarization* are via receptor potential and synaptic potential. In the first case, various sensory neurons on the surface of the skin, the retina, cochlea of the ear, etc. react to physical stimuli to open Na^+ channels (or close K^+ channels): stopping the flow of K^+ outward or allowing the flow of Na^+

---

an obligatory chapter on the brain that can serve as a good overview.

into the cell decreases the negative charge inside the cell and is therefore depolarizing. Changes in *synaptic potential* arise when neurotransmitters bind to receptor proteins at the synaptic junction. Some neurotransmitters cause Na^+ (or calcium, Ca^{2+}) channels to open, letting Na^+ or Ca^{2+} ions flow in and contribute to depolarization: they are excitory. In contrast, some *inhibitory* neurotransmitters increase the permeability of the cell membrane to K^+, letting it flow out and thus causing *hyperpolarization*, an increase in the cumulative negative charge inside the cell. (Just to keep the story complex, GABA (gamma-aminobutyric acid), an important inhibitory neurotransmitter, increases the permeability of the cell membrane to chlorine ions, Cl^-, which have a high concentration outside the cell. This produces a net *influx* of negative ions into the cell that increases the polarization and is therefore inhibitory.) Both receptor and synaptic potentials are *additive*: gently pressing a sensory neuron on a finger, for example, causes a slight depolarization; pressing it more produces greater depolarization. Similarly, twice as many neurotransmitters binding to neighboring receptor sites will produce twice the effect.[5] If both an excitory and an inhibitory transmitter bind in close proximity at roughly the same time, they cancel each other out.

Local depolarization of the cell membrane causes nearby Na^+ channels to open and increase the depolarization. If the depolarization reaches a certain

---

[5] This is only roughly true, and there are no doubt many variations. For example, there are bipolar cells in the brain that are tuned to detect simultaneous input at a given frequency from each ear. The response to simultaneous binding on each of the two dendritic branches is far greater than two signals on the same branch.

threshold, it stimulates a cascade effect where more and more Na^+ channels open. Na^+ rushes in and produces a spike. This rapid local depolarization produces a well-defined change in voltage across the cell membrane called the *action potential* and strongly affects the nearby Na^+ channels, causing them to reach the threshold and also spike. Since the action potential quickly dissipates as slower-acting K^+ channels also open and counterbalance the Na^+ influx, the net result is a wave of depolarizing action potential that travels down the cell membrane from the dentritic synapses to the axonal terminals and their synapses. The action potential spike causes the cell to release neurotransmitters into the axonal synaptic junctions, where it binds with the receptors on the dendrites of other neurons. The cycle of response then repeats itself. Since the action potential in neurons remains largely a fixed value, however, the *size* of the spike cannot represent the neuron's level of activation. Instead, activation is measured by *spiking rate*, the rate at which a neuron cycles through the process of depolarization and recovery.

In addition to neurotransmitters that directly influence the polarization at the synapses (the *ionotropic*), some neurotransmitters like serotonin are *metabotropic*. They act indirectly by creating a chain of reactions with molecules active inside the neuron that eventually influence the local membrane potential. These reactions usually are much slower but also much more long-lasting than those produced by direct-acting neurotransmitters. Some neurotransmitters, like glutamate, have both types of receptors. Neurons in the brain also release a host of other neuromodulators, chemicals that influence the dynamics of membrane potential at the synapse. Nitric oxide is pervasive, as are various neuropeptides, small chains of

amino acids. They can influence such factors as the sensitivity of a channel to neurotransmitters, which ion the channel works with, and the length of time the channel remains open. Recent research has discovered many neurotransmitters whose precise functions remain largely unclear and are the subject of intensive study.

Synaptic Modification: How Neurons Build Networks

Neurons are important not because they create waves of depolarization but because they transmit this spiking information to other neurons through synaptic junctions. However, they do not simply released their neurotransmitters into the synaptic junction for the receptors on the postsynaptic dendrites to detect: they change as they fire. Chapter Two discussed the centrality of the Hebbian learning rule for changing connection strengths in artificial neurons. For functioning networks to evolve, weights must be able to change according to the general law: units that fire together wire together. The basic assumption about real neurons is also that learning occurs through activity-dependent synaptic plasticity. Neurons, that is, must change their behavior depending on the events that lead to their reaching the activation threshold.

Activity-dependent synaptic plasticity comes via several different mechanisms. Increasing the number of receptors, increasing the sensitivity of the channels, changing the physical structure of the synaptic junction (or deleting it altogether), and turning on the expression of genes in the cell nucleus all seem to participate. For example, a combination of synaptic pruning (the deletion of inactive synapses) and the reorganizing of receptor sites seems to play a central role in shaping the

connections between the retina and the lateral geniculate nucleus (LGN) in the thalamus. At birth, each cell in the LGN receives input from about twenty retinal ganglion cells, but the connections are weak. After three weeks, the number of input source cells drops to less than four, but the connections become much stronger. The decrease in input sources comes through eliminating connections. The increase in connection strength, however, is a bit more complex and relies on a property of glutamate receptors worth stressing because it seems to be an important mechanism for increasing synaptic connection strength throughout the cortex.[6]

Glutamate receptors can be either ionotropic or metabotropic; there are, moreover, three major types of ionotropic receptors. One (the AMPAR) has a high affinity for glutamate (i.e. it binds strongly) and will open its Na^+ channel when the neuron is at normal resting potential. Another type (the NMDAR), however, both has a weaker affinity and does not activate its channel unless the neuron already has reached its activation threshold and has depolarized the region around the receptor. These characteristics create what are called "silent synapses," synapses that have NMDA receptors (and perhaps inactive AMPA receptors) that do not function until the neuron fires at the same time that glutamate is released into the synaptic junction.[7] When the NMDA receptor opens its channel, it seems to cause AMPA

---

[6] For this account of the developmental changes I rely on chinfei chen and Wade G. Regehr, "Developmental Remodeling of the Retinogeniculate Synapse" *Neuron* 28 (Dec. 2000):955-66.

[7] The discussion by Roberto Malinow et al. provides a very good overview, even though they discuss the hippocampus rather than the LGN. Roberto Malinow, Zachary F. Mainen, and Yasunori Hayashi, "LTP mechanisms: from silence to four-lane traffic" *Current Opnion in*

receptors to attach to the same site, thus making the synaptic junction responsive even without simultaneous depolarization. This mechanism is a form of coincidence detector that provides an elegant implementation of the "fire together/wire together" rule: silent synapses of a spiking neuron become active at all sites where the presynaptic neurons also have just fired. The effect is a form of *long term potentiation* that heightens the responsiveness of a neuron and provides a neuronal correlate to changing the weight matrix in artificial networks.

Changing the actual shape of the neurons through pruning synapses is one way the brain makes early, permanent adjustments to neural organization. The transformation of silent synapses into active ones also is important and not so drastic, but are the changes equally permanent? The consensus is that the LTP mediated by activation on NMDA receptors relies not only on changes at the receptor site but also on changes in gene expression in the neuron's nucleus. Exactly how this happens remains unclear but is the focus of intensive research because it is crucial to understanding not only how memories are formed but also how they decay. Moreover, these general properties of neurons are applicable throughout the brain: long-term potentiation and long-term depression are the bases for cortical self-organizing from the earliest stages of visual processing to the most elaborate and

---

Neurobiology 10(2000):352-57. For an excellent account of NMDA receptors, see Jan Pláteník, Nobuyuki Kuramoto, and Yukio Yoneda, "Molecular mechanisms associated with long-term consolidation of the NMDA signals," *Life Sciences* 67(2000):335-64.

subtle forms of reverie.[8] Neurons not only spike, they change: layer after layer, year after year, they grow into the beings who we are.

