Memory in Neuroscience: Rhetoric Versus Reality

Jonathan R. Wolpaw
Wadsworth Center, New York State Department of Health, and State University of New York, Albany, NY

The central point of this article is that the concept of memory as information storage in the brain is inadequate for and irrelevant to understanding the nervous system. Beginning from the sensorimotor hypothesis that underlies neuroscience—that the entire function of the nervous system is to connect experience to appropriate behavior—the paper defines memories as sequences of events that connect remote experience to present behavior. Their essential components are (a) persistent events that bridge the time from remote experience to present behavior and (b) junctional events in which connections from remote experience and recent experience merge to produce behavior. The sequences comprising even the simplest memories are complex. This is both necessary—to preserve previously learned behaviors—and inevitable—due to secondary activity-driven plasticity. This complexity further highlights the inadequacy of the information storage concept and the importance of extreme simplicity in models used to study memory.

Key Words: behavior, conditioning, learning, memory, plasticity

1. INTRODUCTION

1.1 Prescientific Ideas in Science

Scientific problems often begin as popular, religious, or philosophical questions. The nature of the heavenly bodies and of their movements, the position and composition of the Earth, and the origin, structure, and function of humans and other living things were subjects for speculation, argument, and dogma long before they became objects of scientific inquiry. For each question, the central event in its transition to a scientific problem is the development of a more or less consistent body of theory and data that is henceforth continually modified by successive cycles of hypothesis and data collection. In this transition, ideas of popular, religious, or philosophical origin are either abandoned or they are transformed into scientific concepts that fit into the new theoretical framework. Until the transition is completed, the high stature and persistent appeal of these prescientific ideas can generate issues that appear at the time to be legitimate scientific concerns but are in reality products of alternative or complementary approaches to understanding the world. Prescientific ideas can intrude into the formulation of hypotheses and into the design and interpretation of experimental results.

The influence of prescientific ideas was strong in physics and chemistry during the 17th and 18th centuries. Galileo was the central figure in the most famous example. Early in the 17th century, he initiated modern analysis of motion by extending the concept of inertia (which had previously been applied only to bodies at rest) to bodies in motion, stating that a moving body will continue to move unless acted on by an external force, and he applied this concept to both celestial and terrestrial moving bodies (Galilei, 1914; Wolf, 1950). However, unlike the current concept of inertia—that a moving body subjected to no external forces will move at a constant speed in a straight line—Galileo’s concept was that a moving body subjected to no external forces will move at a constant speed in a circle. His concept of circular inertia came from a philosophical and religious assum-
tion that had prevailed since classical times. This ancient belief was that the “integral parts of the universe [the planets and other celestial bodies] are constituted in the best arrangement,” that “perfect order” maintains among them (Galilei, 1967, p. 32). Only circular motion, because it returns perpetually to the same set of locations, was considered to be perfect and unchanging. Thus, it was believed that the parts of the universe must naturally move in circles. If they were to move in any other way, for example, in straight lines, they would pass from one state into another and would not be perpetually perfectly arranged (Hall, 1981).

Although Galileo’s own studies with the newly invented telescope provided evidence against the assumption of celestial perfection, he continued to assume that no forces acted on the celestial bodies, and he ascribed their circular motion to circular inertia (Finocchiaro, 1980; Hall, 1966; Wolf, 1950). Galileo was unwilling or unable to “conceive of breaking the magic circle of the Cosmos, without which any order appeared unthinkable” (Galilei, 1953, p. 23). It was left to Descartes to develop the modern concept of inertia and to Newton and Hooke to extend Galileo’s analysis of terrestrial motion and the concept of gravity to explain the movements of the celestial bodies (Butterfield, 1961; Hall, 1966, 1981).

Since the scientific revolution began, the abandonment or modification of prescientific ideas has often occurred rapidly. For example, in the 16th century the prevailing scientific analysis of the circulation of the blood was the 1400-year-old theory of Galen, who taught that blood went from the right ventricle to the left by seeping through the intraventricular septum. His hypothesis was entangled with critical theological and philosophical questions, such as how the divine spirit entered the body. In 1553, in a wildly heretical theological tract titled “Christianismi Restitutio” (“The Restoration of Christianity”), Michael Servetus answered this question with the new hypothesis that blood passed from the right ventricle to the left by way of the lungs, where it acquired the divine spirit from the air (Bainton, 1953; Bayon, 1938, 1939; Boas, 1962). Later the same year, Servetus and his book were burned at the stake in John Calvin’s Geneva. (It seems that in the 16th century, the rule was not “publish or perish,” but rather “publish and perish.”) Nevertheless, by early in the next century, shortly after William Harvey described the entire circulation (except for the capillaries), cardiovascular physiology had become a purely scientific subject (Boas, 1962; Butterfield, 1961; Hall, 1966; Harvey, 1957).

1.2 Prescientific Ideas in Neuroscience

The problem of the function of the nervous system has not fared so well. Although generally recognized as a purely scientific problem early in the 19th century, understanding of the nervous system remains, nearly 200 years later, strongly influenced by ancient ideas. More than any other natural phenomena, the operations of the nervous system inspired an extensive roster of prescientific concepts. These came from introspection, that is, from each person’s conscious perceptions of events inside himself or herself and from the inferences this consciousness prompted about events within others. Introspection led to the recognition of an immaterial entity called the soul, or the mind, that possessed functions, often referred to as faculties. The list of these functions differed from one description to another but always included in some guise the functions of perception, volition, emotion, consciousness, and memory.

In the past, these functions were also projected into other objects and events. Early cultures commonly believed that spirits, or mindlike entities, were responsible for many natural phenomena and could inhabit inanimate objects (Lloyd, 1987; Lowie, 1948; Tylor, 1871). Such animistic concepts disappeared long ago from every area of science except one. They still survive where they began: in ideas about the operations of the nervous system. Their introspective origin (the fact that they seem self-evident to everyone) gives these concepts immense appeal. It has made neuroscience the last refuge of animism.

Deeply embedded in the fabric of everyday life, where they can be extremely useful, the functions of the mind intrude also into neuroscientific theory and experiment, where they have no \textit{a priori} right to be. As a result, the accommodation of these ancient concepts into modern neuroscience—the process through which they, like the assumption of the perfection of the celestial order, are either abandoned or modified into scientific concepts—has been prolonged and difficult and remains unfinished.

1.3 The Present Article

This article is an attempt to understand the concept of memory that operates in neuroscience today, to specify what we are studying when we study memory, to describe the strategies we use to study it, and to identify the tactics likely to be most effective at the present time. The primary impetus for this effort was a persistent uncertainty: What exactly does memory mean in a neuroscientific context? Does it have a legitimate place in neuroscience? Is such an idea needed to make sense of available data? Is it a logical basis for hypothesis and experiment? (Or should the idea of memory be discarded, as was the idea of celestial perfection?) If it is a logical basis, what experimental approaches are most appropriate now and for the immediate future?
These questions came from discomfort with the common assumption that the central nervous system (CNS) does two things: (a) it interacts with its environment, and (b) it stores information about those interactions. This stored information is “information” in the sense of a record or representation of the external world—it is about something. It includes, for example, a map of the route from home to work, a list of important events over the past year, the multiplication tables, the movements of a dance step, and the faces of friends and relatives. The assumption that such information is stored implies that there must be processes that encode, maintain, and retrieve it. These processes become the theoretical framework of memory research, and the elucidation of their realization in brain structure and function becomes the experimental goal. However, the assumption that the nervous system stores information may only be a belief imposed on neuroscience by ideas derived from introspection.

Seen from the outside, without the benefit or the impediment of introspection, the brain is simply a physical system that interacts with its environment. In this, it is no different from other physical systems, animate or inanimate. Certainly, many of its interactions involve stimuli and responses widely separated in time, but so do those of other systems, animate or inanimate. Why assume that the brain alone is also engaged in the storage of information, in the preservation of records of the past or representations of the external world? Why not begin by assuming only that the brain interacts with its environment and then try to define the processes underlying those interactions? And if this is the proper approach, does the concept of memory have a place in it? And if it does have a place, what phenomena does it embrace, and how can we best study them?

This article tries to answer these questions. Because the questions challenge ideas that have evolved over many centuries and are deeply embedded in both popular and scientific thought, addressing them requires understanding this evolution. Thus, the article begins by reviewing the ancient prescientific concept of memory and the theoretical basis of modern neuroscience and then considers the terms under which the former has been and ought to be integrated into the latter. It tries throughout to avoid assuming the concepts and conclusions that originate in introspection. The intention is to approach nervous system function as any other scientific problem is approached, that is, from the perspective of an external observer who is not influenced by ideas or assumptions derived from a private and ill-defined position inside the phenomenon under study.

This effort should not be construed as a new foray into the old controversy in psychology between behaviorist and cognitive approaches to understanding behavior. That controversy concerned the importance of considering conscious thoughts, feelings, and intentions in understanding behavior and traditionally ignored or largely ignored the brain itself. In contrast, the present focus is on understanding brain function per se, and the intention is to proceed unbiased by traditional ideas about brain function that come from introspection.

The endeavor arrives at a definition of memory different from that suggested by introspection. It is a definition that is consistent with the theoretical basis of modern neuroscience and is free of the problems that follow from the assumption of information storage. With this definition in hand, the article goes on to identify the essential components of memory and to analyze the experimental strategies used to study them. Finally, it discusses the new insights emerging from current studies and considers their implications for memory research now and in the future.

2. THESAURUS IMAGINUM (THE STOREHOUSE OF IMAGES)

2.1 The Popular Concept of Memory

At the climax of the movie *Casablanca*, when Humphrey Bogart consoles Ingrid Bergmann with the now famous line, “We’ll always have Paris,” he implicitly invokes what St. Augustine, writing in the 4th century, called the “thesaurus imaginum,” the storehouse of images. This storehouse and the images in it are personal, no one else has access to them, and thus no one else can confirm their existence. But because everyone seems to have a similar awareness, the reality of this storehouse was accepted long ago, and its nature has been the object of speculation for at least 2,500 years.

Augustine (1961) described the “thesaurus imaginum” in considerable detail:

> Memory . . . is like a great field or a spacious palace, a storehouse for countless images of all kinds which are conveyed to it by the senses . . . In it are the sky, the earth, and the sea, ready at my summons, together with everything that I have ever perceived in them . . . In my memory, too, are all the events that I remember, whether they are things that have happened to me or things that I have heard from others . . . It also contains all that I have ever learnt of the liberal sciences [and] the innumerable principles and laws of numbers and dimensions . . . In it I meet my self [sic] as well. I remember my self and what I have done, when and where I did it, and the state of my mind at the time. (Book X)

Augustine also included habits and skills as objects of memory and described the salient characteristics of remembering and forgetting.
In the course of his extensive discussion, three categories of memory emerge: memory of external objects and events, that is, memory of images conveyed by the senses; memory of skills, habits, and knowledge (such as principles of logic); and memory of previous states of mind, including past thoughts and emotions. Thus, the storehouse of memory contains records of interactions with the environment and records of internal events.

Augustine’s description is similar to earlier ones by Aristotle and other classical figures both in the image of storage and in the inclusion in memory of both interactions with the outside world and internal events such as thoughts. In 1690, John Locke (1825, pp. 101,71-72) offered a comparable discussion: “Memory . . . is . . . the storehouse of our ideas. For the narrow mind of man not being capable of having many ideas under view and consideration at once, it was necessary to have a repository to lay up those ideas which at another time it might have use of” (p. 101). For Locke, as for Augustine, the ideas that are the content of the memory storehouse can originate either externally or internally. The first source is the external world: “Our senses . . . do convey into the mind several distinct perceptions of things, according to those various ways wherein those objects do affect them . . . This great source of most of the ideas we have, depending wholly upon our senses . . . I call SENSATION” (p. 71). The second source is the “perception of the operations of our own mind within us” (p. 71), these operations furnish the understanding with another set of ideas, which could not be had from things without; and such are perception, thinking, doubting, believing, reasoning, knowing, willing, and all the different actions of our own minds . . . This source of ideas every man has wholly in himself . . . I call this REFLECTION. (p. 72)

Early in the 19th century, James Mill (1878) made a comparable distinction more succinctly: “There are two cases of Memory . . . The first is, when we remember what we have seen, felt, heard, tasted, or smelt. The second is, when we remember what we have thought, without the intervention of the senses” (p. 328). For Mill, as for Locke and Augustine and their philosophical colleagues over the centuries, memory was a storehouse containing records of previous interactions with the environment as well as records of previous internal events (i.e., thoughts). This storehouse was wholly personal. It was accessible only to a single privileged observer—the person who possessed it. It could not be observed or examined by others. Thus, from classical times until the early 19th century, memory was a philosophical issue, and the philosophers, in a rare instance of unanimity, agreed with each other.

2.2 Pre–19th Century Theories About Memory Processes

Along with these consistent philosophical analyses came speculations about the physical basis of memory. None departed from the popular concept of a storehouse. They simply went beyond frankly metaphorical descriptions like that of Augustine to consider the composition of the storehouse. These speculations drew on the dominant scientific and technical knowledge of their times and fell into two categories. Some proposed that the storehouse was composed of structural changes in the brain (or in whatever organ the theorizer believed to be the site of memory), whereas others proposed that the storehouse consisted of persistent motion or activity. Thus, in modern terms, some theories were structural or anatomical, and some were functional or physiological.

Perhaps the earliest anatomical hypothesis was that of Parmenides, who in the 6th century B.C. proposed that each experience produced a particular mixture of light or heat and dark or cold in the body, and that memories consisted of the persistence of such mixtures, while their destruction was the mechanism of forgetting (Gomulicki, 1953).

Despite the fact that he lodged memory in the heart rather than the brain, Aristotle may have been the first to propose a physiological theory of memory (Burnham, 1889; Gomulicki, 1953; Neuburger, 1981). He maintained that sensations passed from sensory organs to the heart by movements of the animal spirits, or pneuma, the agent through which the soul ruled over the body. In the heart, the movements of the pneuma became ideas. The persistence of these movements constituted memory, and their gradual diminution accounted for forgetting.

Classical theories of anatomy and physiology, including ideas about memory, culminated in the voluminous writings of the 2nd-century Roman physician, Galen (Finger, 1994; Gomulicki, 1953). He agreed with Aristotle that memory was a function of the animal spirits. However, he placed the spirits in the brain rather than in the heart. Throughout the Middle Ages and into the Renaissance, the sequential functions of sensation, idea production, and memory were usually placed front to back in the ventricles of the brain. Animal spirits passed from the body to the anterior ventricles, where they produced sensations. In the next compartment, the middle ventricle, sensations gave rise to ideas. Finally, these ideas were stored as memories in the posterior ventricle.

