

Is Neuroscience Relevant to Philosophy?

PATRICIA SMITH CHURCHLAND

To cite this article: PATRICIA SMITH CHURCHLAND (1990) Is Neuroscience Relevant to Philosophy?, Canadian Journal of Philosophy, 20:sup1, 323-341

To link to this article: <http://dx.doi.org/10.1080/00455091.1990.10717230>



Published online: 01 Jul 2013.



Submit your article to this journal [↗](#)



Article views: 112



View related articles [↗](#)

Is Neuroscience Relevant to Philosophy?

PATRICIA SMITH CHURCHLAND
University of California, San Diego
La Jolla, CA 92093
U.S.A.

Many questions concerning the nature of the mind have remained intractable since their first systematic discussion by the ancient Greeks. What is the nature of knowledge, and how is it possible to represent the world? What are consciousness and free will? What is the self and how is it that some organisms are more intelligent than others? Since it is now overwhelmingly evident that these are phenomena of the physical brain, it is not surprising that an established empirical and theoretical foundation in this domain has eluded us for so long. For in order to understand what we are and how we work, we must understand the brain and how it works. Yet the brain is exceedingly difficult to study, and research on any significant scale is critically dependent on advanced technology.

Unlike celestial mechanics or evolutionary biology, where groundbreaking achievements could be made with relatively simple instruments, neuroscientific progress was severely restricted until the advent of cellular physiology, the light microscope, techniques that would selectively stain single cells, the electron microscope, micro-electrodes, and modern electronics generally. To be sure, physicians since ancient times have made important clinical observations about their human patients with brain damage, but these observations could not transcend the merely intriguing stage so long as we remained ignorant of the micro-organization of nervous systems and the computational capacities of neural networks.

Within this century, and especially within the last two decades, a number of scientific developments have dramatically altered the status of mind-brain questions, moving them from the shelf labeled 'well-nigh-impenetrable-mystery' to that labeled 'difficult-but-tractable.' There is a gathering sense that some of the major pieces of the puzzle are on the brink of falling into place, and that basic neurobiological explanations for certain psychological phenomena are now within reach.

The general developments are threefold. (1) There has been a spectacular blossoming of data describing the nervous system. New neuroscientific techniques have made possible very detailed structural and functional descriptions of nervous systems at many levels of organization. This has helped to dislodge certain misconceptions about brain mechanisms, for example that past the first cortical stage of sensory processing the next stages of cortex perform the 'association of ideas' from different sensory modalities. But it has also functioned to catalyze theorizing. For example, the detailed results concerning patterns of connectivity and cell types in the visual cortex have inspired hypotheses concerning how circuits compute the direction and velocity of a moving stimulus,¹ and how the curvature of a line can be detected by cells using signals from orientation-selective cells in cortical area V1.² (2) Increasingly subtle and sophisticated behavioral studies in experimental psychology and ethology have greatly deepened our understanding of what exactly are the psychological capacities, thereby clarifying the molar phenomena for which neurobiology seeks the mechanisms.³ (3) Computer

1 See, for example, E. Hildreth and C. Koch, 'The Analysis of Visual Motion: From Computational Theory to Neuronal Mechanisms,' *Annual Review of Neuroscience* 10 (1987) 477-534.

2 See S. Zucker, A. Dobbins, and L. Iverson, 'Two Stages of Curve Detection Suggests Two Styles of Visual Computation,' *Neural Computation* 1 (1989) 68-81.

3 See, for example, D. Schacter, 'The Psychology of Memory,' and W. Hirst, 'The Psychology of Attention,' both in J. LeDoux and W. Hirst, eds., *Mind and Brain* (Cambridge: Cambridge University Press 1986); on vision, see V. Ramachandran, 'Perceiving Shape From Shading,' *Scientific American* 259 (1988) 76-83.

modeling approaches that permit effective simulation of neural networks have led to computational discoveries concerning how networks of neuron-like units, with synapse-like connections and a parallel organization, can accomplish certain complex tasks such as associative memory and pattern recognition. Since cognitive functions appear to be systems-level properties, and in that sense are emergent properties, this research promises to be an important bridge between basic neuroscience and experimental psychology.⁴

It is not that we now have clear and complete neurobiological answers to traditional philosophical questions. Rather, the philosophical significance derives from the nature of the progress made by means of a variety of techniques, broadly neuroscientific in scope, that are converging on problems about the mind. The convergent data have begun to suggest plausible theories (see below), and newly developing techniques, such as positron emission tomography (PET), magnetoencephalography (MEG), and voltage sensitive dyes, are providing information about brain function at a level above the level of the single cell. PET scans, for example, have a spatial resolution of about 1 millimetre, and can reveal which areas of the brain are differentially active given a particular task, such as looking at a light, or, with greater complexity, reading silently, or with yet greater complexity, generating a word to associate with a stimulus word.⁵

Neurobiological data bearing upon such favourite philosophical issues as representation, consciousness, and perception are beginning to be available (see below), and philosophers cannot now afford to speculate on these issues in ignorance of the data. That the data are relevant in this context has the same rationale as the relevance of evidence in any context of inquiry. Thus philosophers con-

4 See S. Lehky and T. Sejnowski, 'Neural Network Model for the Representation of Surface Curvature from Images of Shaded Surfaces,' in J. Lund, ed., *Organizing Principles of Sensory Processing* (Oxford: Oxford University Press 1988); P.S. Churchland and T. Sejnowski, 'Neural Representations and Neural Computations,' in L. Nadel, L. Cooper, P. Culicover and R.M. Harnish, eds., *Neural Connection Mental Computations* (Cambridge, MA: M.I.T. Press 1988) 15-48.