The Organization of the Brain

Early in fetal development—between the third and fourth weeks—a region of the outer layer of skin on the back (dorsal) part of the fetus begins to fold in on itself. The cells form a long hollow cylinder called the neural tube. In the sixth week, three lumps appear in the front (anterior) region of the tube. These lumps gradually form the forebrain, midbrain, and hindbrain. The rear (posterior) section becomes the spinal chord. This sequential development—though with different timings—is as true for fish, frog, and rats as it is for us. In humans, the forebrain in particular begins to grow rapidly. The best-know component of the forebrain is the cerebral cortex. As it grows, the cortex begins to fold in a complex but well-structured manner. Although our usual image of the "grey matter" of the cortex (that covers most of the rest of the brain) is of a rather massive lump of tissue, in fact the neocortex is only about 3-4 mm (0.12-0.16 inches) thick. The folding—produced by the right crimps appearing at the right time as the brain grows—manages to get about 2,400 cm^2 (2.5 square feet) of cortex inside the average adult skull and

---

[8] This discussion of synaptic connections argues that neural systems have the properties required to implement Hebbian rules for changing synaptic weights. Although biological systems surely depart from simple network models, the basic approach continues to find support from the available evidence. For reviews of the issues, see S. J. Martin, P. D. Grimwood, and R. G. M. Morris, "Synaptic Plasticity and Memory: An Evaluation of the Hypothesis" in *Annual Review of Neuroscience* 23(2000):649-711, and Guo-qiang Bi and

produces the impression of massiveness. The cortex rests on top of a complex set of subcortical cerebral structures like the corpus callosum, thalamus, hypothalamus, hippocampus, etc. The midbrain in humans is a set of structures that sits between the cerebrum and the pons (literally, "bridge") that connects the two halves of the cerebellum ("little brain," part of the hindbrain). The hindbrain consists of the pons, the two halves of the cerebellum, the medulla, and a few other structures.[9]

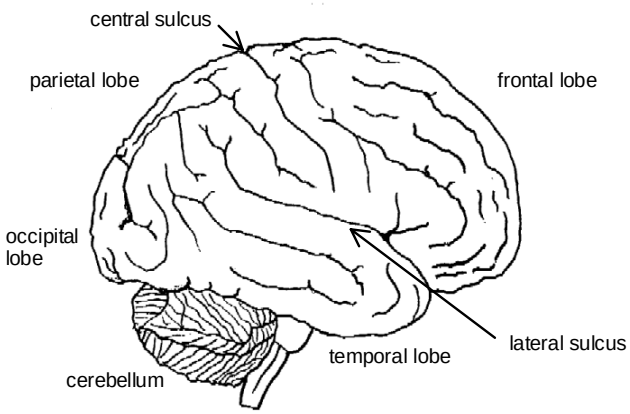
The thin sheet of cells of the cerebral cortex has six layers. One peculiar and extremely important feature of this cortical layering is that the cells develop in an inside-out manner. The neurons of layers 5 and 6 migrate along trails of glial cells first, then layers 4, 3 and 2. Maturation, i.e. arborization and myelation, also follows this order. The outermost Layer 1 has few neuron cells of its own and instead seems to serve primarily as a site for interconnections among neurons from other cortical regions. Layers 2 and 3 have both horizontal connections like Layer 1 and project below to other regions. Layer 4 is where most input connections end. The innermost Layers 5 and 6 contain the major output connections to subcortical regions and also participate in larger cortical pathways.[10]

Mu-ming Poo, "Synaptic Modification by Correlated Activity: Hebb's Postulate Revisited" in *Annual Review of Neuroscience* 24(2001):139-66.

[9] Although recent research has revealed the cerebellum to be more important in issues of learning than had been previously thought, I largely leave it out of this overview.

[10] Many books provide an overview of these structures. For an account that centers on the cognitive importance of developmental issues here, see Mark H. Johnson, *Development Cognitive Neuroscience: an Introduction* (Oxford and Cambridge, MA: Blackwell, 1997), pp. 24-39.

The sheet of cortex folds into two hemispheres connected in the middle by the corpus callosum. Anatomists traditionally divide each hemisphere into four lobes: the frontal, parietal, occipital and temporal. Major sulci (fissures) mark off several of these lobes. In particular, the central sulcus separates the frontal and parietal lobes, and the lateral sulcus separates the parietal and temporal lobes.



Our current understanding of the relationship between large-scale cortical structure and function comes from three sources: lesion studies, direct electrical probing of the brain, and more recently, from various forms of brain imaging.[11] Over the years, neurologists have accumulated a wealth of detailed information on the behavior of patients with brain injuries. These data allowed researchers to draw approximate functional mappings. If the loss of a section of the brain led to a loss of

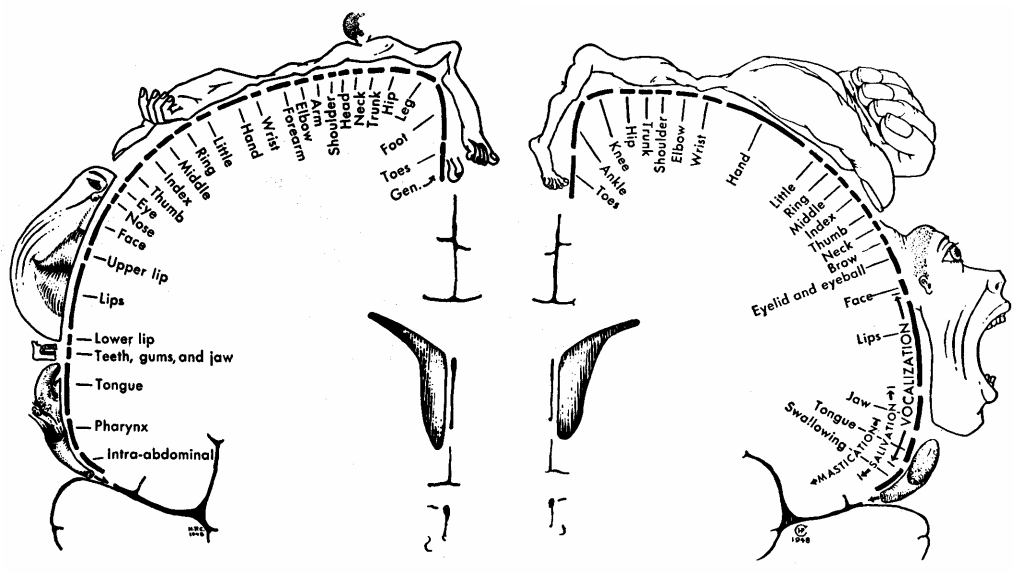
---

[11] For a good perspective on the history of understanding brain function, see Edwin Clark and Kenneth Dewhurst, *An Illustrated History of Brain Function* (San Francisco: Norman, 1996), 2nd edition.

previous abilities—the ability to read, for example—then surely, the argument went, that brain area participated in the systems that made reading possible. There were doubts about this simple correspondence, and researchers were aware of cortical plasticity. Still, the account was plausible and reasonably well supported by the preponderance of the data. Furthermore, the work of Penfield and Rasmussen, reported in 1957, tended to confirm these findings. They used electrical current to directly stimulate brain tissue of conscious patients and had them report the responses. They probed the regions on either side of the central sulcus and confirmed that the ridge in front of the sulcus is the primary motor cortex, and that behind the sulcus is the primary sensory cortex. Working their way from one end of these regions to the other, they determined what neurons caused what part of the body to either twitch (primary motor) or report sensation (primary sensory). They created the famous "homunculus," a drawing that illustrates the relative sizes of the neural real estate devoted to each part of the body.[12]

---

[12] From Wilder Penfield and T. Rasmussen, *The Cerebral Cortex of Man: A Clinical Study of Localization of Function* (New York: Macmillan, 1957), Figures 17 and 22, on pp. 44 and 57, as printed in Clark and Dewhurst, *An Illustrated History of Brain Function*, p. 131