Speculations about the physical bases of memory finally advanced beyond Galen in the 17th century, when Descartes incorporated the concept of animal spirits into a new anatomical theory based on hydraulic principles. He believed that the immaterial soul—the single privileged observer—accessed the nervous system
through the pineal gland, which controlled the flow of animal spirits through the pores of the brain. Memories were changes in the pores. These changes modified the flow of the spirits, and this altered flow caused movements in the pineal that led the soul to remember (Descartes, 1988, pp. 233-234).

As the scientific revolution advanced, new discoveries led to new speculations about the basis of memory. In the 18th century, inspired by the relatively recent advent of Newtonian physics, David Hartley and Charles Bonnet proposed vibratory theories of memory (Finger, 1994; Gomulicki, 1953). Hartley believed that memories were stored as vibrations of particles in the white matter of the brain. “Natural vibrations,” established before birth, were subsequently modified “in degree, kind, place, and line of direction” by external stimuli (Gomulicki, 1953, p. 4). With sufficient repetition of the stimuli, the new vibrations became permanent and constituted memory traces. These memory images were identical to sensations except that they were not as strong. In a related theory, Hartley’s contemporary Bonnet proposed that sensations caused nerve fibers to vibrate and that they vibrated more readily with repeated use. “The feeling accompanying this increase in suppleness or mobility constitutes reminiscence, which acquires strength in proportion as the fibres become more supple or mobile” (Gomulicki, 1953, p. 5).

Also in the 18th century, the long history of speculation about the nature of the memory storehouse culminated in an effort to quantify its physical substance. The Swiss physiologist Albrecht Haller conducted experiments designed to determine the duration of the process necessary to produce one idea and concluded that the answer was one third of a second. Starting from this point and assuming that the brain minus vasculature and cortex (which was not thought to retain impressions) weighed 2 pounds, others calculated that after 50 years of life, each grain (i.e., 0.065g) of brain contained 205,542 memory traces (Burnham, 1889; Gomulicki, 1953; Huber, 1878).3

2.3 The Appearance of a Separate Scientific Concept

For thousands of years, the concept of memory as a storehouse of images remained stable and noncontroversial. It had arisen, along with the other functions of the mind, from introspection. These concepts were considered self-evident and, as such, were simply refined and formalized by philosophers. The only new developments in regard to memory were merely periodic upgrades in the technical sophistication of speculations about the physical substance of the storehouse.

Through the same thousands of years and in sharp contrast to this placidity, controversy and uncertainty dominated a related issue: Neither philosophers nor scientists nor the two together were able to produce a comprehensive and consistent theory as to the functions of the brain, the mysterious organ located inside the head. Although the brain seemed to be somehow connected to the functions of the mind, it was also thought to have a number of other functions, including cooling the blood and producing mucus.

In the 19th century, the situation reversed: The functions of the brain ceased to be a problem, and the functions of the mind, including memory, became a problem. A comprehensive hypothesis of brain function developed and was widely accepted. This hypothesis stated that the entire function of the brain was to connect sensory input, or experience, to appropriate motor output, or behavior. The acceptance of this hypothesis signaled the birth of modern neuroscience. Furthermore, it required that ancient ideas about the activities occurring within the brain, particularly the operations of the mind, be recast in terms of the new hypothesis or be simply discarded. In the context of the new hypothesis, the concept of memory as a personal and private storehouse was no longer acceptable. If memory was to remain a viable concept in neuroscience, it had to be redefined in a way that made it accessible to scientific study and consistent with the new hypothesis. This necessary redefinition, begun more than a century ago, remains unfinished. Understanding the development and implications of the new hypothesis of brain function is essential for understanding how the concept of memory has changed and must change further to become a useful neuroscientific concept.

3. MEMORY IN NEUROSCIENCE

3.1 The Riddle of Brain Function

In the ancient world, the function of the brain was not clear. Most cultures thought that the heart was the most important organ, and many paid little attention to the brain (Dudai, 1989; Finger, 1994; Plum & Volpe, 1987). For the Chinese, it was an auxiliary organ, less important than the five major organs, which were heart, lungs, kidneys, liver, and spleen. The Egyptians, despite their detailed case reports describing the devastating effects of head injuries, lodged consciousness and intelligence in the heart. The Old Testament never mentions the brain. According to the prophets Jeremiah and Isaiah, the soul was in the kidney and the heart.

Among the Greeks, whose thought was most central to the subsequent development of Western biology and medicine, opinions were divided (Clarke & O’Malley, 1996; Finger, 1994; Neuburger, 1981; Sigerist, 1933). Controversy centered on whether the brain or the heart was the home of perception, intellect, emotion, imagina-
tion, memory, and other functions of the mind. In the 5th and 4th centuries B.C., Alcmaeon, Plato, and others of the Hippocratic school placed intelligence in the brain (although Plato, in agreement with Democritus, maintained that emotions resided in the heart and liver). During the same period, Empedocles and Aristotle believed that intellect and perception were to be found in the heart. Aristotle adopted this position for several reasons. The brain was cold to the touch and the heart was warm, and the Greeks associated warmth with sensation and life in general. The heart responded vigorously to emotions. Furthermore, touching the brain produced no sensation, which seemed to imply that it could not be the organ of feeling or consciousness. The issue was largely resolved in favor of the brain in the 3rd century B.C. by Herophilus and Erasistratus, known respectively as the “Father of Anatomy” and the “Father of Physiology.” They worked during a brief highly productive period when the usual stricture against human dissection was lifted. Erasistratus even linked the complexity of the cerebral gyri to intelligence. Primarily as a result of their work and the subsequent work of Galen, drawings from the Roman era and continuing through the Middle Ages and into the Renaissance showed perception, intellect, imagination, and memory in various CNS locations, especially the cerebral ventricles.5

At the same time, however, the paucity of anatomical and physiological knowledge impeded formulation of testable hypotheses about what the brain does and rendered both futile and unnecessary attempts to link the functions of the mind to the structures of the brain in any substantive way. Thus, although it was generally recognized that the mind resided in the head, the mind and the brain could remain quite separate subjects for speculation. Any desire to understand the functions of the brain was largely frustrated by its delicate substance and bewildering structure. The mind remained as it always had been: the province of philosophy and theology.

This peaceful arrangement unraveled in the 17th and 18th centuries, as researchers began to make substantive progress toward localizing specific functions in the brain and philosophers turned their attention to interactions between the mind and the physical world (Finger, 1994; Locke, 1825; Neuburger, 1981; Young, 1990). Particularly important were Thomas Willis’s studies that focused attention on the brain’s regulatory roles by linking respiratory and cardiac functions to the cerebellum and adjacent structures. At the same time, he placed intellect in the cerebrum. The philosopher Locke, who had been Willis’s pupil at Oxford, stressed the central role of sensory input in the development and continuing operations of the mind. These approximately concurrent developments produced confusing and inconsistent analyses of brain function. The material brain and the immaterial mind or soul were both residents of the cranium, and the brain had two different jobs. On one hand, it interacted with its physical environment—the body and the outside world—whereas on the other hand, it interacted with the mind or soul.

This unsatisfactory situation is well illustrated in the writing of the prominent late-18th-century scientist Jirí Procháška. His book, A Dissertation on the Functions of the Nervous System (Procháška, 1784/1851), begins by summarizing these functions:

The nervous system is the seat of the rational soul, and the link by which it is united to the body; it is the instrument by which the soul, so long as it is united to the body, produces its own actions ... and by which it acts on the rest of the body, and the body in turn acts on it. But, however great may be the importance of the nervous system in these respects, it is of further importance, because it possesses in addition the singular faculty of exciting in the human body various movements without the consciousness or assistance of the soul. . . . It can and does excite them without intermission through the whole of life. The nervous system also influences other functions of the human body . . . digestion, nutrition, and secretion, which functions do not remain long undisturbed if the nerves be injured. (p. 1)

According to this representative late-18th-century formulation, the brain had two different functions. First, it was “the seat of the rational soul,” the “link” between the immortal soul and the mortal body, and the “instrument” that supported the activities of the soul. These activities comprised the functions (or faculties) of the mind. Second and separate, the nervous system had “the singular faculty” of exciting various bodily movements and controlling bodily functions “through the whole of life . . . without the consciousness or assistance of the soul.”

3.2 The Sensorimotor Hypothesis of Brain Function

In the 19th century, these two disconnected functions were brought together. New discoveries drove the transition to a single comprehensive hypothesis of brain function. This new hypothesis ignored the brain’s traditional role as the material link to the immaterial soul and focused entirely on the brain’s interactions with its internal and external environments. It incorporated the functions of the mind, or soul, only as terms applying to the most complex of those interactions.

The most important of the new discoveries was that of Charles Bell and François Magendie. In the second and third decades of the century, they showed that the posterior spinal roots were sensory, conveying input to the brain, whereas the anterior spinal roots were motor, car-
ry output to the muscles (Clarke & O’Malley, 1996; Finger, 1994). The existence of separate motor and sensory nerves had been known to Herophilus and Erasistratus 2,000 years earlier (Clarke & O’Malley, 1996; Neuburger, 1981; Sigerist, 1933). However, all their original reports were lost, and their knowledge survived only in fragmentary secondary references that had little influence.

The Bell-Magendie discovery of the different functions of the posterior and anterior spinal roots may appear at this distance less significant than achievements later in the 19th century, such as the development of the neuron doctrine or the discovery of cortical excitability. Nevertheless, it has been justly described as the most important physiological advance since Harvey’s description of the circulation of the blood 200 years earlier (Young, 1990). Bell and Magendie focused attention on the brain’s role as the interface between sensory inputs and motor outputs. Their discovery was the central event in a complex of scientific and philosophical developments that settled the ancient issue of the brain’s function and turned the old and nonspecific question “What does the brain do?” into a new and specific question: “How are the brain’s inputs connected to its outputs?”

This question contains a hypothesis of brain function. The hypothesis is that the function of the brain is to connect inputs from the internal environment (i.e., the body) and from the external environment to appropriate outputs: that the nervous system is an organ that connects stimuli—activity in sensory nerves—to responses—activity in motor nerves.

As with other successful ideas, the general acceptance of this comprehensive “sensorimotor” hypothesis of brain function over the past 150 years has obscured its fundamental importance (as well as obscuring the fact that it still remains a hypothesis, albeit a very fruitful one). Due to its widespread acceptance, contemporary textbooks do not need to state it formally. A century ago, it was still new enough to warrant explicit statement. Thus, in 1890, William James (1890/1983) began his famous textbook, The Principles of Psychology, with this paragraph:

If I begin chopping the foot of a tree, its branches are unmoved by my act, and its leaves murmur as peacefully as ever in the wind. If, on the contrary, I do violence to the foot of a fellow-man, the rest of his body instantly responds to the aggression by movements of alarm or defence. The reason of this difference is that the man has a nervous system, whilst the tree has none; and the function of the nervous system is to bring each part into harmonious cooperation with every other. The afferent nerves, when excited by some physical irritant, behave as gross in its mode of operation as a chopping axe or as subtle as the waves of light, conveys [sic] the excitement to nervous centres. The commotion set up in the centres does not stop there, but discharges itself, if at all strong, through the efferent nerves into muscles and glands, exciting movements of the limbs and viscera, or acts of secretion, which vary with the animal and with the irritant applied. These acts of response have usually the common character of being of service. (p. 25)

In his classic text of 1906, The Integrative Action of the Nervous System, Sherrington provided a physiologically oriented explication of the hypothesis. He described the CNS as the “central ‘exchange’ organ” in which “the afferent paths from receptor-organs become connected with the efferent paths of effector-organs.” Furthermore, it is not merely a meeting place where afferent paths conjoin with efferent, but is, in virtue of its physiological properties, an organ of reflex reinforcements and interferences, and of refractory phases, and shifts of connective pattern. . . . It is, in short, an organ of co-ordination [sic] in which from a concourse of multitudinous excitations there result orderly acts, reactions adapted to the needs of the organism, and. . . . These reactions occur in arrangements (patterns) marked by absence of confusion, and proceed in sequences likewise free from confusion. (p. 313)

Such formal statements of the sensorimotor hypothesis are uncommon in the contemporary literature, but do occur. An article on neural network analyses of brain function restates it:

The neurons of the central nervous system are engaged in the following three operations: (1) reception of sensory signals from outside and from within (input), (2) planning and execution of motor acts (output), and (3) intermediary processing interposed between input and output. (Mesulam, 1990, p. 610)

In these descriptions of brain function, the dichotomy evident in Procháška’s 18th-century statement is gone; there is no mention of brain/mind or brain/soul interactions separate from interactions between the brain and its internal and external environments. The brain has a single, well-defined function as “the central ‘exchange’ organ.”

In its simplest interactions with the environment, the brain connects a single stimulus, an action potential in one afferent neuron, to a single response, an action potential in one efferent neuron. Most interactions, however, involve sets of stimuli interspersed with sets of responses and are usually referred to as experiences or behaviors. “Experience” focuses on the interaction as seen from the brain and can be defined as a specific set of stimuli interspersed with a specific set of responses, whereas “behavior” focuses on the interaction as seen
from the environment and can be defined as a specific set of responses interspersed with a specific set of stimuli. The brain connects the set of stimuli constituting the experience to the concurrent set of responses constituting the behavior. It also establishes connections between interactions that are separated in time. The experience of a past interaction may be connected to the behavior of a present interaction, and the experience of the present interaction may in turn be connected to the behavior of a future interaction. For example, the interaction of meeting a stranger is both a behavior that is affected by past encounters with strangers and an experience that affects future behaviors in which the former stranger is greeted as an old acquaintance.

In accord with this new consensus as to the function of the brain, the functions of the mind—perception, imagination, volition, memory, and so forth—became terms that described the most complex connections between sensory input and motor output, between experience and behavior. One of the earliest statements of this new conception of the mind was by the psychiatrist Wilhelm Griesinger in 1845:

The psychical life of man ... commences in the organs of sense, and ... passes out again into the organs of movement. ... Between these two fundamental acts [sensory input and motor output] ... [is] an accessory sphere ... within which moves the whole mental life of the man ... all the various mental acts which were formerly designated separate faculties (imagination, will, emotions, &c.) [i.e., the functions of the mind] are only different relations of the understanding with sensation and movement. (Griesinger, 1965, p. 26)

Sixty years later, Sherrington (1906) was more specific. He pointed out that the functions of the mind are focused in those sensorimotor transformations that begin with input from what he called the “distance-receptors,” that is, eyes, ears, and nose. In keeping with the then-new theory of evolution, he attributed the existence of these functions to their importance in ensuring the survival of the organism.