5 See M. Posner, S. Petersen, P. Fox, and M. Raichle, 'Localization of Cognitive Operations in the Human Brain,' *Science* 240 (1988) 1627-31.

cerned with the nature of space and time cannot ignore Einstein's theory of special relativity and the physics of space and time, and philosophers concerned with the nature of life cannot ignore molecular biology. Although this parallel is now generally adopted within the sub-field of philosophy of science, it is frequently rejected within the wider discipline, where philosophers are apt to assume that humans are special in ways that put human mental life beyond scientific understanding entirely⁶ or at least beyond the reach of neuroscience.⁷

Humans *are* special in certain respects, and the human brain appears to be the most complex brain on the planet. Remarkable capacities notwithstanding, we, like every other organism, are products of natural selection; our brains have evolved from simpler brains. Consequently, our cognition, awareness, and motor control are not likely to be radically distinct from the cognition, awareness, and motor control of other mammals. From a biological perspective, those features that do distinguish us from our nearest living relatives are far more likely to be minor modifications on the basic design than a top-to-bottom innovation. Evolution, as François Jacob⁸ has reminded us, does not proceed by redesigning from scratch, but by modifying what is already in place.

Additionally, the biological perspective invites us to see higher functions within the wider framework of the basic requirements for survival, and survival is crucially dependent on sensorimotor control. Constraints on cognitive design derive from a world of hungry predators and unwilling prey, and the organization of cognitive functions is not independent of the motor functions they serve. Cognition and consciousness were made not in Plato's heaven, but in the competitive Darwinian world where small improvements in sensorimotor control could give an organism a predatory and reproductive edge. Having the Darwinian lesson truly sink in is perhaps the most important element in loosening the bonds of traditional

6 C. Taylor, *Human Agency and Language* (Cambridge: Cambridge University Press 1985)

7 J.A. Fodor, *Language of Thought* (New York: Crowell 1975)

8 F. Jacob, *The Actual and the Possible* (Seattle and London: University of Washington Press 1982)

philosophical approach that dismisses neurobiology as irrelevant to understanding our nature.

Whilst the Darwinian lesson provides the basic biological perspective, to appreciate the sense of impending scientific revolution in cognitive neuroscience, it is necessary to comprehend the progress that has been made in unscrambling the neurobiological mechanisms for psychological phenomena. The breadth and depth of research, at all levels of nervous system organization, is obviously far too vast to be surveyed in short compass,⁹ but I shall try to evoke the sense of being-on-the-scent by pointing to selected examples of research on vision, consciousness, and learning, this choice being guided by relevance to traditional philosophical questions.

Vision research spans work at many levels, from individual neurons and neuronal circuits, to behavior (Figure 1). Although there are still many unanswered questions, there is now so much known about the microorganization of the visual system that it is possible to generate testable hypotheses to explain function, and thus to begin to bridge the explanatory gulf between the neurobiology of the visual system and the psychology of vision. For example, the discoveries showing spatially segregated visual pathways that appear to subservise distinct functions has laid the groundwork for explanations for certain psychophysical phenomena. The segregation begins in the retina, and is preserved at the next synaptic station, the lateral geniculate body (LGN), and is preserved in the various maps in visual cortex (Figure 2). Thus one pathway (the parvo-interblob-pale-stripe pathway) appears to be specialized for information about borders and probably shapes. The second pathway (the magno-4B-thick-stripe-MT pathway) subserves motion and stereo perception, whilst the third (the blob-thin-strip-V4 pathway) seems specialized for information about colorvision.¹⁰

9 For a thumbnail sketch, see T.J. Sejnowski and P.S. Churchland, 'Brain and Cognition,' in M. Posner, ed., *Foundations of Cognitive Science* (Cambridge, MA: M.I.T. Press 1989).

10 For a very readable account, see M. Livingstone, 'Art, Illusion, and the Visual System,' *Scientific American* 258 (1988) 78-85; for an excellent and philosophically insightful introduction to the neurophysiology of vision, see D. Hubel, *Eye, Brain and Vision* (New York: W.H. Freeman 1988).