Sensory Homunculus Motor Homunculus

Similar techniques explored other areas of the cortex where previous lesion studies identified functional mappings. The results confirmed the earlier data and held out the possibility of developing a detailed understanding of the organization of the cortex. As technique improved and more studies accumulated, however, this elegantly simple picture began to unravel. Detailed analyses of lesions in a large population of patients gave inconsistent results. Neurologists developed staining techniques that allowed them to trace what neurons communicated to what regions. It turned out that the direct interconnectivity between regions was much broader than the simple picture of localized functions would allow. Moreover, studies in transcranial stimulation began to confirm the implications of the neuroanatomical data. The most recent development has been brain imaging through CAT scans, positron emission tomography (PET), and various forms of nuclear magnetic resonance imaging (MRI). These techniques are still works in progress, but they tend

to repeat the disquieting results. When experimental subjects are given a task—reading, hearing, or thinking about saying a word, for example—the appropriate areas light up, but so do others. In the last paragraph of *An Illustrated History of Brain Function*, the authors conclude:

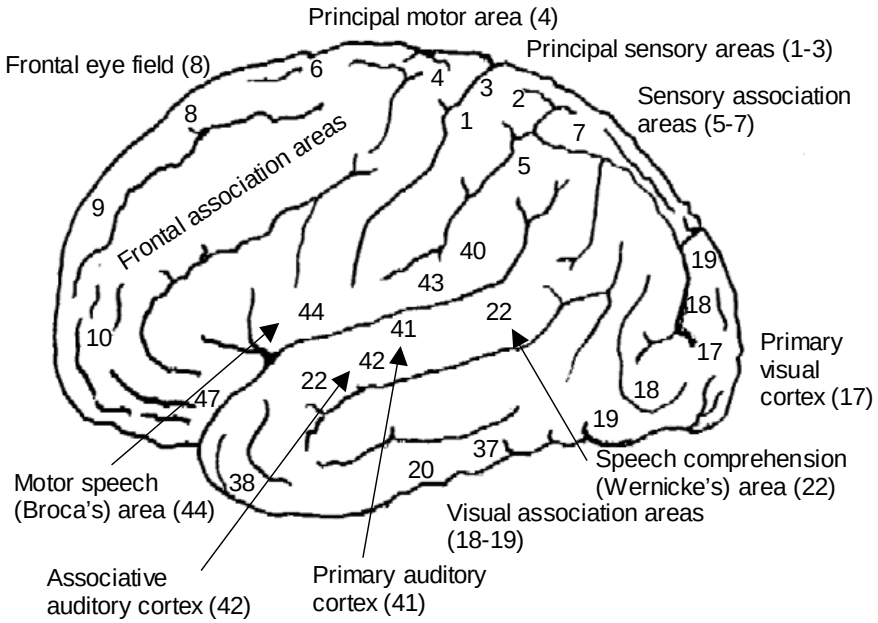
> It is becoming increasingly clear that the operation of the nervous system during the performance of a particular function depends on the pattern of excitation and inhibition in neuronal populations that are widely dispersed in different regions of the cerebral cortex. The belief that distinct functions are restricted to distinct cortical regions, as envisioned in the earlier part of this century, has become increasingly untenable. Recent work has emphasized the plasticity of the cortex, which is not simply a hard-wired array of neural circuits but a dynamic system that is modified or reorganized by preceding activity and feedback, and the importance of parallel distributed cortical networks with mutual reciprocal interconnections between different cortical areas.[13]

Given the above caveats, what then *is* the standard first-order account? The map below, adapted from a recent medical textbook on neuroanatomy, is representative.[14]

---

[13] Clark and Dewhurst, *An Illustrated History of Brain Function*, p. 170.

[14] The numbers are (approximately) the traditional Brodmann area designations. Adapted from Stephen G. Waxman, *Correlative Neuroanatomy* (Stamford, CT: Appleton & Lange, 1996), p. 148.



The prefrontal cortex usually is associated with long-term planning and with the integration of information from the various senses. The *frontal lobe* maps to movement control. As mentioned above, the primary motor area (Brodmann area 4) is along the central sulcus, and Broca's area (area 44) serves in the physical control of speech. The *parietal lobe* is dominated by the primary sensory area. The *occipital lobe* primarily serves in visual processing but also contains Wernicke's area (Brodmann area 22), which is important in the understanding of language. Finally, the *temporal lobe* contains the primary auditory cortex and other regions that help process hearing and language.

This sort of list, however, does not convey much sense of the actual relations between systems within the brain or between the brain and the rest of the body. How is the sensory stream carried into the brain? How is the information in that

stream structured and then integrated into higher cortical processes for evaluation and response? Exploring the visual system, which has been intensely studied, can help set out current working hypotheses about these relationships and also suggest the value of neural net modeling in understanding and extending the empirical data.[15]

Seeing the Light

Vision is our window to the world. The eyes report the real: open them, see a ball rolling toward you, reach for it. The rubber ball is smooth but slightly tacky with the summer heat. Unfortunately, there is no little person looking out from within the eyelids, and the direct access—even if an everyday occurrence—is illusory. The issue here is of great significance: our sense of our self, our knowledge, and our world hangs upon whatever we develop to replace the simple story.

The revised version will have to begin with the sensory neurons of the retina. The elegance of design discovered there may serve as a reassuring introduction.[16] The retina consists of three layers of cells: ganglion cells, a middle layer (amacrine, bipolar, and horizontal cells), and rods and cones. The rods and cones are light-

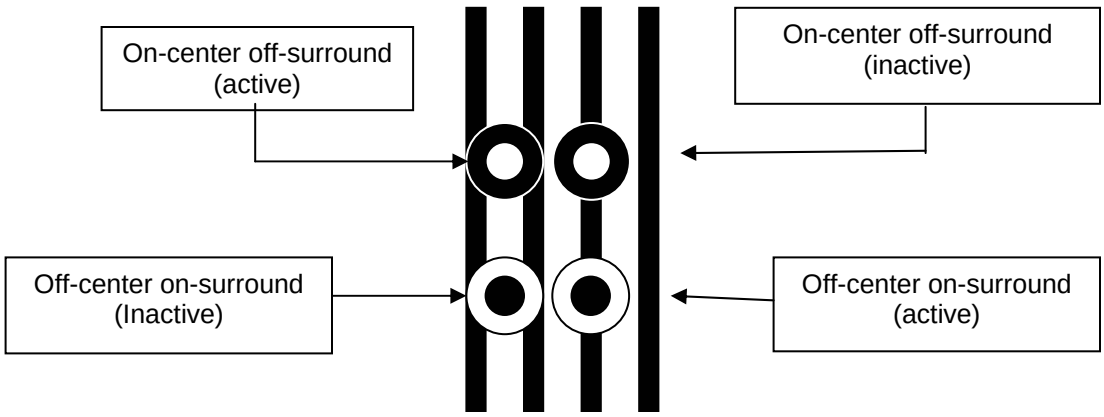
---

[15] My strategy here is essentially the same as both Patricia Churchland and Terence Sejnowski in *The Computational Brain* and Mark H. Johnson in *Developmental Cognitive Psychology*. The visual system provides a good introduction to the logic of sensory processing and its eventual integration into higher order cognitive systems.

[16] I base this account on David Hubel's *Eye, Brain, and Vision* (New York: Scientific American Library, 1995), pp. 36-57, and John E. Dowling *Creating Mind* (New York: Norton, 1998), pp. 101-22.

sensitive. Rods, which are far more numerous, respond to dim daylight but quickly become saturated in bright. Cones, in contrast, do not detect dim light well but are color sensitive and make it possible for us to see fine detail. There is a small spot at the back of the eye where the retina is all cones and no rods. This is the *fovea*, where vision is most acute. In the order of cells, however, rods and cones are *last*: two layers of cells (plus two layers of axons and dendrites) stand between the light and the detectors. This odd arrangement is true everywhere except at the fovea, where the mediating layers are displaced and give the cones direct access. As a result, the fovea appears as a tiny pit in the back of the eye, slightly above the center.