By the “distance-receptors” are initiated and guided long series of reactions. ... It is in [these] reactions ... that the most subtle and complex adjustments of the animal ... arise. ... [They] allow most scope for the selection of brute organisms that are fittest for survival in respect to elements of mind. ... Nothing, it would seem, could better ensure the course of action taken ... being the right one than memory and anticipatory forecast. (Lecture IX)

In accord with the sensorimotor hypothesis—that the whole function of the nervous system is to transform sensory input into motor output—the mind is no longer an independent, nonphysical entity that interacts with the material brain through the pineal gland or some other interface. Rather, the term mind means the processes underlying the most complex connections between experience and behavior.

3.3 The Neuroscience Agenda

The general acceptance of the sensorimotor hypothesis was the beginning of modern neuroscience. Its acceptance was a scientific revolution—the advent of a new paradigm—in the sense defined by Kuhn (1962). Whether neuroscience as a field of endeavor with generally recognized goals and methods existed before or whether there were simply individual scientists who studied the brain according to their own lights is debatable. What is clear is that a coherent new neuroscience arose in the 19th century and that the first principle of its paradigm was and remains the hypothesis that the function of the brain is to connect experience to behavior. Ever since this hypothesis coalesced a century and a half ago, the assumption underlying neurobiological research has been that complete understanding of the connections between sensory inputs and motor outputs would constitute complete understanding of nervous system function. The hypothesis sets the agenda of neuroscience: to understand the sequences of events that lead from input to output, from experience to behavior. In this agenda, the functions of the mind suggested by introspection have a place only to the extent that they prove to be essential for describing the events connecting experience to behavior.

In practice, this agenda has traditionally emphasized the anatomy and physiology of the connections between experience and behavior, that is, where the sequences of events occur and what they are. Nevertheless, it embraces also the genetic, developmental, metabolic, hormonal, and environmental factors that shape and control the anatomy and physiology as well as the pathological processes that can damage or disrupt them. Thus, the sensorimotor hypothesis can be seen to be the first premise underlying all aspects of neuroscience research.

3.4 The Sensorimotor Hypothesis, the Neuroscience Agenda, and Memory

Prior to the 19th century, when memory was a faculty of the mind and the business of philosophers, the popular concept of memory as a private storehouse of images was acceptable and useful. That the storehouse was inaccessible to objective study was of no concern, for in this it was no different from other traditional subjects of philosophical inquiry.

In the 19th century, when memory, along with all the other functions of the mind, became a property of the brain and the business of scientists, the prescientific
storehouse concept became a problem. As described by Augustine, Locke, Mill, and others, it was the private possession of an individual. Therefore, it could not be addressed experimentally. It was no more accessible to scientific study than was justice or beauty. Furthermore, the assumption that the brain created a private record of remote experience was inconsistent with the sensorimotor hypothesis, which maintained that the whole function of the brain was simply to connect experience to behavior. Thus, for the concept of memory to be consistent with the sensorimotor hypothesis and for the study of memory to be included in the neuroscience agenda, the prescientific concept of the “thesaurus imaginum” had to be changed.

The solution came from one who, although trained as a philosopher, was dissatisfied with treating memory as a philosophical problem. In 1885, Hermann Ebbinghaus described objective techniques for measuring memory and thereby took the problem of memory away from philosophers and presented it to scientists (Ebbinghaus, 1913). In his well-known studies, he learned lists of nonsense syllables and at later times tested his recall by measuring the time necessary for relearning. He described the impact on his recall of the duration and ordering of initial exposure, the time elapsed since exposure, and other variables. Thus, he began to define the features of memory.

In the next decade, Edward Thorndike (1965) described similarly objective methods for measuring memory in cats and other laboratory animals. He placed an animal in a box fitted with an internal release and measured the effects of repeated exposures on the time elapsed before the animal triggered the release and escaped from the box. Thus, he initiated the development of laboratory animal models for studying memory, a development that has continued into the present.

In their experiments, Ebbinghaus and Thorndike assessed memory by measuring its impact on behavior. By this means, they provided access to what had previously been the province of a single privileged observer. It was now possible to study the storehouse and the processes that produce, maintain, and degrade its contents by measuring behavior.

Their solution integrated the storehouse concept into neuroscience. It provided the concept of memory that operates in neuroscience today: Memory is a storehouse of information that is derived from previous experience and affects behavior. Thus, contemporary neuroscientific definitions of memory do not, like the definitions of Augustine or Locke, focus on private perception of the records of experience but rather on the effects that these records exert on present behavior. How is information acquired and stored in the brain? . . . Animals and humans are capable of modifying their behavior as a result of experience. . . . What physical changes are induced in the organism by learning and what changes must occur for there to be learning? . . . We call this area of study the neurobiology of learning and memory. (Martínez & Kesner, 1991, p. xv)

Memory is the retention of experience-dependent internal representations over time . . . internal representations [are] neurally encoded structured versions of the world which could potentially guide behavior. (Dudai, 1989, pp. 5-6)

What has preserved the records of our past and is continually recalled in our present? . . . How are patterns of sensed and behaviorally expressed information stored? . . . How is stored information subsequently recalled? (Alkon, 1987, p. 1)

Learning is the acquisition of knowledge about the world. Memory is the retention or storage of that knowledge. . . . [Learning] can only be inferred from changes in behavior. (Kupfermann, 1991, p. 997)

Learning is the process of acquiring new information, while memory refers to the persistence of learning in a state that can be revealed at a later time. (Squire, 1987, p. 3)

In these contemporary definitions, the “image” of Augustine is replaced by terms such as “stored information,” “internal representation,” or “engram,” but the storehouse concept remains. It has been accommodated into neuroscience by providing it with a connection to behavior.6

This scientific form of the storehouse concept has supplied the framework for studies of human and animal memory for the past 100 years. Over this period, as in the preceding 2,500 years, theories about the nature of the storehouse have continued to undergo periodic revisions. And, as in the past, these theories have followed concurrent scientific and technical advances. Thus, in the middle of the last century, new understanding of the structure and function of nucleic acids and proteins spurred speculations that memories were coded in macromolecules (e.g., Griffith & Mahler, 1969; Monné, 1949; Nakajima & Essman, 1973; von Förster, 1948). In the same period, the possible similarities between memory in the brain and computer memory drew attention (von Neumann, 1958). Most recently, neural network principles have come to underlie hypotheses as to the nature of the memory store (Buonomano & Merzenich, 1999; Churchland & Sejnowski, 1992; deCharms & Zador, 2000; Hasselmo & McClelland, 1999; O’Reilly & Rudy, 2001; Pouget & Snyder, 2000; Sommer &
Wennekers, 2000). In these hypotheses, the basic elements of the storehouse are a standard repertoire of synaptic and/or neuronal modifications, and the contents of the store lie in the distributions of these elements.

3.5 Problems With the Storehouse

The link to behavior solves the problem posed by a memory storehouse that is accessible only to the single privileged observer within whom it resides. By tying the storehouse to observable phenomena, the link to behavior opens it to scientific study. At the same time, however, the retention of the storehouse concept preserves a prescientific assumption about how the brain connects remote experience to present behavior. The assumption is that the brain stores information about experience and that this information is consulted in formulating present behavior. This assumption is based only on introspection; it is not part of the sensorimotor hypothesis, which simply says that the brain connects experience to behavior, without specifying how.

Furthermore, as the current edition of the popular textbook *The Cognitive Neurosciences* acknowledges, “Information [cannot] be readily defined. [It] simply is the intangible, ineffable, unknown ‘stuff’ that is somehow created, transferred, transformed, preserved (‘processed’) in the mind/brain, which when appropriately ‘converted,’ determines behavior and conscious thought” (Tulving, 2000, p. 729). The assumption of the existence of this “intangible, ineffable, unknown ‘stuff,’” is reminiscent of the assumption by 19th-century physicists of the existence of the luminiferous ether, a similarly intangible substance that filled all matter and space and was the medium for transmission of electromagnetic waves (see Bartusiak, 2000; Lodge, 1925, p. 127; Preston, 1912, p. 561). The ether concept became increasingly difficult to reconcile with experimental results, and the need for the idea disappeared with the advent of the special theory of relativity (Bartusiak, 2000).

Most important, the assumption that the brain stores information creates the problems described in the next four subsections. First, it raises the question of what exactly is recorded or represented from an experience, and it implies the existence of an observer who has access to the record. Second, it ignores important aspects of the plasticity associated with the connections that the brain establishes between remote experience and present behavior. Third, it imposes a specious distinction between the processes of memory and the processes that underlie other long-term effects on nervous system function. Fourth, it dictates a simplistic theoretical framework that promotes an inadequate and cumbersome research agenda.

3.5.1 What Is Represented and to Whom. According to the definitions in the Section 3.4, the information in the storehouse of memory is not merely the “information” of information theory, which is simply the reduction of uncertainty (Pierce, 1980; Shannon & Weaver, 1964) (see Note 1 and Section 1.3). The “information” supposedly stored as memory is a representation of something else—it is about something. It is “the records of our past” or “knowledge about the world” or “neuronally encoded structured versions of the world.” This concept of memory raises difficult questions that are ultimately superfluous for neuroscience.

First, it is not clear what the information stored from an experience is supposed to be about. Is it a description of the experience, a blueprint for future behavior, a set of facts to be consulted in deciding behavior, all of these together, or something else entirely? Second, the concept of memory as representation implies the existence of an observer who can use the information, and this observer seems to be indistinguishable from the single privileged observer of Augustine’s (1961) “thesaurus imaginum” or from the soul that Descartes (1988) believed interfaced with the brain through the pineal. Alternatively or in addition, the implied observer could be the scientist who can recognize and understand the information. In either case, this present-day assumption of information storage in the nervous system differs from Parmenides’ mixtures of light and dark or Descartes’ hydraulics only in its technological sophistication, and the implied observer takes the place of the soul found in pre-19th-century theories about brain function.

Third and most important, it is not clear why information about or a representation of anything should be stored. The assumption of information storage is simply irrelevant to the sensorimotor hypothesis, which proposes that the function of the brain is to connect experience to appropriate behavior, not to store information for actual or theoretical observers. The nature of the processes underlying this function is an issue to be resolved by experiment, not by introspection or tradition. Although it seems clear that connections between remote experience and present behavior involve persistent structural or functional effects of experience on the nervous system, there is no reason to assume that these effects constitute stored information or representations of anything.

3.5.2 The Several Categories of Plasticity Underlying Memory. The memories that are presumed to be the contents of the storehouse depend on the persistent effects of experience. That is, they depend on the plasticity that experience produces in the nervous system. Prescientific concepts of memory, typified by that of Augustine (see Section 2.1), usually sequestered this plasticity; They saw the storehouse of memory as a facility that was distinct from the parts of the brain responsible for moment-to-
moment function, a facility that did nothing other than store records of experience. This “filing cabinet” conception of the storehouse has appeared most recently in the proposal that memories are encoded in macromolecules and in the parallels drawn between the storehouse and computer memory (see Section 3.4). Nevertheless, it has become increasingly clear over the past half century that the idea of the storehouse as a separate repository is no longer tenable. The plasticity that is produced by experience and that in turn affects future behavior is not limited to a few brain regions nor to specialized neuronal or synaptic populations; it can occur throughout the nervous system, and it involves the same neurons and synapses responsible for ongoing brain function. As a result, the memory of an experience is likely to involve three categories of plasticity (Wolpaw & Lee, 1989; Wolpaw & Tennissen, 2001). Only the first of these is compatible with the concept of memory as representation of experience.

The first category of plasticity, which might be called “primary,” consists of those nervous system changes that underlie the obvious or, in an experimental context, the rewarded (or intended) behavioral outcome of the experience (i.e., intended by the experimenter). In normal life situations, an experience often has several obvious behavioral outcomes. Thus, learning a skill produces the plasticity responsible for later exercise of the skill and also produces the plasticity responsible for later recollection of the learning experience. For example, a person who learns a new game can play the game later and can also state the time, place, and other circumstances of the learning experience. These behaviors are the products of primary plasticity. If experience produced only primary plasticity, that is, if it produced only the plasticity responsible for the obvious or intended change(s) in behavior, the concept of memory as representation of experience might be adequate. However, this is not the case. A second category of plasticity is essential to preserve other behaviors, and a third category is inevitable due to the ubiquity of activity-driven plasticity in the nervous system.