Levels of Investigation

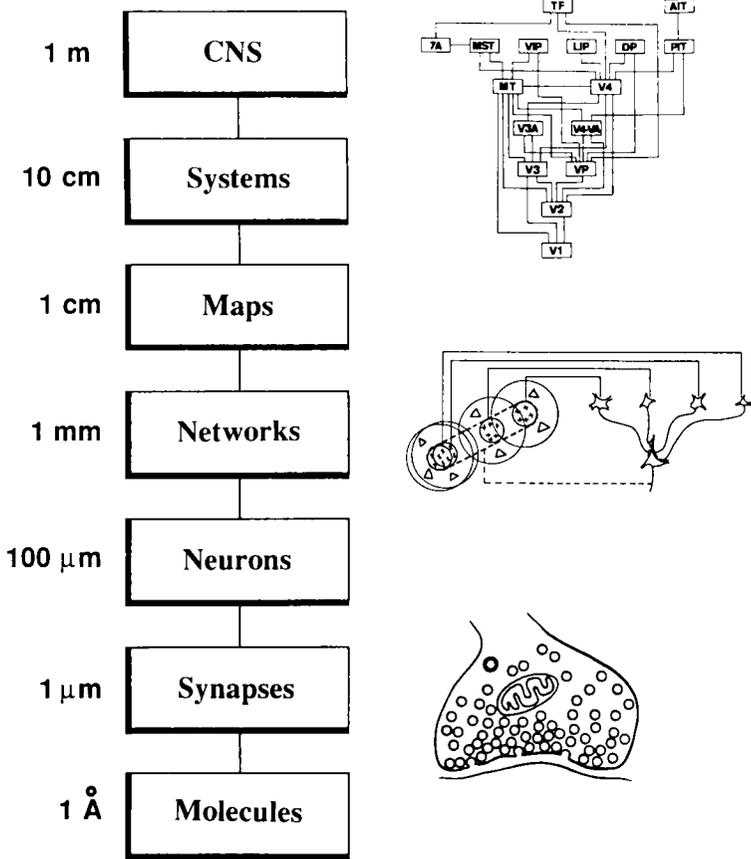


Figure 1

Structural levels of organization in the nervous system. The spatial scale at which anatomical organizations can be identified varies over many orders of magnitude. Schematic diagrams illustrate (top) a processing hierarchy of visual areas in the monkey visual cortex; (centre) a small network model for the synthesis of oriented receptive fields of simple cells in visual cortex; (bottom) the structure of a chemical synapse. From Churchland and Sejnowski, *Science* 242 (1988).

This organization begins to explain certain puzzling facts of human visual experience, such as that a person can detect where something is without knowing what it is,¹¹ and that in cortical color blindness (cerebral achromatopsia), the subject has intact stereoptic, motion, and form perception, but retains only high-contrast gray level discrimination.¹² It helps to explain why, in unusual conditions where colors in a scene are isoluminant, stereo vision disappears¹³ and motion perception is seriously degraded; it also explains how someone's perception of motion can be selectively destroyed by a lesion.¹⁴ Collectively, these and many other discoveries in vision research constitute major inroads to the problem of how we see. They are, to be sure, only a beginning. Yet knowing this much, further experiments both in neurobiology and in psychophysics will be provoked, and a fruitful co-evolution of theory and experiment can occur.

Consciousness often seems a deeply puzzling phenomenon, and our capacity to be aware of some things and not others, to experience the richness of colors, sounds, and smells, and so forth, may appear very remote from neurons, action potentials, and transmitter release. Such differences have convinced some philosophers¹⁵ and scientists¹⁶ that being conscious is not a state of the physical brain at all, but a state of a nonphysical soul.

11 D. Sagi and B. Julesz, "'Where' and 'What' in Vision," *Science* **228** (1985) 1217-19

12 See A. Damasio, T. Yamada, H. Damasio, J. Corbett and J. McKee, 'Central Achromatopsia: Behavioral, Anatomical, and Physiologic Aspects,' *Neurology* **30** (1980) 1064-71.

13 See C. Lu and D.H. Fender, 'The Interaction of Color and Luminance in Stereoscopic Vision,' *Investigative Ophthalmology* **11** (1972) 482-90.

14 See J. Zihl and D. Von Cramon, 'The Contribution of the "Second" Visual System in Directed Visual Attention in Man,' *Brain* **102** (1979) 835-56.

15 For example, R. Descartes, *Les Passions de L'ame*, trans. Haldane and Ross (Cambridge: Cambridge University Press 1911); R. Swinburne, *The Evolution of the Soul* (Oxford: Oxford University Press 1986).