The rods have the greatest response to green light but do not participate in encoding color. Cone cells have three subgroups that respond—in a rather broad tuning curve—to different frequencies. For some the maximal response (the highest point of the tuning curve) is in the red-yellow (565nanometers) region; some respond maximally to green (535 nm), and some to blue (440 nm). Both rods and cones transmit their activation to bipolar cells, but the response patterns of the bipolar cells reveal that a first layer of processing has occurred. Bipolar cells are of two types: on-center and off-center. An on-center bipolar cell spikes actively if light shines on the retinal receptors in the center of its "receptive field," and no light shines on the receptors around the edge. Off-center cells are just the opposite. Both types of cells, moreover, spike slowly in darkness and a bit more rapidly when both the center and surround cells are illuminated. Bipolar cells measure *relative* rather than absolute brightness, which turns out to be both a form of data compression



Representative Ganglion Cells' Response to a Grid Pattern

and a way to assure decent acuity over a relatively wide range of lighting conditions. This sort of "center-surround" receptive field is created by the inhibitory role of the large horizontal cells that span the receptors included in the bipolar cell's receptor field. It is important to note that simple Hebbian training of such a system of receptors, horizontal cells, and bipolar cells with mutual inhibition easily produces this pattern of activation. It turns out, however, that amacrine cells contribute yet another layer of processing within a subset of cells at the next layer. Amacrine cells receive information from bipolar cells but respond strongly only to *changes* in light.[17] Connected in a network to a subpopulation of ganglion cells, they train those ganglion cells through mutual inhibition to respond not to steady brightness but to changes in intensity.

Thus the patterns of spiking transmitted through the optic nerve carry information about relative brightness, color, and change over time that preserves the two-dimensional spatial mapping of the surface of the retina. This information next goes to a section of the thalamus called the lateral geniculate nucleus (LGN), which then feeds its responses to the primary visual cortex, V1. However, there are about ten times as many downstream connections from V1 to the LGN than there are going up from the LGN to the visual cortex. What is the point of this massive recurrent network? What is the LGN doing? The answer, unfortunately, remains unclear. From recent work, the LGN seems to serve several functions, and others probably will appear.

First, the LGN appears to control selective visual attention. Research from lesion studies and the like has long acknowledged that the thalamus plays an important role in controlling attention. The question has become what the neuronal correlates are. Here, as elsewhere in neuroscience, the interplay between mathematical neural modeling and biological research has proven very powerful. The basic idea of recurrent networks presented the possibility of solving some fundamental problems in vision research, but early models were simple and not very biologically plausible. Detailed information about neural structure, connection patterns, and spiking behavior has prompted the creation of more sophisticated models. These models in turn make predictions about behavior that can be tested.

---

[17] The retinal system still is not entirely understood, but I rely here on findings reported in Leon Lagnado, "Retinal Processing: Amacrine cells keep it short and sweet," *Current*

The recurrent connections to the LGN have proven a particularly inviting object for modeling because both the retina and V1, the primary visual cortex, have been well studied. The modeling of attention—where one region of the visual field is enhanced and the others are suppressed—is one result of this collaboration: simulations suggested that yet another part of the thalamus, the thalamic reticular nucleus, was part of the inhibition circuit for LGN and V1, and recent physiological evidence supports the existence of the necessary connections.[18] Other studies have proposed yet another function for the LGN: predictive coding of naturally occurring visual patterns. The idea is that more frequently seen patterns excite less attention and can be coded more simply because they have become part of the patterning of the neural connections themselves. This saves "band-width" for conveying details about unusual input. Although this logic appears most clearly higher in the visual processing stream, recent work suggests that the LGN already begins this transformation. Again, what makes this predictive coding possible is the feedback connection from the primary visual cortex.[19]

Biology 1998, 8:R598-R600.

[18] Here I draw upon the work of John Bickle et al., "A Functional Hypothesis for LGN-V1-TRN Connectivities Suggested by Computer Simulation," *Journal of Computational Neuroscience*, 6 (1999):251-261. In a recent study, S. Murray Sherman has correlated the two response mode—tonic or burst firing—of the LGN with attentional states. He suggests, "Tonic firing minimizes distortions in the relay, thereby supporting a more faithful recosntruction of the visual world. Burst mode maximizes initial stimulus detection, perhaps as a sort of 'wake-up call' that something has changed in the environment." See S. Murray Sherman, "Tonic and burst firing: dual modes of thalamocortical relay," in *Trends in Neuroscience* 24.2 (Feb. 2001):122-26.

[19] See Rajesh P. N. Rao and Dana H. Ballard, "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive field effects," *Nature* Neuroscience v. 2, no. 1 (Jan. 1999):79-87.

The basic idea of recurrent loops has proven extremely powerful in modeling information processing throughout the visual system. For example, Daniel Pollen writes:

> Evolutionary pressures to develop such feedback loops have probably been based in part upon the need of an organism to discriminate and interpret sensory data on the basis of its past experience and motivational state (Pandya and Yeterian, 1995). At the basic level of object recognition, Grossberg (1994) and Mumford (1994) recognized that, on the one hand, it is difficult to segregate an object from background without prior recognition of that object. On the other hand, recognition often follows only after the representation of the object has been segregated from its background and its boundaries defined. Thus, figure-ground segregation and object recognition cannot progress in a simple bottom-up serial fashion, but have to occur concurrently and interactively within recursive loops....

Pollen then extends the model (I have omitted the citations):

> Thus, such loops have been proposed to employ active use of higher-level knowledge to disambiguate lower-order percepts, to mediate the play of selective attention upon early image representation, to correlate and synchronize the activity of interrelated groups thereby facilitating the continual updating of the perceptual image, to permit parallel exploration and selection of multiple alternatives, to facilitate binocular fusion by suppressing non-corresponding retinal images in the LGN, to provide spatial 'shifter circuits' for the computation of fine stereo vision and disparity hyperacuity, to pre-attentively separate figure from ground, to modulate cortical output across cortical areas, to sustain 'temporal buffering' when there must be integration of clues otherwise 'hidden' over immediately preceding and succeeding spatial or temporal events, and to mediate contrast gain of LGN neurons.[20]

That is, predictive loops from higher to lower levels are probably part of a fundamental strategy of processing not merely in the visual system but throughout

the cortex. Such predictions are only possible, however, if the system extracts the patterns to predict. We already have seen tuning to basic patterns in the retina: movement, dark center/bright surround, and bright center/dark surround. The primary visual cortex has cells to effect a next level of aggregation. So-called "simple cells" respond to the angle of a line segment based on information from a cluster of retinal ganglion cells. Some spike for horizontal bars; some prefer 72°; other react to 45° and so on. They use a system of "coarse coding" where there are not cells to represent all possible angles. Instead, since the response of the cells is graded—falling off quickly for angles above and below the optimal angle for the particular cell, but still responding a little—linear combinations of the responses can represent any possible orientation. A similar logic is true for the response curves of the cones: output from cones broadly tuned to three wavelengths is capable of representing all the colors of the visible spectrum. This retinal response is built in to the biophysics. In contrast, the organization of the simple cells can be created from a homogeneous network of neurons through simple Hebbian learning. They are self-assembling feature detectors. Since the retina also produces ganglion cells that respond to changes in light intensity, there are corresponding movement detectors in V1. There are, moreover, "complex cells" that discriminate combinations of features abstracted from the simple cells. Some, for example, combine movement detection with line orientation. Among the complex cells (of layers 2 and 3 in particular), there also

---

[20] Daniel A. Pollen, "On the Neural Correlates of Visual Perception," *Cerebral Cortex* 9.1 (Jan.-Feb. 1999):9-10.

seems to be a population that uses the simple-cell responses to develop a coarse coding for representative composite features of the natural world as interpreted by the retina. This organization of simple-cells and ever higher-order complex-cells looks very much like the distribution of increasingly complex response patterns produced by introducing a wave of plasticity into the training of simple perceptrons discussed in Chapter 2. The inside-out maturation of the layers of the cortex—seemingly no more than a curious fact—may create precisely this logic of development. Again the confluence of neural modeling and biological data produces unexpected but fascinating possibilities.