The effects of experience that constitute primary plasticity do not normally occur at sites devoted solely to the rewarded or intended behavior. Most neurons and synapses in the CNS participate in numerous behaviors. Thus, the primary plasticity created by an experience is almost certain to disturb other behaviors, and this disturbance will necessarily generate additional or “compensatory” plasticity that preserves or restores these behaviors. For example, for a rat, learning a new maze is likely to affect the processes responsible for transit through a previously learned maze, especially if the mazes are similar in appearance but different in organization. Preservation of performance in the original maze will require additional or altered processing of sensory inputs and changes in the activity governing selection and sequencing of motor outputs. Similarly, for a human, learning a new language is likely to affect the processes underlying use of a previously learned language. Continued use of the earlier language despite acquisition of the new language is likely to require alterations in the responsible brain activity, particularly when words, meanings, or structures in the two languages are very similar or very different. Such alterations can only result from additional, compensatory plasticity. Primary and compensatory plasticity may presumably depend on similar processes occurring in comparable locations. They are most readily recognized and differentiated by their effects on behavior.\footnote{The ubiquity in the nervous system of the capacity for activity-driven change ensures that primary and compensatory plasticity, by chronically modifying the activity reaching other sites in the nervous system, will inevitably result in additional plasticity, which can be labeled “reactive.” Unlike compensatory plasticity, which occurs in response to the disturbances in the interactions between the brain and its environment caused by primary plasticity, reactive plasticity results from events wholly within the nervous system. For example, when primary plasticity increases the tonic firing rate of a neuron, the increased activity may decrease the effectiveness of its synapses on other neurons by desensitizing postsynaptic receptors (e.g., Huganir & Greengard, 1990). Like compensatory plasticity, such reactive plasticity may bear no obvious relationship to the behavior that is the obvious or intended outcome of an experience. It might have no apparent effect on behavior, or it might even appear to be maladaptive (e.g., reactive postsynaptic desensitization might reduce the beneficial behavioral effects of increased presynaptic activity).}

Because reactive plasticity may eliminate or distort the behavioral effects of primary or compensatory plasticity, it may generate additional plasticity. Furthermore, the plasticity associated with growth, aging, and trauma is likely to compel continual adjustments in the primary plasticity underlying a particular connection between remote experience and present behavior, and this should in turn trigger further compensatory and reactive plasticity. In sum, an experience is likely to produce a complex pattern of plasticity that changes over time and connects to present behaviors in a variety of different ways. Indeed, some of these connections may not affect present behaviors at all; they may simply affect how these behaviors are produced. It appears to be neither reasonable nor useful to try to view this complex plasticity as a representation of the original experience.
3.5.3 Memory Versus Other Long-Term Effects on Nervous System Function. As noted above, research over recent decades has destroyed the traditional image of the hard-wired nervous system that changes rarely and only at certain sites. It is now clear that plasticity is the rule rather than the exception, that it occurs continually throughout the CNS, and that it is driven by activity. The state of the nervous system at any point in time is a product of recent and remote activity, and the persistence of that state, as well as any change it undergoes, depends on present and future activity. Both short-term plasticity, such as synaptic facilitation and post-tetanic potentiation (PTP), and long-term plasticity, such as long-term potentiation (LTP) and long-term depression (LTD), can occur at numerous locations from autonomic ganglia and spinal cord to hippocampus and neocortex (Calabresi, Pisani, Mercuri, & Bernardi, 1996; Fisher, Fischer, & Carew, 1997; Froc, Chapman, Trepel, & Racine, 2000; Maren, 2000; Maren & Baudry, 1995; Muir & Steeves, 1997; Oda, Charpier, Murayama, Suman, & Korn, 1995). Even the stereotyped plasticity that occurs during early development, once thought to be largely of genetic origin and independent of external influence, is often critically dependent on the interactions of genetic endowment with activity generated by sensory inputs (e.g., Sharma, Angelucci, & Sur, 2000; von Melchner, Pallas, & Sur, 2000). Furthermore, the synaptic, biochemical, and hormonal mechanisms of developmental plasticity underlie the plasticity that occurs during learning and in response to trauma throughout life (Bailey, 1999; Cusick, 1996; Doherty, Fazeli, & Walsh, 1995; Garcia-Segura, Chowen, Parducz, & Naftolin, 1994; Gu, 1995; D. G. Jones & Harris, 1995; Mceachern & Shaw, 1996; Rutishauser & Landmesser, 1996; Schwab, 1996; Spitzer, Vincent, & Lautermilch, 2000; Stiles, 2000; Sur & Leamey, 2001; Sweatt, 2001; Woolf & Doubell, 1994).

Impressive evidence for the close relationships among learning-related plasticity and other plasticity has come from the many recent studies of activity-driven change in the primary motor and sensory cortices of animals and humans (Buonomano & Merzenich, 1998; E. G. Jones, 2000; Kaas, 2000; Nudo, Plautz, & Frost, 2001; Sanes & Donoghue, 2000). This work has dissolved the concept of the cortical sensory and motor homunculi as fixed internal representations of the body’s sensory and motor interfaces with the outside world. It has transformed the homunculi from stable features of cortex into features that are continually maintained and modified by activity throughout life.

The primary observation is that peripheral or central events that affect neuronal activity in sensorimotor cortex and associated thalamic and brainstem regions can modify the cortical sensory and motor topographies: A point in sensory cortex that previously responded most to stimulation of one finger can come to respond more to stimulation of another, and electrical or magnetic stimulation of a point in motor cortex that previously most excited one set of muscles can come to excite another set more. Both cortical and subcortical events underlie such changes, and these events are thought to include intracortical LTP and/or LTD, modifications in N-methyl-D-aspartate (NMDA) and other receptors, release of previously silent intracortical connections from GABAergic inhibition, transneuronal degeneration, and other as yet obscure processes.

In the present context, the relevant point is that these processes and the changes they cause in cortical sensory and motor topographies can result from experience (e.g., stereotyped sensory input, repetitive performance of a particular movement, practice and acquisition of a new sensorimotor skill), from peripheral lesions such as nerve transection or limb amputation, or from ablation of specific CNS pathways or structures. There is no reason to make a mechanistic or theoretical distinction between changes in cortical topographies due to actual experience and those due to peripheral or central lesions.

As such examples illustrate, to say that the nervous system stores representations of experience makes no more sense than to say that it stores representations of development, growth, aging, or trauma. The changes produced by experience do not have a unique status; they are simply part of a spectrum of activity-driven plasticity. Thus, the concept of memory as information storage or representation of experience imposes a meaningless distinction between the plasticity produced by experience and that produced in other ways.

3.5.4 The Misdirection Of Theory and Experiment. The concept of memory as information storage implies that information is encoded, maintained, and retrieved. Since the original studies of Ebbinghaus and Thorndike, these secondary concepts have been the framework for psychological theories and experiments seeking to understand the effects of remote experience on present behavior. At the level of behavior, the concepts are adequate: The encoding and maintenance of a record of an experience and its retrieval as a behavior, whether an athletic performance or a verbal description of a past event, can be inferred and characterized by measuring behavior. However, these concepts are not adequate when the goal is to describe the effects of experience on the brain and to understand the impacts of these effects on behavior.

Even the simplest experience is likely to have a variety of more or less persistent effects, each with its own anatomical, physiological, and biochemical mechanisms. For example, as noted in the previous section, the expe-
perience-driven changes in cortical sensory and motor topographies reflect multiple forms of plasticity at multiple levels. Furthermore, each of the behavioral effects of an experience is likely to reflect multiple forms of plasticity at multiple levels. Although each behavioral effect may itself be conveniently described in terms of encoding, maintenance, and retrieval, the investigation of the neural events underlying the behavior must inevitably confront the existence of multiple processes.

In response to the multiple behavioral effects of an experience, the concept of memory as information storage has promoted the hypothesis that these different effects reflect different sets of encoding, maintenance, and retrieval processes (Kim & Baxter, 2001; Rolls, 2000; Squire & Zola-Morgan, 1991; Stanton, 2000). Thus, an experience, such as exposure to a puzzle, is said to be encoded as an “implicit” or “procedural memory,” that is, the ability to solve the puzzle, and also encoded as an “explicit” or “declarative memory,” that is, the knowledge of when and where the experience occurred. This approach to the multiple effects of an experience—a different set of encoding, maintenance, and retrieval processes for each behavioral effect—recalls the ancient effort of Ptolemy to explain the elliptical orbits of the planets as combinations of multiple circular orbits (Boas, 1962). Just as Ptolemy assumed that circles are the basic units of a planet’s motion, this approach assumes that representations that are encoded, maintained, and retrieved as particular behaviors are the basic units of the memory of an experience. By assuming the existence of multiple representations of experience, this Ptolemaic solution compartmentalizes research into separate investigations of the encoding, maintenance, and retrieval of each one.

Memory encoding, maintenance, and retrieval make sense only at the behavioral level. At the level of brain function, there are simply activity-driven processes (e.g., long-term potentiation and depression, synaptic growth, gene activation, etc.) that may reflect experience and may contribute to behavior. Each of these processes has its own requirements, dependencies, and time constants, and each may affect different behaviors in different ways at different times. All these are the proper objects of memory research and their elucidation will contribute to understanding how particular experiences affect particular behaviors. To search instead for the information that is encoded from an experience, maintained in the brain, and retrieved to produce a behavior is to pursue a phantom no more real than the luminiferous ether of 19th-century physics.

3.6 The Implications for Neuroscience

In sum, the ubiquity of the capacity for activity-driven plasticity in the nervous system, the continual occurrence of such plasticity necessitated by the need to maintain a large repertoire of behaviors, the inevitable complication of reactive plasticity, and the additional plasticity imposed by growth, aging, and trauma ensure that any remote experience contributes to a continually changing set of behaviors and that any behavior reflects a numerous and continually changing set of remote experiences. This reality provides neither place nor need for the concept of information storage or representation of experience in the nervous system.

Recognition of this situation is not important in everyday life, nor is it necessarily essential for analyzing or predicting the relationships between experience and behavior from the outside, that is, without examining the processes occurring within the nervous system. For these purposes, the simple assumption that the brain stores some record of an experience, that it derives from the experience specific items of information about the world to be used in determining future behavior, is acceptable and even useful. Many other similarly false assumptions—that the sun rises and sets, that the table supporting my computer is a solid object rather than mostly empty space—are similarly sufficient and convenient for everyday life. Recognition of the actual nature of the plasticity produced by an experience—recognition of the existence of primary, compensatory, and reactive effects and of the continual interactions between them as growth, aging, trauma, and new experiences occur—is essential only in neuroscience, that is, when the goal is to define the processes occurring within the nervous system and to understand their roles as connections between input and output, between experience and behavior.

For neuroscience, it is not useful or even meaningful to speak of the storage or representation of experience in the nervous system. An experience simply has a variety of transient and persistent effects on the nervous system. These effects and the roles they play in subsequent behaviors depend on previous experiences and are continually modified by subsequent experiences and other influences, such as growth, aging, and trauma. Thus, a change in transmitter concentration, receptor density, synaptic number, dendritic branching, or neuronal membrane properties caused by an experience may be eliminated or augmented by later experience. Furthermore, depending on the impact of associated compensatory and reactive plasticity, the original change and any subsequent modification in it may or may not affect the behavior that was the intended or obvious outcome of the original experience and may or may not also affect other behaviors or the manner in which other behaviors are produced.

Within the nervous system, the plasticity produced by a learning experience comprises a shifting set of effects...
at a shifting set of sites and contributes to a shifting set of behaviors. There appears to be nothing in this plasticity that qualifies as a record or representation of the experience. And without these terms, that is, without the storehouse concept, the terms memory encoding, maintenance, and retrieval also lose their meaning. They may be useful in everyday life and in the study of behavior without reference to the brain, but they are useless and misleading for studies of the nervous system. They are misleading because they divert hypothesis and experiment from the search for the connections between experience and behavior and focus them instead on a search for persistent and stable records of experiences (i.e., the contents of the memory store) and for processes that create, maintain, access, and degrade these records. They are misleading because they ignore those effects of experience that are compensatory or reactive. And they are misleading because they imply a distinction between the mechanisms underlying the long-term effects of experience and those underlying the long-term effects of other influences, such as growth, aging, and trauma.

For the concept of memory to accommodate the ubiquity of activity-driven plasticity in the nervous system and the complexity of the long-term effects of experience, it must be freed of assumptions derived from the prescientific concept of the storehouse and defined purely in terms of the connections between experience and behavior, that is, in terms of the sensorimotor hypothesis. The rest of Section 3 describes such a definition, and Section 4 analyzes, in terms of this definition, the essential features of memory and the principal strategies of memory research. In the course of these sections, the problems raised in Section 3.5 are addressed further.

3.7 Memory as Part of the Spectrum of Experience-Behavior Connections

According to the sensorimotor hypothesis, the nervous system connects experiences to appropriate behaviors, and that is all it does. In the simplest examples, such as pupillary constriction in response to a flash of light, the behavior is strongly connected to a single experience and follows it after a very short and predictable delay. However, few behaviors are so strongly connected to a single experience; most are connected to more than one (indeed, even the pupillary response to a light flash is connected to, i.e., affected by, earlier light exposure). In Sherrington’s (1906) words, the brain produces behavior—“orderly acts, reactions adapted to the needs of the organism”—by coordinating “a concourse of multitudinous excitations” or “experiences” (p. 313). Although some of these “excitations” immediately precede the behavior, others occur in the more or less remote past.

In terms of the sensorimotor hypothesis, which says that the entire function of the brain is to connect experience to behavior, a behavior is connected to all the experiences that contribute to it, and observation indicates that these experiences fall on a spectrum going from those that immediately preceded the behavior to those that occurred a short time ago to those that occurred long ago. For example, when a rat placed at the entrance to a maze moves unerringly through it to reach food at the other end, its performance is connected to the recent experience of being placed at the entrance and being hungry and is also connected to the experience days before when it wandered aimlessly through the maze and encountered food at the end. Similarly, when a person greets by name an old acquaintance, the greeting is connected to the person’s sudden appearance and also connected to the introduction long before in which the name was first encountered.

By this analysis, which does not mention information storage, the phenomena ascribed to memory are the connections that fall at the far end of the spectrum, connections that span longer periods of time. Memory becomes the term applied to the connections between remote experience and present behavior, the term applied to those sequences of events that bridge a substantial period of time. This definition is implicit in the experiments of Ebbinghaus, Thorndike, and their successors. Their experiments are called memory experiments because they studied the effects on behavior of experience that occurred in the past. The identifying feature of memory phenomena is the long time between experience and behavior, not the participation of a hypothetical storehouse of information that is encoded, maintained, and retrieved.

Defined in this way—as the connections between remote experience and present behavior—memory is a subset of all the connections the brain establishes between experience and behavior, and learning is the process or processes that create this subset. These definitions of memory and learning cover the phenomena to which the terms are traditionally applied and conform to the sensorimotor hypothesis. Viewed from the starting point of the experience, which may contribute to multiple subsequent behaviors, memory is responsible for those contributions that do not follow immediately. Viewed from the ending point of the behavior, which usually reflects multiple previous experiences, memory is responsible for the contributions of those experiences that occurred long before the behavior.

The traditional vantage point from which to study the nervous system’s production of a behavior is the time of, or the time immediately preceding, the behavior. Underlying this tradition is the assumption that behaviors are generated in the present, that they are products of the
3.8 The Difficulty of Memory Research

When memory is defined as the processes underlying connections between remote experience and present behavior, understanding memory means understanding the entire sequence of events leading from remote experience to present behavior. This is among the most difficult problems confronting neuroscience research. The difficulty arises from the combination of three factors. The first is the complexity and relative inaccessibility of the nervous system, particularly in humans and other vertebrates. The second is that the defining feature of memory—the temporal separation between experience and behavior—ensures that the sequence of events leading from the one to the other may occur anywhere or everywhere in that complexity. The third is the still primitive state of knowledge about the nervous system. Although important principles and considerable basic information have emerged over the past 150 years, understanding of brain function is still limited and fragmentary. As the next section discusses, recognition of and accommodation to these factors are crucial for the design and execution of effective memory research. At present and for the immediate future, designs based on the simplest possible memories are likely to prove most productive, whereas investigations of complex memory phenomena may produce little of lasting value.