16 For example, J.C. Eccles, *The Self and Its Brain*, Part Two (Berlin: Springer 1977).

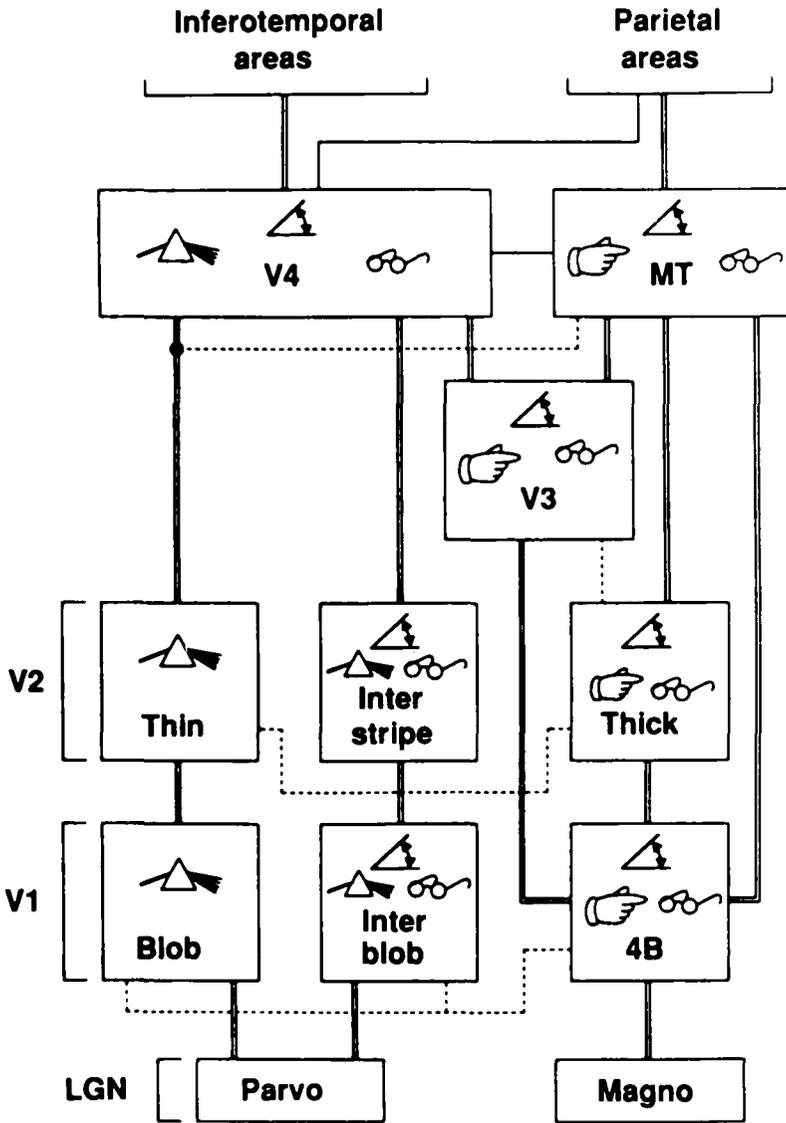


Figure 2

Apart from the inconsistency of this conviction with biology and physics, it rests on a very dubious assumption; to wit, if a phenomenon detected by one means *seems* very different from a phenomenon detected by another, the two phenomena cannot in fact be identical. The assumption has well-known counter-examples: temperature seems subjectively to be very different from mean molecular kinetic energy, yet that is what it is. Light, as detected by our visual system, seems to be very different from electromagnetic radiation as detected by a photoelectric cell; nevertheless, visible light *is* electromagnetic radiation. There are counter-examples to the converse as well: rusting and burning do not seem to be instances of the same phenomenon, yet they are both instances of oxidation. Indeed, science is replete with discoveries about the

Figure 2

Schematic diagram of anatomical connections and response selectivities of neurons in the early visual areas of the macaque monkey. Visual information from the retina is split into two parallel streams – the parvocellular and the magnocellular – at the level of the lateral geniculate nucleus (LGN). The parvocellular stream projects to two divisions of primary visual cortex; the areas rich in the enzyme cytochrome oxidase (called blobs), and second, the cytochrome oxidase poor regions (interblobs). The Magnocellular stream projects to layer 4B of V1. These three divisions of V1 project into corresponding areas of V2: the ‘thin strip,’ ‘interstripe,’ and ‘thick stripes’ of cytochrome oxidase-rich and –poor regions of V2. These areas in turn project to V3, V4, and MT (middle temporal areas, also called V5). Heavy lines indicate robust primary connections, and thin lines indicate weaker, more variable connections. Dotted lines indicate connections that require additional verification. Not all projections from these areas to other brain areas are represented. The neurons in each visual area respond preferentially to particular properties of visual stimuli, as indicated by the icons: Prism means responsive to colored stimuli; Angle means selective for orientation of the stimulus; Spectacles mean selective for binocular disparity or for strong binocular interactions; Pointing hand means selective for direction of motion of the stimulus. Reprinted with permission from E. De Yoe and D. Van Essen, *Trends in Neuroscience* 11 (1988).

nature of things that are radically at odds with what naive intuition bid us assume.¹⁷

Two main lines of neuroscientific research seem to me particularly important for resolving the puzzles about consciousness.

The first focuses on the changes in state from waking to dreaming (REM sleep) to nondreaming (NREM) sleep.¹⁸ Because awareness changes dramatically in these different states, determining the brain mechanisms regulating the shifts is a key element in understanding consciousness. Neuroscientific techniques are beginning to reveal the microdynamical features distinguishing the three states and the neuronal generator for producing shifts.¹⁹

More specifically, neurons in a brainstem structure called the locus coeruleus that project very widely all over the cortex, are maximally active during the waking state, show declining activity during deep sleep, and are largely inactive during dreaming. With just the opposite phase are giant neurons in another brainstem area called the pons (see Figure 3).

The second line studies selective attention using a variety of techniques, ranging from clinical and PET studies of patients with hemineglect (where one hemisphere, usually the left hemisphere, and its objects seems not to exist for the subject) and schizophrenia,²⁰ to single cell studies showing cells responding as a function of at-

17 P.M. Churchland, *Matter and Consciousness*, 2nd ed. (Cambridge, MA: M.I.T. Press 1988)

18 See also P.S. Churchland, 'Reduction and the Neurobiological Basis of Consciousness,' in A. Marcel and E. Bisiach, eds., *Consciousness in Contemporary Science* (Oxford: Oxford University Press 1988).

19 See J.A. Hobson, R. Lydic, and H.A. Baghdoyan, 'Evolving Concepts of Sleep Cycle Generation: From Brain Centers to Neuronal Populations,' *Behavioral and Brain Sciences* 9 (1986) 371-448. For a very readable and beautiful introduction, see J.A. Hobson, *Sleep* (New York: W.H. Freeman 1989).