One other feature of the primary visual cortex suggests the surprising behavior that can arise from self-assembling systems that confront streams of data from an implicitly structured environment.[21] We have, of course, two eyes rather than one, and depth perception and stereoscopic vision are important for humans. The information sent from each retina, however, corresponds to its two-dimensional surface of receptors. The processing that makes stereopsis possible begins in the lateral geniculate nucleus (LGN) of the thalamus, which has alternating rows of cells responding to each eye. In the primary visual cortex, this organization is more clearly developed: the arrays of simple-cell angle detectors are arranged in alternating rows for input from each eye. (Once more, simple competitive training can produce this result.) Moreover, there are complex-cells that build upon this

organization and become disparity detectors. some are "far-cells" which fire when the corresponding feature from one eye is behind the other. Some are "near cells," and some are tuned to precise convergence. To complete the picture, some of these complex cells only fire when there is movement as well.

The primary visual cortex begins organizing the input stream into larger units of information. V1 connects in a very orderly manner to V2, in which color, form, and motion appear to be processed separately. V2 in turn connects to V3, V4 and the medio-temporal area (MT, also referred to as V5), all of which have more specialized functions. V3 processes information about form; V4 handles color, and the MT analyzes movement. The higher levels create neurons that respond to increasingly complicated patterns even as they dissociate sensory aspects into different streams.

Top-down and Bottom-up, or Interactions Everywhere

The coding structures of the visual cortex are not an autonomous sensory processing module. One half of the story—the part presented so far—has been the reliance of these structures on developmental factors to shape them into ever more complex populations of feature detectors. The other half, however, are the top-down recurrent activations that participate in training the feature-detectors, control

---

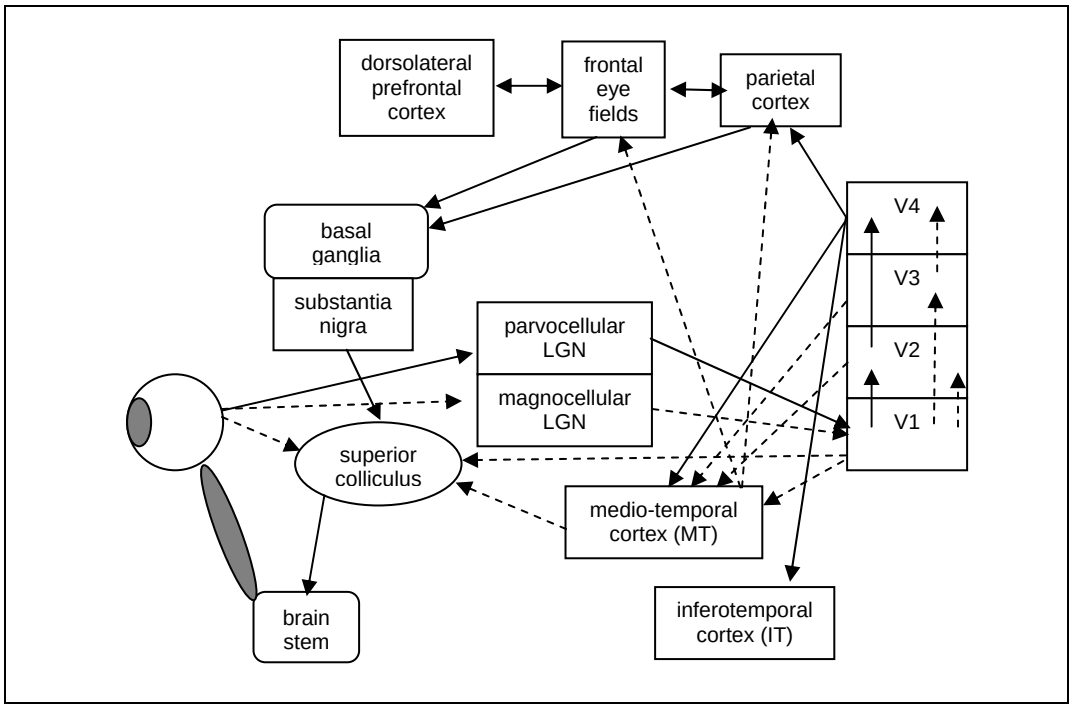
[21] There is a self-consistent if circular logic here. We consider it quite remarkable that neural networks, including the eye, can be trained to extract the sorts of regularities we detect in the world. We know nothing of the regularities we cannot detect.

attention within the visual field, and seemingly underlie what we experience as visual consciousness.[22] Attention and consciousness introduce larger questions of memory, recall, action planning, language, objects, and other such "upstream" phenomena. Yet, as mentioned in discussing the LGN, they also reach down into all the subsystems for visual processing.

To give one example of these interactions between levels, consider the control of tracking as the eyes follow a red ball bouncing across the floor. The seemingly simple operation requires a complex coordination of brain functions (see diagram).[23] The superior colliculus sends signals to the brain stem to control eye movement. The superior colliculus in turn receives inputs from three sources: the basal ganglia (via the substantia nigra), the medio-temporal cortex (MT), and the primary visual cortex (V1). The MT, which analyzes motion, receives input from all four visual areas and

---

[22] The mechanisms and implications of attentional modulation within the visual cortex are important sites of current research. For a good overview, see Sabine Kastner and Leslie G. Ungerleider, "Mechanisms of Visual Attention in the Human Cortex," *Annual Review of Neuroscience* 23(2000):315-41. For aspects of current research and theory, see, for example, Roy E. Crist, Wu Li and Charles D. Gilbert, "Learning to See: Experience and Attention in Primary Visual Cortex" (*Nature Neuroscience* 4.5 (May 2001):519-25), Carl R. Olson, "Object-based Vision and Attention in Primates" (*Current Opinion in Neurobiology* 11 (2001):191-79), Victor A. F. Lamme and Pieter R. Roelfsema, "The Distinct Modes of Vision Offered by Feedforward and Recurrent Processing" (*Trends in Neuroscience* 23 (2000):571-79), Victor A. F. Lamme et al., "The role of primary visual cortex (V1) in visual awareness" (*Vision Research* 40 (2000):1507-21), and Stefan Treve, "Neural correlates of attention in primate visual cortex" (*Trends in Neuroscience* 24.5 (May 2001):295-300). Lamme also has a more speculative article that echoes Grossberg's earlier proposal about the centrality of recurrent networks in awareness: "Neural Mechanisms of Visual Awareness: A Linking Proposition" (*Brain and Mind* 1 (2000):385-406).



generates output not only to the superior colliculus but also to the higher cortical regions: the parietal cortex (spatial analysis, planning of "eye-centered" movement), the frontal eye fields (in the frontal cortex, for anticipatory eye movement and volitional control), and the prefrontal cortex (makes time delays possible). These three regions in turn talk to one another and produce output to the basal ganglia, which seems to suppress eye movements. When an adult tracks a ball, the brain calculates where the ball will be, uses information about where the eye currently is focused, and moves the eye in an anticipatory manner so that the tracking becomes smooth. In infants, however, not all the elements are yet "on-line." As discussed above, the cortical layers develop in an inside-out manner. Dendritic arborization

---

[23] Adapted from Mark H. Johnson, *Developmental Cognitive Neuroscience* (Cambridge, MA: Blackwell, 1997), p. 78. I follow his discussion of the control of visual attention and eye

and myelation of the axons proceed from layers 5 and 6 up to layers 2-4. In the primary visual cortex, however, connections to the superior colliculus arise from layers 5 and 6; connection to the MT is from layer 4, and connection to the rest of the visual cortex starts from layers 2 and 3. Thus, in a new-born, control of eye movement involves just the loop from the eyes to V1 to the colliculus back to the eyes. Movement is very jerky. Soon connections from layer 4 to the MT mature, but there still is little information being passed to the parietal cortex. Movement becomes smoother but is still automatic. Only later, after connections to the rest of the visual cortex mature does voluntary control become possible. Yet where the infant looks and what it sees will strongly influence the initial shaping of the coding structures of the visual cortex, so eye movement control feeds back into the structuring of V1 as well.[24]

Red ball: the Transition from Perception to Memory

Looking at a system like visual tracking introduces many new terms into the discussion. In considering the visual cortex, we dealt with coding within feature domains. Processing tended toward splitting the visual information into separate components. However, there was no mention of "things" like balls or of such categories as volition. How do such entities become part of our account? Volition is part of the larger story of the structuring of creaturely values that is discussed in the

---

movement.