4. STUDYING MEMORY

4.1 The Problems of Distance and Time

About the year 1500, European explorers began to search for the Northwest Passage, the legendary water route that led across the wilderness of North America to the lush Orient (Crouse, 1934; Thomson, 1975). The explorers knew where they were starting and where they were going, but the way between proved elusive nonetheless. Four hundred years passed before Roald Amundsen (1908) became the first to follow the Northwest Passage all the way from the Atlantic to the Pacific. Since the emergence of the sensorimotor hypothesis less than 200 years ago, neuroscience has been engaged in a similar search, a search for the passages from stimulus to response, from experience to behavior. In that time, much has been learned about the beginning—the generation of sensory inputs and their progress to and through successively higher levels—and about the ending—the activity in motor cortex and related areas that passes to spinal motoneurons and coalesces into behavior. However, like the Northwest Passage, the middle stages, the events through which sensory input becomes motor output, have proved far more difficult to locate and define.

For only a few simple examples is the full passage from experience to behavior largely understood. Thus, in vertebrates, the connection leading from sudden muscle stretch to the earliest component of subsequent muscle contraction is known because the separation between experience and behavior is short in distance, or number of neurons (i.e., two), and is also short and predictable in time. Consequently, it has been described anatomically and physiologically, and the relationship between the experience and the resulting behavior has been defined quantitatively as well as qualitatively (Henneman & Mendell, 1981; Matthews, 1972).

Such examples are unusual. For most experience-behavior connections, the delay is longer, implying that...
more complex neuronal pathways are involved, and more variable, implying that other factors are influential. Longest and most variable are the connections described as memory. Latency is often limited only by the lifetime of the organism. The long and variable latency between experience and behavior ensures that any or all of the neurons in the brain may participate and that the events that constitute their participation are hard to detect. Experiences essentially disappear into the brain to reappear at some later time in behaviors.

At the same time, however, it is clear that the sequence of events comprising a connection between remote experience and present behavior contains events of two distinct types. These two kinds of events are here called “persistent” and “junctional,” and they are described in the next section. They are the key features of memory, and as discussed in subsequent sections, they underlie the two principal strategies that have developed for studying its neural mechanisms.

4.2 Persistent Events and Junctional Events

The elementary sensory and motor processes that are respectively the beginnings and endings of experience-behavior connections are usually transient events. However, the long time between experience and behavior is the distinguishing feature of memory. That one or more of the sequential events connecting these extremes be persistent. For example, the initial transient events might lead to a persistent change in enzyme activation that causes a persistent increase in the transmitter store at a particular synaptic terminal. This lasting change in activation and the increased transmitter store that it produces could be a persistent event that bridges the bulk of the time between experience and behavior. Persistent events are the first essential feature of memory.

That memory requires persistent events has been recognized at least since Parmenides 2,500 years ago. From the earliest times, as noted in Section 2.2, theories about these events have fallen into physiological and anatomical categories, according to whether the theories propose that the events are functional or structural changes. Physiological theories evolved from Aristotle’s belief that memories were continuous movements of animal spirits in the heart, through Hartley’s theory that they were vibrations of particles in the white matter of the brain, to the 20th-century hypothesis that they were patterns of excitation continuously cycling around closed loops of neurons (Lorente De Nó, 1934, 1938; Rashevsky, 1938). (The last theory, although biologically plausible, is not consistent with the observation that memory can survive the profound disruption or complete cessation of neuronal activity associated with generalized seizures, pharmacological depression, or deep coma of other origin.)

This would seem to leave the field to anatomical theories, which usually postulate changes in synaptic number, size, or distribution, and more recently have come to encompass changes in neurotransmitter receptors and other ultrastructural features. However, as analysis proceeds from neuronal to subcellular to molecular levels, the distinction between anatomical and physiological (or structural and functional) theories evaporates. What is structure at one level is function at a more reduced level (Churchland, 1986). Although a change in synaptic level is a structural change at the neuronal level, at the intraneuronal or synaptic level it is a functional change in the biochemical processes that determine and maintain synaptic size. Similarly, a change in protein phosphorylation is a structural change from the perspective of the protein and a functional change from the perspective of the enzyme activation that is responsible. The traditional distinction between structural and functional—or anatomical and physiological—theories of memory has little meaning. This fact underlies the widespread adoption of the general term plasticity to refer to lasting change in the nervous system. Using this term avoids the specious distinction between structure and function. In regard to memory, plasticity means the persistent events that underlie connections between remote experience and present behavior.

Although persistent events have long been recognized as an essential feature of memory, memory has a second equally essential feature that has been less well recognized and has only recently become a central factor in research design. A behavior that is the termination of a connection from a remote experience—a behavior that displays memory—does not occur at random times. It occurs only when, in addition to its connection to a remote experience, it has a connection to a recent experience. This recent experience determines when the behavior occurs. A rat traverses a maze only when it is placed at the entrance, a person greets by name an old acquaintance only when that individual appears, a student answers a question only when the question is asked. A behavior that displays memory is normally a product of recent as well as remote experience.

This means that at some point or points in the nervous system, the connections from remote and recent experiences merge in one or more “junctional events” from which proceed single sequences of events leading on to behavior. The long connection from the remote experience begins earlier than the short connection from the recent experience. Thus, as pictured in Figure 1, a junctional event occurs when and where the connection from the recent experience overtakes that from the remote experience. For example, when the persistent
event is maintenance of an increased transmitter store in a synaptic terminal and the recent experience produces an action potential that reaches the terminal, the junctional event is increased transmitter release, which, by causing the postsynaptic neuron to fire, can produce behavior. Junctional events are the second essential feature of the connections from remote experience to present behavior that constitute memory.

This description of memory differs from its traditional description as the encoding, maintenance, and retrieval of information about experience. The traditional description came originally from introspection, which seems to reveal to each person a personal “storehouse of images” that are encoded from experience, maintained over time, and sometimes retrieved to influence behavior. This traditional assumption implies that the main goal of memory research is to locate and describe these records. In contrast, the description of memory in terms of persistent and junctional events comes not from introspection but rather from the sensorimotor hypothesis that the whole function of the nervous system is to connect experience, recent or remote, to appropriate behavior. This hypothesis implies that the task of memory research is to define the connections that span substantial time. It changes the research goal from finding the records of experience to explaining how remote experience affects present behavior.

4.2.1 The Taxonomy Of Memory. The standard taxonomy of memory reflects the range of the remote and recent experiences that join in junctional events and the range of the possible relationships between these experiences. The distinction between associative and nonassociative memory is illustrative. In associative memory, a behavior is linked to at least two different and specific experiences, one remote and one recent. Thus, in classical (or Pavlovian) conditioning, the remote experience is one in which a sensory input, called a conditioned stimulus (e.g., a tone), is followed closely by another sensory input, called an unconditioned stimulus (e.g., food); the recent experience is the most recent presentation of the conditioned stimulus alone, and the behavior, called a conditioned response (e.g., salivation), is connected to both the remote and recent experiences. The past sequence of conditioned and unconditioned stimuli produces the persistent event(s) that combine with the recent conditioned stimulus to produce the junctional event(s) that lead to the conditioned response. In operant (or instrumental) conditioning, the remote experience consists of the combination of exposure to a specific environment and the occurrence of a specific behavior followed quickly by a reward (e.g., a rat is presented with a bar press, presses it by chance, and receives food), the recent experience is reexposure to the same environment, and the behavior (e.g., more frequent bar presses) is linked to both experiences.

Nonassociative forms of memory, on the other hand, do not fulfill the criterion of two different and specific experiences. Thus, in habituation, the remote and recent experiences are identical: For example, the repeated presentation of a loud sound decreases the arousal that immediately follows each one. The past occurrences of the stimulus produce the persistent event(s) that combine with the most recent occurrence to produce the junctional event(s) that lead to the behavior (i.e., reduced arousal). In sensitization, the remote experience is specific, but the recent experience is not. The past presentation of a specific sensitizing stimulus (e.g., a painful electric shock) produces the persistent event(s) that combine with a wide variety of recent stimuli to produce the junctional event(s) that lead (s) to an abnormally strong response (e.g., inappropriate startle).

4.2.2 The Distinction Between Voluntary and Reflex Behavior. Prior to the 19th century, what is now called “voluntary behavior” was considered to be the product of the rational soul, which used the nervous system as a conduit for producing “its own actions” and what is now called “reflex behavior” was considered to be the product of interactions between the nervous system and its environment that occurred “without the consciousness or assistance of the soul” (Procháska, 1784/1851) (see Section 3.1). Although overt references to the soul have disappeared from neuroscience, the original distinction between reflex and voluntary persists: Voluntary behavior is most often defined as conscious behavior (e.g., Prochaska [apparently no relation], Clarac, Loeb, Rothwell, & Wolpaw, 2000). However, consciousness is directly accessible only to introspection. Its supposed behavioral manifestations are readily described in much more limited and tractable terms such as attention or perception, and thus, its value as a neuroscientific concept is at best uncertain. In reality, the accommodation into
neuroscience of the traditional distinction between voluntary and reflex behaviors rests on the degree to which they depend on junctional events.

The sensorimotor hypothesis has no place for the soul and “its own actions” nor for consciousness as a phenomenon distinct from behavior and accessible only to introspection. The whole function of the nervous system is to produce behavior, and behavior is the product of experience combined with the characteristics and capacities of the nervous system as specified by genetic endowment and shaped by subsequent growth, trauma, aging, and other influences. These characteristics and capacities (which include the mechanisms that generate stereotyped outputs such as locomotion and the sleep/wake cycle) define the substrate of neural structure and activity that processes sensory inputs into motor outputs. In this conceptual setting, behavior is the product of experience, and the difference between voluntary and reflex behavior is the timing of the experience relative to the behavior.

Reflex behaviors are the products of recent experience: If recent experiences are known, reflex behaviors are predictable. For example, a tendon tap normally ensures that muscle contraction will occur a short and predictable time later. In contrast, voluntary behaviors are the products of both recent and remote experience: They cannot be predicted simply from knowledge of the present situation. Recent experience serves only a permissive or enabling function. For example, an old acquaintance’s sudden appearance introduces the possibility of a greeting, but it does not determine what the greeting will be or even ensure that it will occur. The behavior (i.e., the greeting or the lack of one) is shaped by an extensive and incompletely defined body of remote experience, including the original introduction, subsequent interactions, the many experiences that determine an individual’s characteristic interactions with others (i.e., personality), the experiences with both internal and external environments that control mood, and so forth. Recent experience, in the form of the acquaintance’s appearance, leads to the junctional events through which these remote experiences produce behavior. The dominant role of remote experiences and, thus, the importance of junctional events is still greater in other voluntary behaviors, such as the composition of this article. Recent experience—the availability of time and essential equipment—provides mainly the opportunity for the behavior, but it has little role in determining what is written.

The distinction is not absolute—reflex behaviors such as the tendon jerk can be affected by remote experience (Meyer-Lohmann, Christakos, & Wolf, 1986; Nielsen, Crone, & Hultorn, 1993; Segal & Wolf, 1994; Wolpaw, 1997; Wolpaw, Braitman, & Seegal, 1983), and voluntary behaviors such as a greeting or the composition of a manuscript reflect recent experience (e.g., the loudness of a greeting or the pace of writing may be affected by aspects of the immediate environment, such as ambient noise). Nevertheless, the presence and importance of junctional events are the mark of voluntary behaviors. These behaviors are determined in large part by remote experiences, and they are mediated by the connections called memory.

### 4.2.3 Prejunctional and Postjunctional Parts of Memories

In principle, junctional events could occur at any place from the primary sensory neurons to the spinal motoneurons. In fact, the events near the input side are usually limited to the time of the experience, whereas those near the output side usually do not begin until just before the behavior. Thus, junctional events normally occur in the still-undefined middle portions of the connections that lead from remote experiences to behavior.

A junctional event divides the connection constituting memory into two parts: the part from remote experience to junctional event and the part from junctional event to behavior (i.e., Figure 1). The first, or prejunctional, part is unique to the memory and contains the persistent event(s) that bridge most of the lengthy time between remote experience and behavior. The second, or postjunctional, part of a memory is shared with the short connection from a recent experience. It begins when and where the two connections meet in a junctional event, an event different from what either would produce alone. For example, when the persistent event produced by a remote experience is increased neurotransmitter content in a particular synaptic terminal and a recent experience produces an action potential in the presynaptic neuron, the junctional event is greater transmitter release than would have occurred without the remote experience. It is an event that is connected to, or depends on, both the recent and the remote experience. By triggering an action potential in the postsynaptic neuron, this junctional event can initiate a postjunctional connection to behavior.

Most behaviors are connected to many remote and recent experiences and thus depend on many persistent events and many junctional events. Nevertheless, because persistent and junctional events are the two essential and distinctive components of memory phenomena, they define the research strategies aimed at elucidating memory.

### 4.3 Research Strategies

Current memory research embraces a large and varied set of endeavors. It ranges from evaluations of pathological or surgical lesions, to metabolic and electrophysiological imaging studies, to analyses of
pharmacological and genetic manipulations, to \textit{in vivo} and \textit{in vitro} investigations of specific brain regions at neuronal, synaptic, and subsynaptic levels to computer-based modeling efforts. Although these endeavors are often described in and constrained by the introspective rhetoric of information storage and retrieval, they are all aimed in one way or another at elucidating connections between remote experience and present behavior. Furthermore, when they are analyzed in terms of the distinctive features of memory—persistent and junctional events—they can be seen to comprise two distinct research strategies: an older correlational strategy that focuses on persistent events that are correlated with memory and a newer mechanistic strategy that focuses on junctional events that directly underlie behavior.

The correlational strategy is the logical outcome of the concept of memory as information storage. That concept implies that the task is to determine what the records of experience are and how they are created and maintained, and it encourages the expectation that once these records are defined, the relationship between them and behavior, that is, the manner in which they are accessed (or retrieved) by recent experience, will be straightforward. Expressed in terms of the concept of memory as connections from remote experience to present behavior, the correlational strategy is an effort to define the prejunctional portions of the connections, particularly the persistent events that are correlated with memory and essential to it.

In contrast, the mechanistic strategy is a logical outcome of the concept of memory as connection between remote experience and present behavior. That concept implies that the primary task is to define the processes through which remote experience affects behavior rather than to identify hypothetical records of experience. It focuses attention on the junctional events in which recent and remote experience join to produce a behavior different from that produced by recent experience alone. Implicit in this strategy is the expectation that once the junctional events are known, their connections to remote experience, including the persistent events constituting the key element of these connections, will become accessible. Thus, this strategy, by its retrograde progression from junctional events to the persistent events preceding them, eventually encompasses the persistent events that are the focus of the correlational strategy. Persistent events are encompassed purely in terms of an effort to explain behavior, not in terms of an effort to define records of experience.