20 T. Early, E.M. Reiman, M.E. Raichle, and E.L. Spitznagel, 'Left Globus Pallidus Abnormality in Never-Medicated Patients with Schizophrenia,' *Proceedings of the National Academy of Science, USA* 84 (1987) 561-3.

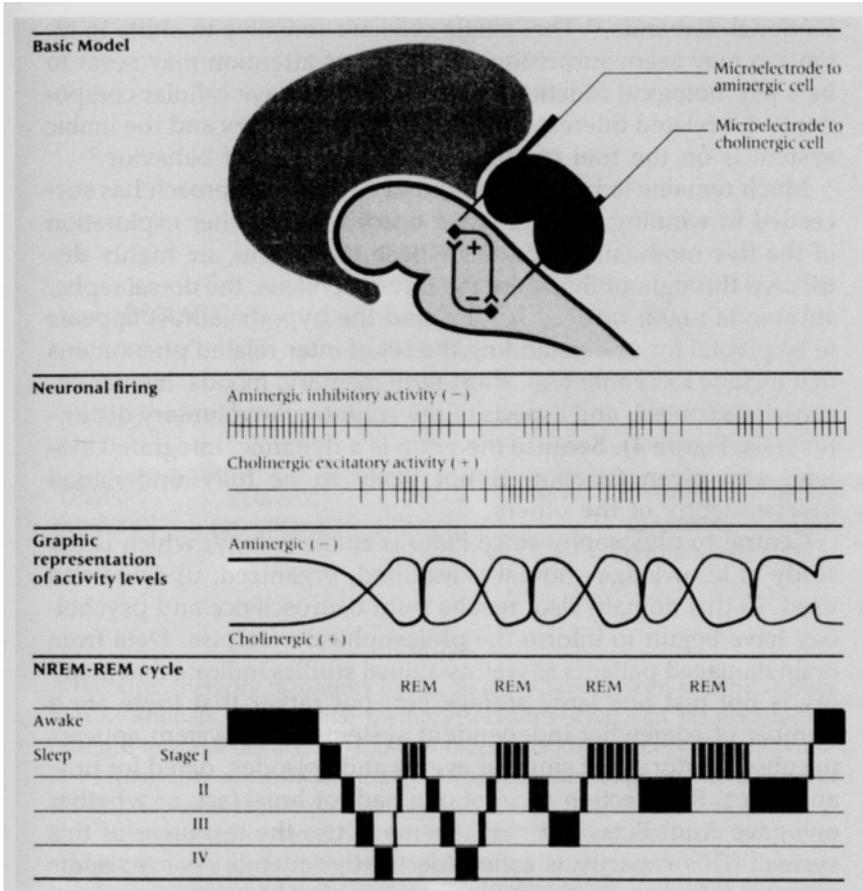


Figure 3

Animal studies indicate that the alternation of NREM and REM phases of sleep is generated by the alternating activity of specific nuclei of neurons in the brain stem. Some of these nuclei are strongly inhibitory and mediated by serotonin and norepinephrine; their actions are opposed by a group of excitatory cholinergic neurons. From J. Allen Hobson, 'The Neurobiology and Psychophysiology of Sleep and Dreaming,' *Discussions in Neurosciences* 2 (1985) 9-50.

tentional demands.²¹ That single cells are sensitive to shifts in attention may seem surprising, inasmuch as attention may seem to be a psychological function too 'cognitive' to have cellular components. Of related interest, research on frontal cortex and the limbic system is on the trail of volitional components of behavior.²²

Much remains to be discovered, but so far the approach has succeeded in winning an impressive beachhead. Further exploration of the five modulatory systems whose projections are highly distributed throughout the brain (the locus coeruleus, the dorsal raphe, substantia nigra, nucleus basalis, and the hypothalamus) appears to be pivotal for understanding the set of inter-related phenomena that include focal attention, short-term memory, moods, mental set, conscious control, and aspects of the voluntary/involuntary distinction (see Figure 4). Because the brain is a dynamic, integrated system, any given function is not likely to be fully understood independently of the others.

Central to philosophy since Plato is epistemology, which is the study of knowledge—how it is acquired, organized, updated and used. In this domain also, results from neuroscience and psychology have begun to inform the philosophical enterprise. Data from brain damaged patients as well as animal studies indicate that memory is not just one large storage bin, but rather that there are a number of somewhat independent systems. One system appears to subserve storage of singular events and episodes, dated for time and place. Recollection of what one had for breakfast, of whether one gave Aunt Betsy her mail, seems to use the resources of this system. This capacity is called 'declarative memory' or 'episodic memory.' A second system, sometimes referred to as 'procedural memory' enables the learning of skills, both cognitive and motor. There are several famous patients whose capacity to learn new skills

21 The classic work here has been done by Desimone and his colleagues. See for example, H. Spitzer, R. Desimone, and J. Moran, 'Increased Attention Enhances Both Behavioral and Neuronal Performance,' *Science* 240 (1988) 338-40.

22 G. Goldberg, 'Supplementary Motor Area Structure and Function: Review and Hypotheses,' *Behavioral and Brain Sciences* 8 (1985) 567-615.