[24] For a recent overview of these processes, see John Colombo, "The Development of Visual Attention in Infancy," *Annual Review of Psychology* 52 (2001):337-67.

next chapter. The shiny red bouncing rubber ball, however, marks the transition from bottom-up perceptual coding to the creation of objects—and other enduring patterns of experience—identifiable through the mediation of memory. Reflecting on a red ball helps complete the presentation of visual processing and the highest levels of sensory coding strategies in general.

The ball can be approached in stages. How does vision identify a distinct entity? How do we know it is a ball? That is, what binds the visual image with the semantic category BALL? What *is* the semantic category BALL? How do we get from the semantic category to the word "ball," spoken or written? And if I know the ball belongs to Jamie, how do I know it? What do I actually know?

Cognitive scientists have been grappling with the neuronal logic of these higher order processes for over a century. The new techniques in imaging and the new mathematical models have contributed greatly to our understanding of the mental phenomena involved in "seeing Jamie's red ball" (object binding, semantic and episodic memory, the phonological lexicon, etc.), but the accounts remain very preliminary. Evidence detailing higher cortical functions is often contradictory, and it is not clear that our current techniques have the resolution needed to articulate these functions, but some tentative conclusions can suggest the broader consensus on cortical dynamics.

The recollection, "This is Jamie's red ball that he left here last night" is a form of *episodic* memory. Episodic memory is about events: it includes contextual information and memory about objects that unfolds over time. It also—quite remarkably, given Hebbian models for learning—is one-shot learning. The medial

temporal lobe (MTL) memory system has to record the event—its context as well as the temporal components—in one exposure: Jamie was playing with the ball in the living room (trying to get one of the cats to chase it), when it came time for him to go to bed. This conjunction of elements occurred only once, yet I remember it vividly. To quickly construct all the associations in memory between objects and contexts and between successive moments of the episode, the brain relies on the specialized hardware of the hippocampus.

Semantic memory, in contrast, is what the brain knows about *things* like "balls," "Jamie," "cats," "chase," and the color "red." What exactly these "things" in the brain *are* is not at all clear. For example, the binding of visual features identified by the bottom-up analyses of the various strata of the visual cortex to create a "shiny red round object" seems to occur in the inferotemporal cortex (IT). The IT relies on spatial information provided by the parietal cortex to assemble the "shiny red" and the "round" into a unitary "thing."[25] (The process may not seem especially worthy of remark, but it turns out that one can routinely force people in a rushed judgment with divided attention to shift color attributes, for example, from one object to another.) Yet this perceptual object is still just a complex shape with colors and textures. It becomes a unitary object (or perhaps several objects, like bubblegum stuck on a ball) in conversation with semantic memories about objects. The simplest

---

[25] See for example, Anne Triesman, "Feature binding, attention, and object perception", Robert Desimone, "Visual attention mediated by biased competition in extrastriate visual cortex," and John Duncan, "Converging levels of analysis in the cognitive neuroscience of visual attention," all in *Proceedings of the Royal Society of London*, B (1998).

semantic information about category—"it's a ball"—seems to reside in the temporal cortex. However, meaning ramifies very quickly and points to information from many different modalities that must somehow remain linked. A BALL is a THING, indeed, a NON-ANIMATE thing (a category revealed to be meaningful through lesion studies), but we do not know very well how the brain abstracts from all the instances of "balls" it has "seen" (baseballs, basketballs, tennis balls, golf balls, beach balls, and on and on) to create the semantic category BALL.[26]

Cognitive psychology—which approaches behavior at a conceptual level above issues of neural implementation—has studied concept formation in great detail, and at the level of brain-as-black-box. The questions posed by cognitive psychologists usually turn upon two major possibilities. Concepts—semantic categories—are perhaps an *averaged* prototype of all the exemplars in the class, or perhaps the prototype is actually a small *set* of typical exemplars. For example, in southern California we have Great Blue herons, ruby-throated hummingbirds, and red-tailed hawks along with robins, finches, crows and pelicans. The average of all these examples of birds would be a strange creature indeed. Not surprisingly, experiments with children indicate that typical examples (rather than the average) define the category. In fact, however, the issue is not between no abstraction from particulars or complete abstraction: a bird still is "a creature with a beak, two wings,

---

[26] Joaqin Fuster frankly states, "Clearly we are not quite ready yet to explore in depth the neural dynamics of episodic or semantic memory, let alone conceptual memory." Joaqin Fuster, "Cortical dynamics in memory," *International Journal of Psychophysiology* 35 (2000): 157.

two legs, and feathers." The process of concept formation seen in children seems closer to simply weak abstraction with simultaneous reference to exemplars.[27] Its neuronal corollary—the contents of semantic categories—translates this tension between specific instances and the processes of abstraction into questions about the processes underlying long-term memory and about the relationship of semantic and episodic memory in particular.

"Jamie rolling the red ball for the cat yesterday" remains a vivid component of episodic memory. The connections between the particular components of this memory are somehow temporarily stored in the cortex immediately next to the hippocampus (the entorhinal cortex in particular) while the fact of memory traces having been stored is registered in the highly recurrent excitatory connections of the CA3 region of the hippocampus. *What* is stored appears to be the patterns of activation of higher order cortical structures ("Jamie," "cat," "ball," etc.) that in turn have newly enhanced reciprocal connections to lower level sensory patterns. The sequences in time and space through which the "episode" unfolded provide components of the recollection's temporal and spatial specificity that may strongly point to the hippocampus' earlier evolutionary role in spatial memory: these sequential components seem to be moved from the hippocampus itself to the temporal and frontal cortex for transformation into long term memory.

---

[27] The literature on this topic is vast. Usha Goswami provides a highly accessible account of concept formation in children in *Cognition in Children* (Hove, East Sussex: Psychology Press, 1998), pp. 73-115. Jerry Fodor has issued a broadside against the entire approach in *Concepts: Where Cognitive Science Went Wrong* (New York : Oxford University Press, 1998), but I do not find his critique especially compelling.

How the hippocampus manages these recodings is unclear, and its precise role is a matter of debate.[28] In particular Endel Tulving, an important researcher in the organization of memory, has argued that episodic and semantic memory proceed by different pathways. In his view, only episodic memory is mediated by the hippocampus. Semantic memory, in contrast, relies on the perirhinal cortex, which nonetheless has dense connections with the hippocampus.[29] Since the nature of memory, and of semantic memory in particular, is central to my discussion, some of the issues in the debate about the locus and mechanism of memory creation and recall are worth exploring in greater detail. Three questions are the focus of current intense discussion. First, does the hippocampus merely facilitate the creation of episodic memory structures elsewhere in the brain? That is, once the synaptic connections necessary to store the memory have consolidated, is the hippocampus still necessary? Secondly, is the hippocampus needed at all for the creation of the synaptic connections that underlie semantic memory? Finally, what is the relationship between the two memory systems?

---

[28] For an overview of the debate, see the set of essays in *Hippocampus* 8 (1998): Endel Tulving and Hans J. Markowitsch, "Episodic and Declarative Memory: Role of the Hippocampus;" Larry R. Squire and Stuart M. Zola, "Episodic Memory, Semantic Memory, and Amnesia;" Mortimer Mishkim et al., "Amnesia and the Organization of the Hippocampal System."