\subsection*{4.4 The Correlational Strategy, or the Search for Persistent Events}

The goal of this approach, expressed in the terminology developed in the preceding sections, is to isolate and identify the parts of experience-behavior connections that are unique to memory—those that are prior to the junctional events and contain the persistent events—and to define these persistent events. Over the past century, this goal has been addressed by studying clinical or experimental lesions that seem to affect only memory, by using various imaging methods to detect prejunctional events, and by identifying phenomena in the nervous system that have properties that qualify them to serve as persistent events.

\subsubsection*{4.4.1 Localization By Lesions.} The oldest form of the correlational strategy uses lesions to determine the location and/or nature of the events comprising the prejunctional portions of the connections that constitute memory. It bore its first fruits in 1887, when Sanger Brown and Edward Albert Schäfer at University College London completed studies of the effects of temporal and occipital lobe lesions in monkeys and the Russian psychiatrist Sergei Korsakoff published the first of a series of papers describing the distinctive deficits of peripheral and central neurological function in a series of patients, most of them suffering from alcoholism.

Among their 13 monkeys with various unilateral and bilateral lesions, S. Brown and Schäfer (1888) described a “fine, large, active \textit{Rhesus} monkey, $\mathcal{E}$” in which the right and then the left temporal lobes had been completely removed in two operations five days apart.

These severe operations were recovered from with marvelous rapidity, the animal appearing perfectly well [the day after] the second lesion. . . . A remarkable change is, however, manifested. . . . Every object with which he comes in contact . . . appears strange and is investigated with curiosity. . . . And even after having examined an object . . . with the utmost care and deliberation, he will, on again coming across the same object accidentally even a few minutes afterwards, go through exactly the same process, as if he had entirely forgotten his previous experiments. (pp. 310-311)

Subsequent work has shown that the deficit produced by such extensive temporal lobe lesions is complex, involving inappropriate behaviors in addition to defects of memory (Kluver & Bucy, 1939). Nevertheless, in the present context, the primary import is that bilateral temporal lobectomy leaves connections from recent experience (which in this case is experience occurring in the last few seconds) to behavior relatively intact while severely disrupting connections from remote experience (in this case, experience occurring a few minutes ago). This implies that events essential for and unique to memory occur in the temporal lobes.

Korsakoff described patients with a memory defect reminiscent of that found in S. Brown and Schäfer’s monkey:
At first . . . the patient gives the impression of a person in complete possession of his faculties . . . Only after a long conversation . . . one may note that . . . he remembers absolutely nothing of what goes on around him: he does not remember whether he had his dinner, whether he was out of bed. On occasion the patient forgets what happened to him just an instant ago: you came in, conversed with him, and stepped out for one minute: then you come in again and the patient has absolutely no recollection that you had already been with him. . . . In conversation, [he] may repeat the same thing 20 times, remaining wholly unaware that [he is] repeating the same thing in absolutely stereotyped expression. It often happens that the patient is unable to remember those persons whom he met only during the illness, for example, his attending physician or nurse, so that each time he sees them, even though seeing them constantly, he swears that he sees them for the first time. (Victor & Yakovlev, 1955, p. 398)

Thus, although their short-term interactions with the environment remain apparently intact, these patients display a striking memory deficit, known as Korsakoff’s psychosis. This impressive deficit is not, however, complete. Certain effects of remote experience remain. In a paper first published in 1911, Claparède (1911/1951) described an interaction with a patient that might not be permitted under present-day guidelines:

I stuck her hand with a pin hidden between my fingers. The light pain was . . . quickly forgotten . . . a few minutes later she no longer remembered it. But when I again reached out for her hand, she pulled it back in a reflex fashion, not knowing why. (p. 69)

Clearly, certain connections between remote experience and present behavior remain in such patients. Clinicopathological correlations eventually linked the memory deficit of Korsakoff’s psychosis to lesions of diencephalic areas, dorsomedial thalamic nuclei, and/or mammillary bodies, which receive temporal lobe outputs (Victor, Adams, & Collins, 1989).

These early animal and human studies and many later clinical and laboratory investigations have shown that medial temporal lobe and related diencephalic structures are essential for certain kinds of memory. They have promoted the distinction between explicit, or declarative, memory (e.g., memory for events), which is disrupted by damage to these areas, and implicit, or procedural, memory (e.g., memory for skills and habits), which is not disrupted (Mishkin & Murray, 1994; Squire & Zola-Morgan, 1991; Thompson & Kim, 1996). Thus, an affected individual who learns a motor performance (such as withdrawing quickly from a proffered hand) may display it without remembering the learning experience.

In accord with the traditional information storage concept of memory, explicit and implicit memories are usually assumed to reflect different records of the same experience. Prescientific analyses of memory handled the explicit-implicit distinction differently. In the past, what is now called explicit or declarative memory was not thought to be memory of external events, that is, of experiences or sets of stimuli reaching the nervous system. It was, in Locke’s (1825) terms, memory of the “perception of the operations of our own mind within us” or, in Mill’s (1878) terms, memory of “what we have thought.” It was what Augustine (1961) described when he wrote that, in memory, “I meet my self as well. I remember my self and what I have done, when and where I did it, and the state of my mind at the time.”

This earlier interpretation of the phenomena now referred to as explicit memory seems to mesh more easily with the lesion data and with the concept of memories as connections between remote experience and present behavior than does the currently prevailing assumption that explicit and implicit memories are different records of the same experience. Translated into terms of memory as connections from remote experience to present behavior and recognizing the effect of damage to medial temporal lobe or associated diencephalic structures, “perception of the operations of our own mind within us” or “what we have thought” could be a sequence of events that occurs in these structures during an experience and is the first part of the connection responsible for future explicit recall of the experience. If this is true, absence of explicit memory after medial temporal lobe damage is similar to absence of visual memory after loss of the optic tract or primary visual cortex. In both cases, connections from remote experience to present behavior are absent because their initial stages, the transient events occurring around the time of the experience, simply did not occur. By this analysis, an individual with medial temporal lobe damage does not remember “what [he] thought” during an experience because the nervous system events that constituted the thought never occurred.

In sum, clinical and laboratory lesion data show that the prejunctional parts of certain connections from remote experience to present behavior—those that underlie behaviors usually described as explicit memory—include events in medial temporal lobe and diencephalic structures. To what extent these are transient events that occur during an experience and lead to persistent events and to what extent they are themselves persistent events remain unclear.

In contrast, implicit memory depends on the brain areas most directly involved in the present behavior. Thus, memory of a visuomotor skill requires that primary visual and motor cortices be intact. The regional
lesion method initiated by S. Brown, Schäfer, and Korsakoff cannot detect prejunctional events that occur in such areas because lesions in these areas affect behavior in general; that is, they affect all experience-behavior connections. Newer lesion methodologies can, at least to some degree, avoid this limitation. Pharmacological agents and genetic manipulations can in principle disrupt connections from remote experience while leaving connections from recent experience relatively intact (Abel & Lattal, 2001; Alberini, 1999; DeZazzo & Tully, 1995; Dudai, 1989; Gerlai, 2000; Izquierdo & McGaugh, 2000; Maren & Baudry, 1995; Mayford & Kandel, 1999; Steele, Medina, Nores, & Mauk, 1998). Particularly promising in this regard is the recent advent of techniques that allow specific genetic abnormalities to be targeted to specific brain regions and to specific periods of nervous system development. Lesions of this kind are more likely to impair memory without impairing all experience-behavior connections (e.g., Mayford, Mansuy, Muller, & Kandel, 1997; Wilson & Tonegawa, 1997).

4.4.2 Localization by Imaging. Noninvasive methods for detecting and localizing brain activity give access to the prejunctional portions of the connections that constitute memory, including both the transient events that occur at the time of remote experience and the persistent events that bridge the time until present behavior. These techniques, which include positron emission tomography (PET), functional magnetic resonance imaging (fMRI), electroencephalography (EEG), and magnetoencephalography (MEG), are beginning to reveal the networks of brain regions in which activity is correlated with the transient and persistent events that make up the prejunctional parts of memories (Baddeley, 1998; Buckner & Koutstaal, 1998; Buckner, Kelley, & Peterson, 1999; Cabeza & Nyberg, 2000; Doyon, 1997; Dunn, Dunn, Languis, & Andrews, 1998; Fletcher, Shallice, & Dolan, 1998; Friedman & Johnson, 2000; Gabrieli, 1998; Heiss, Pawlik, Holthoff, Kessler, & Szélies, 1992; Karni et al., 1998; Klimesch, 1996; Knight & Nakada, 1998; Nyberg, 1998; Okada & Salenius, 1998; Paus, Koski, Caramanos, & Westbury, 1998; Schacter & Wagner, 1999; Skrandies & Fahle, 1994; Stern et al., 1996; Tallon-Baudry, Bertrand, Peronnet, & Pernier, 1998; Vogt, Klimesch, & Doppelmayr, 1998; Yoshida, Ueno, Chyne, & Weinberg, 1995).

In their present forms, these methods have certain limitations. They are as yet macroscopic in space and, in the case of PET and fMRI particularly, in time as well. Thus, they are currently restricted to defining activity in relatively large volumes of tissue and over relatively long periods. The interpretation of their results is complicated by uncertainties about the relationships between the measurements made and the neuronal activity they reflect (e.g., Barinaga, 1997; Roskies, 1994). For example, increased cortical blood flow might reflect increased excitation and/or increased inhibition. The techniques have a more general limitation in that they depend on the existence and use of an appropriate control condition (Kim & Baxter, 2001; Roskies, 1994). Subjects are typically studied during or after two experiences that differ as little as possible except that one establishes a specific connection to later behavior that the other (i.e., the control condition) does not. The areas of significant difference between the images are then hypothesized to be the areas responsible for the connection. This hypothesis assumes that correlation indicates causation. It further assumes that the control condition lacks unique effects of its own that contribute to the difference between the images obtained under the two conditions. These limitations should gradually abate as both technology and data analysis continue to develop.

4.4.3 Candidate Persistent Events. Another form of the correlational strategy developed early in this century. Its goal is to detect and define nervous system phenomena that might serve as persistent events. Probably the earliest example of such a phenomenon was “spinal fixation,” first described by Anna DiGiorgio more than 70 years ago (DiGiorgio, 1929; Gerard, 1961; Kandel & Spencer, 1968; Manni, 1950; Patterson, 1976, 2001). In this phenomenon, a relatively brief period of abnormal activity in spinal cord descending pathways, produced by one of a number of different supraspinal lesions, causes a change in spinal cord function that remains indefinitely after the abnormal descending activity ceases and thus indicates the presence of a persistent event in the spinal cord. However, this striking example of nervous system plasticity has received only limited attention, due perhaps to the persistent influence of the ancient Galenic view that the spinal cord is little more than a conduit to and from the brain (Wolpaw & Tennissen, 2001). For the many neuroscientists still affected by this archaic belief, events in the spinal cord are not relevant to the problem of memory.

Present interest is strongest in the phenomena of LTP and LTD, which although originally described in hippocampus and cerebellum, respectively, occur in numerous CNS locations, including the spinal cord (Bliss & Collingridge, 1993; Buonomano & Merzenich, 1998; Calabresi et al., 1996; C. Chen & Tonegawa, 1997; Levenes, Daniel, & Crepel, 1998; Malenka & Nicoll, 1999; Manabe, 1997; Maren & Baudry, 1995; McCrossan, Withington, & Platt, 1997; Oda et al., 1995; Teskey & Valentine, 1998). In these phenomena, appropriate presynaptic stimulation produces an increase or a decrease in the strength of the synaptic connection that can persist for at least several days. Depending on the synapse and the experimental protocol, the change in
synaptic strength is attributable to presynaptic and/or postsynaptic events.

In addition to their rapidity and persistence, LTP and LTD have other properties that seem to qualify them to support memory (C. Chen & Tonegawa, 1997; Izquierdo & Medina, 1995; Laroche, 2000; Malenka & Nicoll, 1999; Maren & Baudry, 1995; Martinez & Derrick, 1996; Miller & Mayford, 1999; Shapiro, 2001). Both LTP and memory are initially vulnerable to disruption by hypoxia, electroconvulsive shock, and other traumatic interventions, and both become resistant thereafter. The patterns of hippocampal neural activity associated with optimum LTP are similar to those found during certain forms of learning, and LTD-like changes in hippocampus have been described in a variety of associative learning paradigms. Pharmacological agents and highly focused genetic abnormalities that prevent LTP also disrupt learning.

Furthermore, the occurrence of LTP and LTD can depend on the relationship between presynaptic excitation and the concurrent activation state of the postsynaptic neuron (Artola & Singer, 1993; Bliss & Collingridge, 1993; Glanzman, 1994; Paulsen & Sejnowski, 2000; Tsien, 2000). Thus, they could conceivably underlie associative memory, such as classical conditioning. The coincidence of presynaptic excitation, caused by a conditioned stimulus (CS), and postsynaptic depolarization, caused by an unconditioned stimulus (US), could produce LTD, a persistent event. Future occurrence of the CS could then elicit a junctional event. If LTD is presynaptic, this junctional event could be increased transmitter release; if LTD is postsynaptic, the junctional event could be increased receptor responsivity. In either case, the combination of the connection from the remote experience (i.e., the CS-US combination) and the connection from the recent experience (i.e., the CS) could activate the postsynaptic neuron and initiate a postjunctional sequence of events leading to behavior (i.e., the conditioned response [CR]). Recent studies linking fear conditioning in rats to LTD in pathways conveying auditory input to the amygdala support a role for LTD in associative conditioning (LeDoux, 2000; Maren, 2001; McKernan & Shinnick-Gallagher, 1997; Rogan, Staubil, & LeDoux, 1997).

At the same time, the correlations of LTD and LTD with memory are not simple or uniform, and it seems clear that other persistent events are also important (Cain, 1997, 1998; Eichenbaum, 1996; Holscher, 1999; Jeffery, 1997; Martinez & Derrick, 1996; Mceachern & Shaw, 1996; Morris & Frey, 1997). LTD and LTD as yet remain distinctive phenomena that may very well contribute to memory. Definition of their roles awaits description of the events linking them to the remote experiences and present behaviors that are the beginnings and endings of the connections that constitute memory. For example, for classical conditioning, adequate description means delineation of the events leading from the remote experience (i.e., the CS-US combination) to the persistent event (i.e., LTP), from the recent experience (i.e., the CS alone) to the junctional event (e.g., increased transmitter release), and from the junctional event to the behavior (i.e., the CR). This task requires a mechanistic strategy that can define the connections from remote experience to LTD and from LTD to behavior.