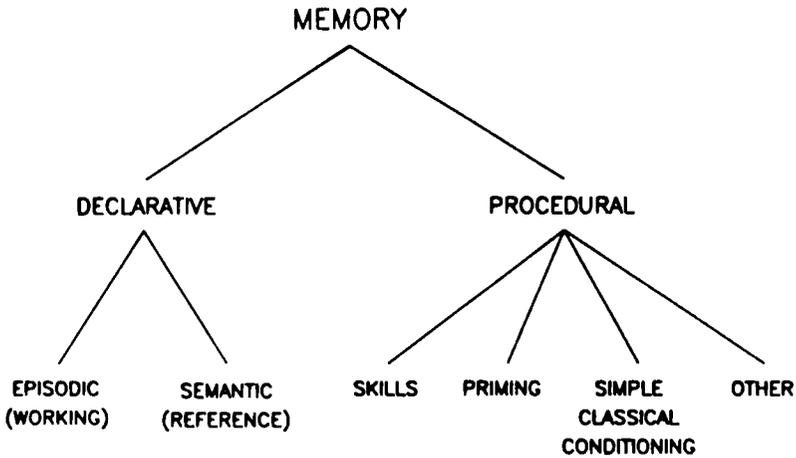


Figure 4

A tentative memory taxonomy based on dissociations seen in clinical cases and in animals. Declarative memory includes what can be declared or brought to mind (awareness). It embraces episodic memory (for individual events) and semantic memory (broadly, knowledge of the nature of things). Other researchers may use the related terms, 'working memory' and 'reference memory.' Under procedural memory are included motor skills, cognitive skills, simple classical conditioning, as well as habituation, sensitization and various perceptual after-effects, and other instances where the facility for engaging specific cognitive operations is improved by experience. Learning of the procedural sort may not be accessible to awareness and different phenomena under this heading may have different neurobiological mechanisms. From L.R. Squire, *Handbook of Physiology—The Nervous System* (1987).

is largely intact but who have completely lost the capacity to store new information about singular events, though their short term memory appears to be quite normal and their memory for events before the lesion is also essentially normal.²³

One particular brain structure, the hippocampus, has been implicated in the loss of the capacity to lay down new information in long term memory. Because the patient's retrograde memory is intact, the hippocampus cannot itself be the long-term storehouse, but it may be the mechanism needed to transfer information to the cortex.

Explorations of structure and function of the hippocampus at the cellular level²⁴ reveals that there are individual cells called 'pyramidal cells' whose responsiveness is enhanced when stimulated with a high frequency burst of pulses (Figure 5). The effect, called 'long term potentiation,' or for short, 'LTP' seems to conform to the model of cellular learning predicted by D.O. Hebb in 1949.²⁵

Hebb suggested that when the presynaptic and postsynaptic cell were simultaneously active, then there would be facilitatory modification of the synapses, and that this could be the mechanism for associating events in memory. This is what seems to happen in the hippocampus. Very recently, what may be the cellular analogue of short-term dissociation has also been found, and is called LTD,

23 For review papers, see L.R. Squire, 'Memory: Neural Organization and Behavior,' *Handbook of Physiology—The Nervous System*, J.M. Brookhart and U.B. Mountcastle, eds. (Bethesda, MD: American Physiological Society 1987), 295-371; M. Mishkin, and T. Appenzeller, 'The Anatomy of Memory,' *Scientific American* 256 (1987) 80-9; for a fascinating discussion of a clinical case, see A. Damasio, P. Eslinger, H. Damasio, G. Van Hoesen, S. Cornell, 'Multimodal Amnesic Syndrome Following Bilateral Temporal and Basal Forebrain Damage,' *Archives of Neurology* 42 (1985) 252-9.

24 G. Collingridge and T. Bliss, 'NMDA Receptor—Their Role in Long-Term Potentiation,' *Trends in Neurosciences* 10 (1987) 288-93

25 D.O. Hebb, *The Organization of Behavior* (New York: Wiley 1949); for neurophysiological results showing Hebbian connections, see T. Brown, A. Ganong, E. Kariss and C. Keenan, 'Hebbian Synapses—Computations and Biophysical Mechanisms,' *Annual Review of Neuroscience* 11 (1989) 475-511

or 'long term depression.' It is anti-Hebbian, in the sense that when the activity of the two connected cells are out of phase, the responsiveness of the postsynaptic cell is depressed with respect to input from the presynaptic cell.²⁶

Whether LTP is the cellular correlate of a mechanism for laying down new information in long term memory is not yet known, but a major piece of luck to aid the investigation is that there is a particular kind of receptor at the synapse, known as the NMDA receptor, that mediates LTP. A chemical, AP5, has been found which selectively blocks the NMDA receptor. This is a boon for all sorts of reasons: the concentration of NMDA receptors throughout the brain can be charted; AP5 can be used as a fine-grained means of dissecting out the memory system.

Although these discoveries at the molecular level are important, what we also badly need are computer models to help reveal what circuits of cells in the hippocampus do, and what their role is in storing information in the cortex. Research of this kind is underway, for example in the Sejnowski laboratory at the Salk Institute in La Jolla.