[29] For a recent summary of the role of the perirhinal cortex is semantic memory, see Elizabeth A. Murray and Barry J. Richmond, "Role of perirhinal cortex in object perception, memory, and associations," *Current Opinion in Neurobiology* 11 (2001):188-93. Lynn Nadel and Veronique Bohbot present an overview of the argument for continuing involvement of the hippocampus in episodic memory in "Consolidation of Memory," *Hippocampus* 11 (2001):56-60. For a more extended account of their model, see L. Nadel, A. Samsonovich, L. Ryan, and M. Moscovitch, "Multiple Trace Theory of Human Memory: Computational,

This third question seems the least contentious: episodic memories are about objects, and objects require semantic memory. Conversely, components of the episodic memories involving an object become part of the semantic domain of the object. "Objects" come to exist—to become differentiated components (attractors) in the semantic representation space—because of the need to capture biologically significant patterns first presented through the sensory and proprioceptive data of episodic memory. The processes of bootstrapping (getting the semantic structure started) and structural redefinition that are part of the assimilation of episodic into semantic memory have received little attention of which I am aware even though almost all the contributors to the debates about memory seem to take very seriously McClelland's argument about the need for slow, interleaved consolidation of semantic memory to avoid catastrophic interference.

The basic distinction between episodic memory as essentially arbitrary associations of synaptic activity captured quickly by the hippocampus and semantic memory as the slow sifting of these associations for significant patterns offers at least a partial answer to the first two questions about episodic and semantic memory. No one doubts the need for the hippocampus in forming episodic memories. Its role as a pattern associator seems well established. The debate has been over what happens next. Do the patterns *stay* in the hippocampus? Central to the arguments has been the nature of retrograde amnesia (forgetting of *past* events, as opposed to *anterograde* amnesia, the quick forgetting of present events) when the hippocampus

---

Neuroimaging, and Neuropsychological Results," *Hippocampus* 10 (2000):

is damaged. Some researchers have asserted that memory loss is temporally graded, i.e. that older memories are better preserved than more recent.[30] This interpretation supports the idea that the hippocampus helps bind together synaptic connections in the cortex, so that the older connections between cortical regions, being better established, are less vulnerable than more recent memories. Other researchers have disputed these data and argued that more careful probing of the neurological patients' episodic memory reveals that old and more recent memories have suffered essentially the same degree of damage.[31] By this account, synaptic connections in the hippocampus are crucial for both the creation and the recall of episodic memories throughout ones lifetime. In the end, however, the outcome of this debate about the long-term role of the hippocampus, although certainly crucial for our larger understanding of the brain and memory, is not especially important for the themes of this book. The hippocampus is needed for the creation and consolidation of episodic memories. For the moment, that is enough.

The second half of the argument—that the hippocampus is *not* needed in semantic memory—is far more crucial because it seems to undermine McClelland's model for memory consolidation. However, here too the relevant facts and

---

352-68.

[30] For an overview of the literature that asserts this position, see Larry R. Squire, Robert E. Clark, and Barbara J. Knowlton, "Retrograde Amnesia," *Hippocampus* 11 (2001):50-55.

[31] See, for example, Indre V. Viskontas, Mary Pat McAndrews, and Morris Moscovitch, "Remote Episodic Memory Deficits in Patients with Unilateral Temporal Lobe Epilepsy and Excisions," *Journal of Neuroscience* 20.15 (Aug. 1, 2000):5853-57. For similar arguments deriving from rat and monkey lesion studies, see Elizabeth A. Murray and Timothy J. Bussey,

interpretation are in dispute. For example, there are cases of patients who suffered severe damage to the hippocampus as children but still have learned much about the world around them. Some researchers argue that their ability to acquire a semantic structure in the absence of the hippocampus and episodic memory shows that the hippocampus is not needed for the building of semantic networks.[32] Others have replied that, first, there remains at least some episodic memory in these patients and at the same time, their semantic structures are still significantly impaired. Secondly, the argument about the role of the hippocampus all along has been that it allows the sort of internal replay needed to integrate new information into the current semantic structure. Such a role for the hippocampus does not preclude the possibility that repeated encounters with semantic patterns *in the world* (i.e. the external sources for the high-level patterns captured by the hippocampus) can also train the necessary synaptic connections as long as the moment-to-moment memory structures of the entorhinal and perirhinal cortices remain intact.[33] Following this line of reasoning, some researchers have argued that without the hippocampus, the level of complexity of associations that can be linked in the perirhinal cortex is significantly diminished,

---

"Consolidation and the Medial Temporal Lobe Revisited: Methodological Considerations," *Hippocampus* 11 (2001):1-7.

[32] The article that began the debate about these patients is F. Vargha-Khadem, D. Gadian, K. E. Watkins, A. Connelly, W. Van Paesschen, and M. Mishkin, "Differential Effects of Early Hippocampal Pathology on Episodic and Semantic Memory," *Science* 277 (July 18, 1997):376-380. They restated their case in Mortimer Mishkin, Faraneh Vargha-Khadem, and David G. Gadian, "Amnesia and the Organization of the Hippocampal System," [Commentary.] *Hippocampus* 8 (1998):212-216

[33] Larry R. Squire and Stuart M. Zola make this argument in "Episodic Memory, Semantic Memory, and Amnesia" [Commentary.] *Hippocampus* 8 (1998):205-211.

but learning remains possible.[34] I confess I prefer this latter version of the hippocampus as recorder of associations between patterns of activations stored in the rhinal cortices that surround it.

I prefer this version because it fits the neurological data reasonably well, it corresponds neatly with the mathematical modeling of how the internal connections of the hippocampus work, and it accounts for research on the role of sleep in learning. For example, work by Michael Hasselmo on neuromodulator levels during the sleep cycle seems to find a function for sleep that supports McClelland's model for the role of the hippocampus in the slow replaying of new patterns to avoid catastrophic interference (discussed in Chapter 2). Hasselmo points out the changes in patterns of the neuromodulators acetylcholine, norepinephrine, and serotonin between waking activity, slow-wave sleep and REM (rapid eye movement) sleep and the impact these changes have on the flow of synaptic activity between the sensory cortex, the entorhinal cortex and the hippocampus. During waking activity, high levels of acetylcholine suppress feedback connections both within the CA3 region and between CA3, CA1 (the region of the hippocampus that mediates between CA3 and the entorhinal cortex), the entorhinal and the sensory cortex. During slow-wave

---

[34] See Rebecca D. Burwell and Howard Eichenbaum, "What's new in animal models of amnesia?" *Behavioral and Brain Sciences* 22:3 (1999)446-47. This report is a brief commentary that is part of a series of responses to the review article by John P. Aggleton and Malcolm W. Brown, "Episodic memory, amnesia, and the hippocampal–anterior thalamic axis," *Behavioral and Brain Sciences* 22.3 (1999):425-44. Eichenbaum gives an extended analysis of the "fused" representations made possible by the rhinal cortices in the absence of the hippocampus in contrast to the identity-preserving associations mediated by the hippocampus in "Declarative Memory: Insights from Cognitive Neurobiology," *Annual*

sleep, acetylcholine levels drop, and the hippocampus replays activation through its recurrent connections as a series of sharp waves originating in the CA3 region. During REM sleep, in contrast, acetylcholine levels become very high, but norepinephrine and serotonin levels drop. Hasselmo speculates that this combination allows the spread of activation from the higher association areas back to the primary sensory cortex without further influence from the hippocampus. This recurrent activity from higher to lower essentially permits "a process of reanalysis, in which this episodic information would be reanalyzed and reinterpreted in relation to previous semantic representations."[35] Dreams, by this analysis, are part of the restructuring of events through the mediation (and modification) of the system of internal semantic interpretation.

Exactly how all the components of the MTL memory systems are tied together awaits future research, but there seems to be agreement on the few basic facts needed for the purposes of this book. First, the hippocampus acts quickly and can capture complex patterns of association that unfold over time. Secondly,

---

Review of Psychology 48 (1997):547-72, and in "The hippocampus and mechanisms of declarative memory," *Behavioural Brain Research* 103 (1999):123-33.