4.5 The Mechanistic Strategy, or the Search for Junctional Events

The correlational strategy tries to separate, by location or mechanism, the prejunctional parts of connections that begin with remote experiences from connections that begin with recent experiences. In contrast, the newer mechanistic strategy starts from the point of their inevitable overlap: the junctional event in which the sequence of events begun by a recent experience joins with a sequence begun by a remote experience to initiate a single sequence that leads to behavior. The primary goal is to delineate the transient and persistent events that together produce behavior rather than to detect and identify persistent events that are correlated with behavior. The practical logic of the approach is that if the location and nature of the junctional event can be established, the preceding and succeeding parts of the connection (i.e., Figure 1) should become more accessible to study; the preceding part can be followed backward from the junctional event and forward from the remote experience, and the succeeding part, which typically consists of a rapid series of brief events, can be followed out to the behavior.

The same imaging methods used to locate prejunctional events (i.e., see Section 4.4.2) are also being used to locate junctional events (Berthoz, 1997; Buckner et al., 1995; Buckner & Koutstaal, 1998; Fletcher, Shallice, Frith, Frackowiak, & Dolan, 1998; Hasegawa, 2000; McIntosh, 1999; Moscovitch, Kapur, Köhler, & Houle, 1995; Nyberg, 1998; Rösler, Heil, & Hennighausen, 1995; Schacter, Alpert, Savage, Rauch, & Albert, 1996; Schacter, Buckner, & Koutstaal, 1998; Skrandies, 1995; Zhuang et al., 1998). At the same time, the limitations imposed by currently available spatial and temporal resolution and the inherent problem of the control condition apply here also.

At present, the greatest value of the mechanistic or junctional event strategy stems from the fact that, for some behaviors that are connected to both remote and recent experience, the connections from recent experience are defined and accessible anatomically and physiologically. As a result, all the possible locations of the
junctional event(s) are also defined and accessible. In an experimental context, the recent experience is a test stimulus that completes a connection between a remote experience and a behavior. Because a junctional event is part of both connections—the connection from the remote experience and the connection from the test stimulus (i.e., Figure 1) —and because the connection from the test stimulus is defined and accessible—the test stimulus specifies all the possible locations of the junctional event. Test stimuli of this kind have formed the basis for development of simple experimental models for locating and defining junctional events and the associated persistent events.

4.5.1 Simple Models with Accessible Junctional Events. One of the earliest examples of this research strategy was classical conditioning in the isolated spinal cord. In this phenomenon, first described in the 1930s, the recent experience or test stimulus is weak skin or peripheral nerve stimulation (Patterson, 1976; Shurrager & Culler, 1940; Thompson, 2001). This test stimulus, the CS, has a short connection through a chain of several spinal cord neurons to a weak flexion withdrawal response. The remote experience is the repeated presentation of a stimulus pair consisting of the CS followed shortly by the US, strong skin or nerve stimulation that produces by itself a strong flexion withdrawal response. The remote experience establishes a connection that merges with the connection from the test stimulus, the CS alone, to produce a CR, a stronger flexion withdrawal response than the CS alone would have elicited had the remote experience not occurred. The fact that the short connection from the CS to the CR is entirely in the spinal cord and comprises only a few synaptic links offers the possibility of locating the junctional event and eventually the persistent event as well. The essential afferent fibers have been defined and junctional events have been localized to interneurons rather than motoneurons (Durkovic, 1986). Recent evidence suggests that NMDA receptors play a crucial role in the persistent event. This persistent event is established by the CS-US sequence and joins with the connection from the test stimulus (the most recent CS) to produce the junctional event that leads to the CR, a stronger flexion withdrawal response (Durkovic, 2001).

In the 1960s, the search for junctional events began to focus on the simpler and technologically more accessible nervous systems of invertebrates (Kandel & Spencer, 1968). These studies, typified by those of the gill and siphon withdrawal responses to tactile stimulation in *Aplysia californica*, have defined junctional and persistent events that appear to underlie both nonassociative and associative forms of memory (Antonov, Antonova, Kandel, & Hawkins, 2001; Bailey, 1999; Bailey, Alberini, Ghirardi, & Kandel, 1994; Kandel, 1991a; Kandel & Schwartz, 1982). At the same time, it has become clear that the neuronal pathways underlying the connections to recent experience that are used in these studies are not so simple and that particular memories depend on junctional events occurring at multiple locations (T. E. Cohen, Kaplan, Kandel, & Hawkins, 1997; Frost et al., 1997; Frost & Kandel, 1995; Gladman, 1995). For example, the siphon withdrawal response to siphon stimulation in *Aplysia* is the product of at least 16 different neuronal types arranged in at least 11 parallel pathways from sensory input to motor output, and it appears that several different junctional events underlie sensitization of this response. Furthermore, as yet undefined neuronal types and/or additional junctional events are likely to contribute.

Somewhat more recently, the pursuit of junctional events has also concentrated on several promising connections from recent experiences in intact vertebrate nervous systems. These experimental models include the conditioned eye-blink response (Kim & Thompson, 1997; Medina, Nores, Ohyama, & Mauk, 2000; Steinmetz, 2000; Woody & Brit, 1992; Yeo, 1991); heart rate conditioning (D. H. Cohen, 1969; Gherlarducci & Sebastiani, 1997); fear conditioning (LeDoux, 2000; Maren, 2001); the vestibulo-ocular reflex (du Lac, Raymond, Sejnowski, & Lisberger, 1995; Miles & Fuller, 1974; Raymond, 1998); and the spinal stretch reflex (SSR) and its electrical analog, the H-reflex (Segal & Wolf, 1994; Wolpaw, 2001; Wolpaw et al., 1983). In terms of the number and accessibility of the neurons in the connection from the recent experience (i.e., the test stimulus) to the behavior, the simplest of these models are the vestibulo-ocular reflex (VOR), which has three neurons, and the SSR or H-reflex, which has two.

The VOR maintains retinal image stability in the face of sudden head movement (du Lac et al., 1995). Naive VOR gain in the monkey is close to one, so that a head movement is matched by an eye movement of equal and opposite amplitude. In a typical conditioning protocol, exposure to reversing prisms or to magnifying or reducing glasses is used to change VOR gain (e.g., Miles & Fuller, 1974). In this design, the remote experiences are those in which the prisms or glasses cause retinal slip during head movements (i.e., those in which the naive VOR does not properly counteract the effect of head movement); the recent experience, or test stimulus, is an imposed head movement in the dark; and the behavior is an appropriately modified VOR. The recent experience is connected to the behavior by a three-neuron, two-synapse pathway consisting of the vestibular nerve afferent, the vestibular nucleus interneuron, and the extraocular motoneuron. Initial studies sought the junctional event underlying gain change in the elements of this pathway, but it is now clear that the search must be wider (du Lac
et al., 1995; Leigh & Brandt, 1993; Raymond, 1998; Raymond & Lisberger, 2000). A number of more complex pathways, some not yet fully defined, can also contribute to the VOR. Furthermore, several junctional events, occurring in the three-neuron pathway as well as elsewhere, are likely to contribute to VOR gain change. Current efforts are focused on defining the components and operation of the network of brainstem and cerebellar pathways that control VOR gain.

The H-reflex is the electrical analog of the spinal stretch reflex, or tendon jerk, which is the earliest response to sudden muscle stretch (W. F. Brown, 1984; Henneman & Mendell, 1981; Matthews, 1972). In both primates and rats, reward for larger (up-conditioning mode) or smaller (down-conditioning mode) H-reflexes causes mode-appropriate change in H-reflex size (X. Y. Chen & Wolpaw, 1994; Wolpaw, 1987). In this operant conditioning protocol, the remote experiences are those in which the H-reflex is elicited and, depending on its size, is or is not followed by a reward; the recent experience, or test stimulus, is the most recent nerve stimulation; and the behavior is an H-reflex that is larger than normal (when the remote experience was exposure to the up-conditioning mode) or smaller than normal (when the remote experience was exposure to the down-conditioning mode). The recent experience is connected to the behavior mainly by a two-neuron, one-synapse pathway consisting of the Ia primary afferent fiber from the muscle spindle, its synapse on the motoneuron, and the motoneuron itself (W. F. Brown, 1984; Henneman & Mendell, 1981; Matthews, 1972). Physiological and anatomical studies have begun to define the spinal cord plasticity associated with H-reflex conditioning and the nature of the supraspinal influence that creates this plasticity.

These studies of H-reflex conditioning provide a good illustration of the usage and yield of the mechanistic strategy. They indicate that at least two junctional events occur in the motoneuron itself and that several persistent events are located in the spinal cord (Carp & Wolpaw, 1994; Feng-Chen & Wolpaw, 1996; Wolpaw, 1997, 2001; Wolpaw & Lee, 1989). These two junctional events and the experiments encouraged and guided by them are briefly reviewed in the next section to show how the mechanistic strategy can elucidate the connections between remote experience and behavior that constitute memory. At the same time, they also illustrate the complexity of even the simplest memories and thereby emphasize the importance of adopting the simplest possible experimental models.

### 4.5.2 Identification and Exploration of Two Junctional Events

The monosynaptic pathway underlying the H-reflex has two likely locations for a junctional event that changes H-reflex size. One location is the Ia afferent-motoneuron synaptic connection, and the other is the motoneuron itself. A junctional event at the synapse could change the excitatory postsynaptic potential (EPSP) that depolarizes the motoneuron, whereas a junctional event in the motoneuron could change the motoneuron’s response to the EPSP. Either event could change the number of motoneurons that are induced to fire and thereby change the behavior, that is, the size of the H-reflex.

Intracellular studies have provided evidence that a junctional event in the motoneuron largely accounts for operantly conditioned decrease in the H-reflex. Successful down conditioning is associated with a positive shift in motoneuron firing threshold and a drop in motoneuron axonal conduction velocity (Carp, Chen, Sheikh, & Wolpaw, 2001; Carp & Wolpaw, 1994). Both these findings support (qualitatively and quantitatively) the hypothesis that the remote experience of being rewarded for a smaller H-reflex produces a persistent event consisting of a positive shift in sodium channel activation voltage throughout the motoneuron (Halter, Carp, & Wolpaw, 1995). According to this hypothesis, the recent experience (i.e., the most recent nerve stimulation) initiates a sequence of events that combines with this persistent event to produce two junctional events. The recent experience elicits in the motoneuron an EPSP that combines with the persistent event to produce the first junctional event: a decreased probability that the motoneuron will fire in response to the EPSP. Because the motoneuron is farther from threshold, an EPSP that would have brought it to threshold prior to downconditioning may no longer do so. In those motoneurons that do fire in response to the EPSP, the resulting action potential combines with the persistent event to produce the second junctional event: slower conduction of the action potential down the motoneuron axon. Because the axonal membrane is farther from threshold, action potential generation at each node of Ranvier takes slightly longer. The connections from these junctional events to altered behavior are straightforward: because a smaller number of motoneurons are excited by the nerve stimulation the H-reflex is markedly smaller and because motoneuron axonal conduction velocity is lower the H-reflex is slightly later. Calculations based on the magnitude of the threshold shift and the distribution of Ia EPSP amplitudes imply that the shift accounts for most of the decrease in H-reflex size and for the decrease in axonal conduction velocity (Carp & Wolpaw, 1994; Halter et al., 1995).

Recent studies have begun to trace the connection backward from these junctional events in the spinal cord to the remote experience, which must involve the supraspinal areas that receive and process the input pro-
duced by reward delivery. As noted, the best current hypothesis is that the persistent event responsible for the changes in motoneuron firing threshold and axonal conduction velocity is a positive shift in sodium channel activation voltage. This could in turn result from a change in protein kinase C activation (Cantrell & Catterall, 2001; Franceschetti et al., 2000; Halter et al., 1995). The issue should be amenable to in vitro studies of tissue from conditioned animals (e.g., Carp, Smith, & Wolpaw, 2001; Hori, Tan, Strominger, & Carpenter, 2001). One of the several structural changes described in particular synaptic populations on the motoneurons of conditioned animals (Feng-Chen & Wolpaw, 1996) may reflect the mechanism that induces the persistent event.

The most important achievement to date has been the identification of the descending spinal cord tract that connects the supraspinal regions involved in the remote experience to the spinal cord motoneuron where the persistent event (threshold change) and its associated junctional events (decreased motoneuron excitation and conduction velocity) occur. The corticospinal tract (CST) is essential for down conditioning: CST transection prevents down conditioning, whereas transection of other major descending tracts does not do so (X. Y. Chen & Wolpaw, 1997, 2002). The nature of the CST activity responsible for down conditioning and the supraspinal origins of the input to sensorimotor cortex that elicits this CST activity remain to be defined. Current studies suggest that cerebellar-cortical connections have an important role (X. Y. Chen, Chen, & Wolpaw, 2001). Future studies will explore the roles of other supraspinal structures in connecting the remote experience to the junctional events and thus to the behavior—a smaller and later H-reflex. The ultimate goal is to define the entire sequence of events that leads from the operant conditioning experience to the altered H-reflex.

The persistent event responsible for the change in motoneuron threshold appears to be primary plasticity as defined in Section 3.5.2: It is probably responsible for the intended change in behavior, a smaller H-reflex. At the same time, this persistent event is likely to affect numerous other behaviors that involve the same motoneuron population and thus to induce compensatory plasticity (Section 3.5.2). For example, after down conditioning, achievement of appropriate motoneuron activation during locomotion probably requires greater excitatory input than it did before in order to overcome the positive shift in firing threshold. This change in input might be accompanied by activity-driven plasticity in the synapses that convey the input and/or in the pathways leading to them. It is also possible that the slight delay in muscle excitation produced by the drop in motoneuron axonal conduction velocity necessitates changes in the timing of activation of other muscles to preserve complex coordinated movements and that these changes too are associated with activity-driven plasticity. Furthermore, as noted in Section 3.5.2, both primary and compensatory plasticity are likely to induce reactive plasticity.

Physiological and anatomical data do indicate that H-reflex conditioning is associated with other persistent events and junctional events, some in the spinal cord and some in supraspinal regions. These appear to include changes in Ia afferent-motoneuron synaptic function, in other synaptic terminals on the motoneuron, and in interneurons that convey oligosynaptic Group 1 input to the motoneuron (Carp & Wolpaw, 1994, 1995; Feng-Chen & Wolpaw, 1996; Wolpaw, 1997, 2001). Whereas some of these events may contribute directly to H-reflex change, others are likely to represent compensatory or reactive plasticity.