Perhaps the most entrenched dogma in epistemology is that beliefs are the prototypical representational states, that what is believed is a proposition, and that what relates beliefs to one another are rules, including rules of logic.²⁷ Yet this proposition-logic model comports very poorly with how knowledge and problem-solving looks from a neurobiological perspective. The really difficult question has been to answer what representations are like, if not propositions; to answer what a neurobiologically plausible theory of representations would look like.²⁸

A major conceptual breakthrough on this front has been made with the development of connectionist models of learning and as-

26 P. Stanton and T. Sejnowski, 'Associative Long-Term Depression in the Hippocampus Induced by Hebbian Covariance,' *Nature* 339 (1989) 215-18

27 Fodor, *Language of Thought* (New York: Crowell 1975)

28 I explored this question in the last chapter of my *Neurophilosophy* (Cambridge, MA: M.I.T. Press 1986).

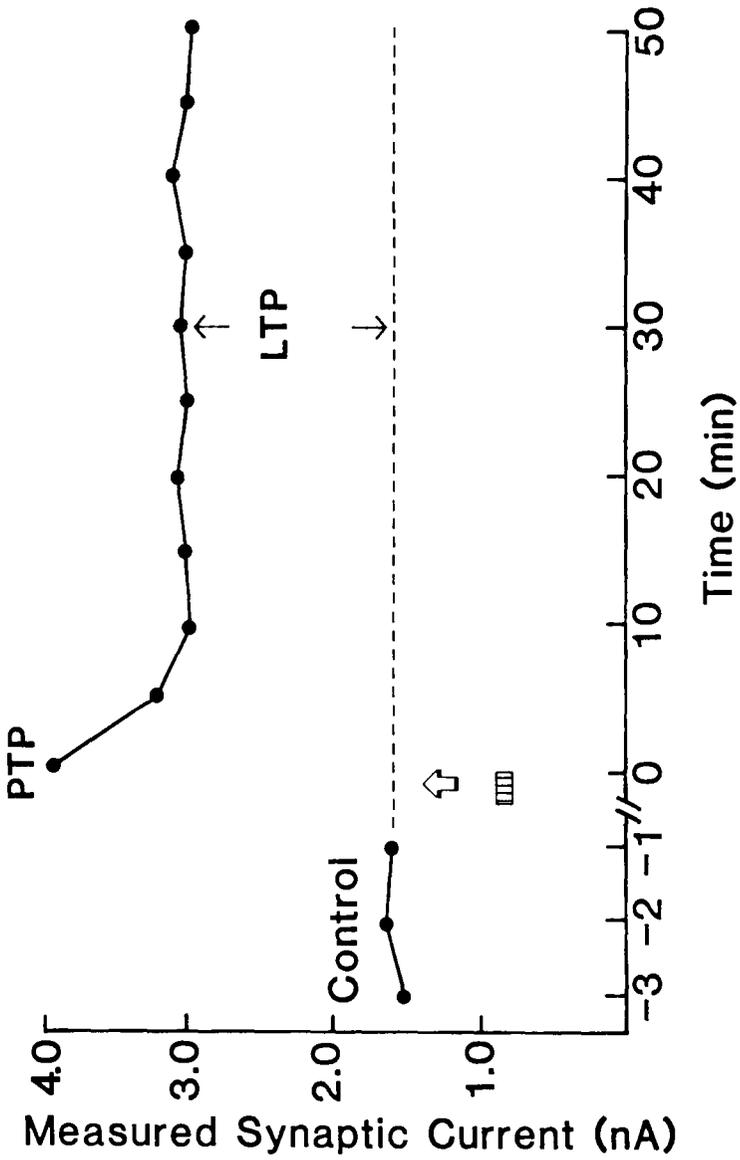


Figure 5

sociative memory.²⁹ They have provided a demonstration of what representations could be, if not sentences, and how computation could be accomplished, if not by following explicit rules. These models honor general neurobiological constraints, and in this respect they are much more realistic than models designed for a serial, digital computer. The models are brain-like only to a first approximation, but the approach has the conceptual resources to design models of increasing biological realism, and thus to co-evolve with neurobiology. If, as it seems, certain macroproperties of mammalian nervous systems, such as learning, emerge at the circuit and the systems levels, then models of networks and their interaction will be essential to explaining those properties.³⁰ Computa-

29 Representations by Error Propagation,' in D. Rumelhart and J. McClelland, eds., *Parallel Distributed Processing* Vol. 1 (Cambridge, MA: M.I.T. Press 1986) 318-62. For a very impressive example of network learning, see T.J. Sejnowski and C.R. Rosenberg, 'Parallel Networks that Learn to Pronounce English Text,' *Complex Systems* 1 (1987) 145-68. For an example in vision that has been very difficult to approach using symbol-manipulation but which is quite successful using a network architecture, see S.R. Lehky and T.J. Sejnowski, 'Network Model of Shape-from Shading: Neural Function Arises from both Receptive and Projective Fields,' *Nature* 333 (1988) 452-4. For some philosophical applications of network models, see P.M. Churchland, *A Neurocomputational Perspective* (Cambridge, MA: M.I.T. Press 1989).