[35] Michael E. Hasselmo, "Neuromodulation and the hippocampus: memory function and dysfunction in a network simulation," in *Progress in Brain Research* 121 (1999):9. Also see Michael E. Hasselmo and James L. McClelland, "Neural models of memory," *Current Opinion in Neurobiology* 1999, 9:184-188 and A. Robins and S. McCallum, "The consolidation of learning during sleep: comparing psuedorehearsal and unlearning accounts," *Neural Networks* 12(1999):1191-1206. Much research recently has been devoted to delineating the neural dynamics of sleep. See, for example, Terrence J. Sejnowski and Alain Destexhe, "Why do we sleep," *Brain Research* 886 (2000):208-23, Kenway Louie and Matthew A. Wilson, "Temporally Structured Replay of Awake Hippocampal Ensemble Activity during Rapid Eye Movement Sleep," *Neuron* 29 (Jan. 2001):145-56, and Laurel

episodic memory is the enduring synaptic representation of those associations. Finally, semantic memory extracts significant patterns within those episodic associations.

The existence of these memory systems—the short term hippocampal and rhinal cortical components and the long-term episodic and semantic systems—that encode a particular event (my seeing Jamie's ball, for example) suggests that the creaturely meaning of the event (the patterns of activations related to it) can be complex indeed. These evolving strata of memory all in turn shape the seemingly direct perception of "Jamie's bouncing red ball." Although the cortical subsystems implicated in each of these facets of memory and meaning are different, the general model for their influence on perception seems to be largely the same. That is, within a particular experiential domain, the particular aspects of Jamie's bouncing red rubber ball are the activation of a vector within the distributed representation space. Activation of that vector within a particular sensory modality (here, vision) appears to occur through a quick set of negotiations between levels of coding, with input from other systems if available. That is, seeing only the side of Jamie's ball in the dark corner of the room at twilight may not produce instant recognition. Instead, the visual cortex will give "red" and a small curved surface. The "what" stream of the temporal cortex (as opposed to the "where" stream of the parietal cortex) may suggest a ball, a helmet, or a balloon. Semantic memory for BALL may suggest several possible types of balls without any clear winner. Episodic memory may add

---

Graves, Allan Pack and Ted Abel, "Sleep and memory: a molecular perspective," *Trends in*

in "I saw Jamie's ball yesterday afternoon" while other possibilities draw blanks. Competition among the possibilities all along the line quickly strengthens the parts of the particular links "red"-"curved surface"-"ball"-"Jamie's ball"-"yesterday" to reach the final conclusion.[36]

This process of settling upon an identification is one more variation of the feedback loop that applies higher-order structures discussed above. In particular, it resembles the "adaptive resonance" model for memory discussed in Chapter 2. Given the regularity of reciprocal excitatory connections between areas of the brain, this strategy of resonance seems to pervade the higher cortical regions. Although neuroanatomical research reveals differences in neuron populations and connectivity from region to region within the neocortex, this processing strategy is largely shared, and the cortical tissue itself is *largely* the same throughout these regions.[37] The

---

Neuroscience 24.4 (April 2001):237-43.

[36] In particular see Desimone, "Visual attention mediated by biased competition in extrastriate visual cortex," *Proceedings of the Royal Society of London*, B (1998): 1245-1255. For an overview of the processes of integration, see S. Zeki, "Localization and Globalization in Conscious Vision," *Annual Review of Neuroscience* 24 (2001):57-86.

[37] There are differences in cell populations which may affect the relative ease with which areas can process different types of sensory input. These functional differences in turn may be important in shaping the connections established during development. On the whole, however, there is significant neural plasticity. Some areas of the cortex can learn to pick up the processing tasks of damaged adjoining regions if the damage occurs early enough. For a good general discussion of the idea of the specificity of functioning within cortical regions arising out of developmental sequencing, see David J. Buller and Valerie Gray Hardcastle, "Evolutionary Psychology, Meet Developmental Neurobiology: Against Promiscuous Modularity," *Brain and Mind* 1 (2000):307-25. However, this view is by no means universally accepted or certain. (Stephen Pinker's *How the Mind Works* (New York: Norton, 1997) presents one well-known counter-argument.) One particularly contentious issue, for example, is the status of human syntax. Are there particular features in Wernicke's area and Broca's area that shape the unique possibilities for syntax? Is Chomsky's Universal Grammar directly coded by the genes, or is it an emergent feature of the developing cortex's

shared character of these organizational features leads into one last speculative theme, that of the neural substrate of consciousness itself. We already have seen that some researchers emphasize the need for an intact sensory cortex for consciousness to be possible in a given sensory modality. Thus recurrent connections subserve conscious experience. Stephen Grossberg is willing to go further and suggest that only cortical systems organized through adaptive resonant connections can participate in consciousness.[38] The neural logic of conscious experience is at the frontier of current research, but neuroscience has begun to ask systematic questions based on models like Grossberg's about the difference between the neural substrate for processes that are simply outside of the possibility of conscious awareness and those to which we have access. A systematic account of the ways in which the stream of sensory information interacts with evaluative, executive and memory functions—and how some aspects of the interaction are conscious—is beyond the scope of this

---

interaction with its environment? Since Universal Grammar also needs a lexicon of words, does genetic information innately determine the internal structure of the lexicon as well? The question of specialized hardware for processing syntax does not matter too much for the issues of this book—and I can remain agnostic—but the structure of the lexicon is of more direct relevance. Is the lexicon hard-wired, with the categories preset, or are the structures self-organizing maps like the ever higher order feature detectors of the sensory cortex? Studies of the types of aphasias caused by damage to the cortex have explored some aspects of the internal naming space in the brain and suggest that the semantic categories arise through the usual self-organizing logic of competitive neural networks. For example, FRUITS AND VEGETABLES is a category that can become inaccessible in aphasia. Inaccessibility, however, is not simple: models of distributed representation suggest that self-organizing representations should degrade rather than disappear altogether. Indeed, a study with one aphasic patient suggests that this is the case: access becomes hard but not impossible. (See Matthew A. Lambon Ralph, "Distributed versus Localist Representations: Evidence from a Study of Item Consistency in a Case of Classical Anomia," *Brain and Language* 64 (1998):339-360.)

[38] Stephen Grossberg, "The Link between Brain Learning, Attention, and Consciousness," *Consciousness and Cognition* 8(1999):1-44.

chapter and this book.[39] For the moment, what is most pressingly relevant is simply the sense in current research that whatever consciousness may be, it draws heavily on the bottom-up processing of the sensory system and is part of a top-down recurrent logic of selecting what matters from the constant flux of sensory information.

Retrospective and Prospective

This chapter began with questions about the cortical events we call "seeing" and about the neuronally conceived content of sensory experience. The account split into two: at first it stressed the self-organizing character of the visual pathway from the retina to the thalamus and the ascending areas of the visual cortex. But recurrent networks pervade the system from the top to the bottom, so the story must also set out the mutually structuring interactions among perception, attention, and memory. Memory is continuous with the upper reaches of perceptual processing, and what matters at the higher levels shapes what is salient at the lower level of cognitive maps.

The cortical substrate for sensory experience is an ordered web of rapidly ramifying activations where all the levels of featural abstraction, memory and evaluation constantly talk to one another simultaneously. This is a world of shades and echoes and drifting focus: this is in fact the world we know from our own

---

[39] For one promising approach to the large-scale integration of cortical sub-systems in active thought, see Stanislas Dehaene, Michel Kerszberg, and Jean-Pierre Changeux, "A neuronal model of a global workspace in effortful cognitive tasks," *PNAS* 95 (1998):14529-14534.

experience. It is richly interconnected but correspondingly highly mediated. Even the most direct encounter follows an inner trajectory shaped by complexly sedimented traces of past experience. The next question we must confront is what controls the logic of sedimentation. There are internal biases to sensory experience—neuroscientists refer to such biases as "biologically relevant" input—that must be at work from birth if the child is to survive. The next chapter explores how these biological biases shape the emergent properties of that system we call our selves.