30 A.I. Selverston and M. Moulins, 'Oscillatory Neural Networks,' *Annual Review of Physiology* 47 (1985) 29-48; for a view of the goals and strategies of computational neuroscience, see P.S. Churchland, C. Koch and T.J. Sejnowski, 'What is Computational Neuroscience?' in E. Schwartz, ed., *Computational Neuroscience* (Cambridge, MA: M.I.T. Press 1990).

Figure 5

Mossy fibre EPSP (excitatory post synaptic potential) amplitude plotted over time, before and after the induction of LTP. Brief tetanic stimulation was applied at the time indicated (striped bar and arrow). Note the change in time scale at the time of stimulation. Each data point is the average of 5 EPSP obtained from a holding potential of -90 millivolts. The tetanic stimulation induced posttetanic potentiation (PTP), which has a short time constant, and long-term potentiation (LTP), which may last for many hours.

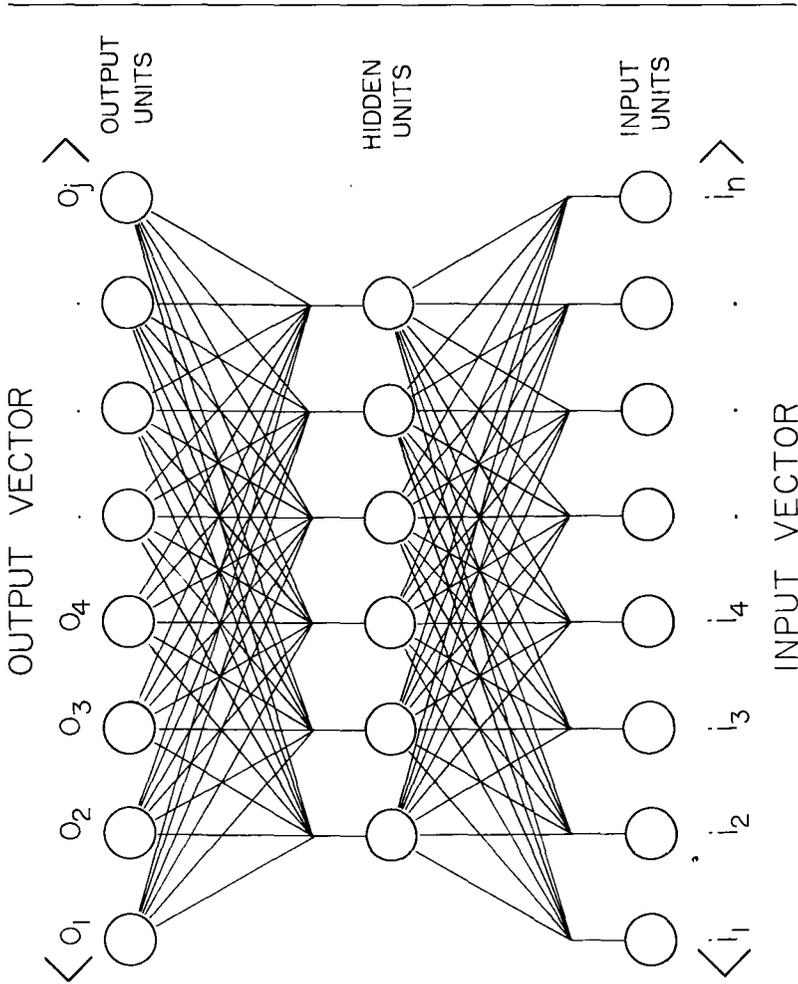


Figure 6

A simple three-layer network showing the input vectors and the output vectors. The activation vector for the hidden units (not shown) after training corresponds to the representation or category found by the units (courtesy Paul Churchland).

tional models are therefore important in bridging the gap between structure and function—between neurons and thought.³¹

Within the connectionist framework, representations turn out to be activation patterns across sets of units; mathematically, they are vectors in a state space (see figure 6). Computations turn out to be vector transformations from one state space to another. If something even roughly like this conception is right, it overturns the conventional wisdom, and brings philosophy, psychology and neuroscience into closer cooperation.

Scientific discoveries frequently cause a profound shift in how we think about our universe, and consequently they may reconfigure the very questions we ask.³² Thus, after Copernicus, it was no longer worth asking what made the crystal spheres turn daily; after Harvey it was no longer worth asking how the heart concocted animal spirits; once Newton framed the laws of motion, no one cared about an object's 'natural place' nor about the properties of 'impetus'; after Lavoisier, the problem of the negative weight of phlogiston could safely be ignored as entirely misbegotten.

Coming to understand representing within the context of computational neuroscience is potentially revolutionary. As we discover the properties of circuits and systems and how they achieve macro effects, doubtless some time-honored assumptions about our own nature will be reconfigured. More generally, it is probable that our commonly accepted ideas about knowledge, reasoning, free will, the self, consciousness, and perception, have no more integrity than prescientific ideas about substance, fire, motion, life, space and time. We still have a long way to go, but the convergence of research in neuroscience, psychology, and computational modeling holds out the promise that at least some basic principles of neural representation can be understood.

31 See my *Neurophilosophy* (Cambridge, MA: M.I.T. Press 1986); also P.M. Churchland, *The Neurocomputational Perspective* (Cambridge, MA: M.I.T. Press 1989).

32 See T. Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press 1962); P.S. Churchland and T.J. Sejnowski, 'A Perspective on Cognitive Neuroscience,' *Science* 242 (1989) 741-5.