Most surprisingly, one of these additional persistent events is located in the contralateral spinal cord (Wolpaw & Lee, 1989). H-reflex conditioning has little or no effect on the contralateral H-reflex of the awake behaving animal. The effect of conditioning is focused on the ipsilateral H-reflex, the behavior that determines whether reward occurs (Wolpaw, Herchenroder, & Carp, 1993). However, when the conditioned animal is anesthetized and the spinal cord is separated from supraspinal influence by transection, the contralateral reflex of a down-conditioned animal is at least twice as large as that of an up-conditioned animal or a naive animal (Wolpaw & Lee, 1989). Anesthesia and cord transection uncover a hidden effect of conditioning: a change on the contralateral side of the spinal cord. In the awake behaving animal, this plasticity is not apparent presumably because its effect on the contralateral H-reflex is canceled out by other spinal or supraspinal plasticity. Anesthesia and transection, which remove supraspinal influence and quiet tonic spinal cord activity, eliminate the canceling effect and reveal the presence of plasticity that changes the contralateral reflex.

What this contralateral plasticity might be is as yet unknown. It may be related to the hyporeflexia that has been noted ipsilateral to sensorimotor cortex strokes in humans (Thilmann, Fellows, & Garms, 1990). In the present context, it is important for two reasons. First, it illustrates the complexity of the nervous system changes associated with even the simplest learning. Second, it illustrates the inadequacy of approaching memory in terms of the encoding, maintenance, and retrieval of information about experience. Not only does this contralateral plasticity lack an apparent relationship to the ipsilateral H-reflex change that is the goal of conditioning, it lacks an apparent relationship to any behav-
ioral change: The contralateral H-reflex of the awake behaving animal is not changed. However, the existence of this contralateral plasticity indicates that conditioning affects production of the contralateral H-reflex—the processes that produce the unchanged reflex are different after conditioning. The contralateral plasticity cannot be easily construed as information about the conditioning experience, nor does it appear to be retrieved as behavior. Thus, it does not fit into the framework of encoding, maintenance, and retrieval. Nevertheless, it is an effect of the experience, and it modifies the nervous system processes that produce behavior.

Although the exploration of the H-reflex model is just beginning, the results to date are both encouraging and daunting. On one hand, application of the mechanistic strategy in this simple model is beginning to trace the connection of remote experience to behavior. On the other hand, even this extremely simple model produces multiple persistent events that affect behavior in complicated ways.

4.6 The Importance of Simplicity

It is clear that the memories established in the simplest vertebrate and invertebrate models are far more complex than was expected, or at least hoped, when the models were first developed. As discussed in Section 3.5.2, because the nervous system is a multitask system and has a ubiquitous capacity for activity-driven plasticity, even the simplest memory involves multiple persistent and junctional events at multiple locations. Thus, realization of the theoretical promise of these models for defining entire connections from remote experience to present behavior is certain to be a long and difficult process.

At the same time, the emerging intricacy of the processes operating in these simple models is further evidence of their importance in studying memory because it implies that the processes underlying more complicated memories are so complex as to defy productive exploration at present. The plasticity underlying a new motor skill, for example, necessarily includes modifications in the plasticity underlying other skills and changes as further experience occurs, and the plasticity responsible for the memory of a picture is probably inseparable from that responsible for memory of previous and subsequent visual experience. Although studies of such complicated memory phenomena may have practical clinical implications, they are likely to be less useful in defining the connections between remote experience and present behavior. Because understanding of brain function is still limited and fragmentary, attempts to study complex memories inevitably fall back on the concepts and the language derived from introspection, and they risk producing only contemporary equivalents of the hydraulic and vibratory memory theories of the 17th and 18th centuries—ultimately sterile combinations of scientific understanding and prescientific ideas. At present, continued pursuit of the simplest memories in the simplest models appears to offer the best hope for the substantive new insights that will eventually support productive study of more complicated memory phenomena. Thus, models that offer access to junctional events and persistent events located in the experience-behavior connections responsible for very simple memories are likely to play an increasingly central role in the study of memory. The access they provide will be needed to define the roles of LTP and LTD in memory and to clarify the contributions of different regions, specific neuronal populations, and particular genes.

4.7 Persistent and Junctional Events Versus the Storehouse

The sensorimotor hypothesis that defines modern neuroscience underlies the concept of memories as sequences of events connecting remote experience to present behavior and implies that the distinguishing features of these sequences are persistent and junctional events. As the preceding sections indicate, this conception of memory in terms of persistent and junctional events provides a logical framework for the wide variety of research approaches that have developed over the past century. Furthermore, the concept is free of the four problems (Section 3.5) that beset the traditional concept of memory as a storehouse of information about experience and prevent it from serving as an adequate basis for research.

First, the persistent and junctional event formulation adheres to the agenda that follows from the sensorimotor hypothesis—to explain the connections from experience to behavior: It focuses exclusively on explaining the effects of remote experience on present behavior. In contrast, the storehouse concept of memory, motivated by introspection, digresses from this agenda to the superfluous and difficult question of what is recorded from experience—of what the representations in the storehouse represent.

Second, the concept of persistent and junctional events has no difficulty accommodating the complex plasticity produced by experience or the fact that this plasticity and its effects on behavior continually change with new experience and with development, aging, and trauma. Thus, for example, it can readily encompass the as-yet mysterious contralateral plasticity associated with H-reflex conditioning (Section 4.5.2), despite the fact that this plasticity is hard to understand as information about the conditioning experience. In contrast, the storehouse concept, because it is tied to the notion of memories as representations of experience, cannot
readily accommodate this complex and changing plasticity.

Third, the persistent and junctional events concept makes no distinction between the plasticity that constitutes memory and the plasticity associated with development, aging, or trauma. As a result, it has, for example, no difficulty in approaching the modifications in the cortical homunculi associated with experience and trauma as closely related phenomena with similar or identical mechanisms (Section 3.5.3). In contrast, the storehouse concept assumes a useless and indefensible distinction between the plasticity produced by experience and other activity-driven plasticity.

Fourth, the concept of persistent and junctional events, because it is not constrained by the notion of memories as representations of experience, can readily address the various processes that are triggered by experience and the different ways in which each may contribute to different behaviors at different times. In contrast, the storehouse concept implies that the multiple behavioral effects of experience reflect multiple representations of experience and thus divides memory research into multiple separate endeavors, each exploring the encoding, maintenance, and retrieval of a particular representation. This compartmentalization is not consistent with the ubiquity, complexity, and continual modification of activity-driven plasticity in the nervous system.

In sum, the concept of memory as sequences of events connecting remote experience to present behavior avoids the superfluities, inconsistencies, and needless constraints of the storehouse concept and provides a logical and useful framework for incorporating memory into the research agenda that follows from the sensorimotor hypothesis that underlies modern neuroscience.

4.8 The Shift From Experience to Behavior

Introspection, which led to the ancient concept of memory as a storehouse of experience, fostered the assumption that the memory of a new word, a picture, a simple motor skill, or the route home is a discrete phenomenon that is caused by a discrete experience and depends on a discrete sequence of storage processes occurring in a few specific brain regions. This assumption focuses attention on the experience (e.g., exposure to the new word) and limited aspects of the nervous system activity occurring during and immediately after it. The research goal becomes localization and description of the supposed storage processes and their product. The subsequent effects of this product on behavior (e.g., later use of the word) become a separate problem. This formulation, however well founded in introspection, is not drawn from the sensorimotor hypothesis of neuroscience, nor is it compatible with the continual occurrence of activity-driven plasticity throughout the nervous system.

The sensorimotor hypothesis of neuroscience—that the whole function of the nervous system is simply to connect experience to behavior—implies that the memory of a remote experience is no more nor less than a set of connections to a set of present behaviors. Furthermore, the fact that activity-driven plasticity occurs continually throughout the nervous system ensures that the memory of an experience (i.e., the set of connections from the experience to present behaviors) is continually changing and that almost every behavior reflects numerous connections from numerous remote experiences. The use of a new word, for example, depends not only on initial exposure to the word but also and probably to a greater extent on the prolonged early experiences responsible for language mastery. Without the persistent events established by such early experience, use of the new word would not be possible. And in comparison to those events, the additional plasticity caused by initial exposure to the new word is likely to be quite trivial, comprising slight adjustments in or minimal additions to the previously established persistent events. In this light, an effort to understand the mechanisms responsible for use of the new word logically begins from the behavior itself rather than from the experience of first exposure to the word.

Together, the sensorimotor hypothesis and the ubiquity of nervous system plasticity shift the focus of memory research away from seeking stable records of specific experiences, which do not appear to exist, and toward defining the processes underlying behavior (Vanderwolf & Cain, 1994). The shift is reflected in the recent prominence of studies based on the mechanistic strategy and the recognition that the correlational strategy alone cannot explain the phenomena of memory. Although it can elucidate events critical to memory, it must be combined with the mechanistic strategy to explain behavior. In accord with other research based on the sensorimotor hypothesis, the goal of memory research is not to describe records of experience but rather to understand the production of behavior.

5. CONCLUSIONS

The impetus for this review was discomfort with the traditional idea that the brain both responds to experience and stores representations of experience for use in formulating future responses. It set out to learn how that idea came about and whether it actually fits into the theoretical structure and practical agenda of modern neuroscience. The goal was to understand the concept of memory that operates in neuroscience today and thus to understand what we are studying when we study memory.
and the strategies that we use to study it. The principal points are as follows.

Long before the beginning of scientific inquiry, memory was conceived, largely on the basis of introspection, to be a personal storehouse of experience accessible only to one individual. In the early 19th century, with the acceptance of the sensorimotor hypothesis that the entire function of the nervous system is to connect experience to behavior, the storehouse concept of memory became a problem. It became a problem scientifically because a personal storehouse was not accessible to observation, and it became a problem neuroscientifically because it implied that the nervous system did more than simply connect experience to behavior.

The standard solution has been to redefine memory as a storehouse of information that is derived from experience and that affects behavior. This solution makes memory accessible to observation. At the same time, however, it retains the ancient assumption that the nervous system stores records or representations of remote experience and uses them to formulate future behavior.

This assumption is based only on introspection. Furthermore, it introduces issues that are, in the context of neuroscience, unnecessary and irrelevant. It introduces the question of what the information is about and implies the existence or the possibility of an observer who can decipher the information. It also implies that the objective of memory research is to define the processes that encode, maintain, and retrieve this information. These implications are irrelevant to the sensorimotor hypothesis—that the brain simply connects experience to behavior—and do not adequately serve the research objective derived from this hypothesis—to understand the connections between experience and behavior.

Most important, the assumption of information storage ignores the necessary and inevitable complexity of the plasticity associated with even the simplest memories and ignores the ubiquity of activity-driven nervous system plasticity in the nervous system. An experience is likely to generate three categories of plasticity: primary plasticity responsible for the obvious or intended behavioral effect, compensatory plasticity necessary to maintain other behaviors that depend on neural elements affected by primary plasticity, and reactive plasticity secondary to the first two categories. These categories of plasticity and their effects on behavior will depend on earlier experience and will be continually modified by later experience as well as by development, aging, and trauma. Thus, the assumption that the nervous system stores information derived from experience is not an adequate basis for addressing the complex processes that determine the effects a remote experience exerts on present behavior. In addition, the assumption creates a meaningless distinction between the plasticity produced by experience and that occurring with development, aging, and trauma.

Defined in terms of the hypothesis that the whole function of the nervous system is to connect experience to behavior and defined as it actually functions in neuroscience research, memory encompasses the sequences of events that comprise connections from remote experience to present behavior. These connections have two distinctive features. The first is the delay between experience and behavior, and the second is the fact that the occurrence of the behavior depends on, or is enabled by, recent experience. Thus, memories have two essential components: persistent events that bridge the lengthy time between experience and behavior and junctional events in which the connections from remote experience and the connections from recent experience merge to initiate single sequences of events leading to behavior.

This definition of memory—as the sequences of events comprising connections between past experience and present behavior—is consistent with the central goal of neuroscience—to understand the connection of experience to behavior—and underlies the entire spectrum of memory research methods. Analyzed in terms of this definition, memory research has two basic strategies: a traditional correlational strategy that focuses on persistent events and a newer mechanistic strategy that focuses on junctional events.

The older correlational strategy, originally motivated by the concept of memory as information storage, developed in the late 19th century. It began with studies that indicated the importance of medial temporal lobe and diencephalic structures for certain memories and thereby suggested that persistent events were located in these regions. It continues at present in studies of phenomena, such as LTP and LTD, that might serve as the persistent events in the connections from remote experience to present behavior. However, persistent events do not by themselves explain memory. It is also necessary to understand their links to behavior. The newer mechanistic strategy, derived from the sensorimotor hypothesis, seeks to locate and describe junctional events and use them as starting points for exploring the sequences of events comprising connections between remote experience and present behavior, including the essential persistent events. It relies principally on very simple connections from recent experience to behavior. In these connections, the possible sites of junctional events are defined and accessible. This strategy is typified by studies of gill and siphon withdrawal in Aplysia and of the VOR and the H-reflex in vertebrates. The complexity of the plasticity found even in these ostensibly simple models
emphasizes the importance of extreme simplicity in experimental models used for studying memory.

The growing prominence of the mechanistic strategy as an essential complement to the correlational strategy is in accord with the sensorimotor hypothesis and with the ubiquity of activity-driven plasticity in the nervous system. It signals a switch in memory research from the traditional focus on the storage of experience to a new focus on understanding the genesis of behavior. It recognizes that behavior, which has traditionally been seen as merely reflecting the storehouse of information, is in fact the only product of the nervous system and thus that the explanation of behavior is the entire goal of research.

NOTES

1. It was discovered only recently that the pulmonary circulation had been hypothesized in the 13th century by the Arab physician Ibn al-Nafis al-Qurashi, but this was almost certainly not known to Servetus (Boas, 1962; Sarton, 1931).

2. The information stored by the nervous system is clearly more than the "information" of information theory, which is simply the reduction of uncertainty (Pierce, 1980; Shannon & Weaver, 1964). (For example, when a synaptic bouton from one neuron produces an excitatory postsynaptic potential in a second neuron, it conveys information because it reduces the uncertainty of the state of the second neuron; e.g., the second neuron becomes likely to fire.)

3. Burnham (1889) and Gomulicki (1953) seem to have translated Huber's (1878) "gran" (i.e., "grain") as "gram," and thus they concluded that each gram should contain about 205,542 traces. This apparent error corrected, calculation from the numbers given by Huber indicates that each grain should contain about 102,771 traces, not 205,542. The origin of this discrepancy is not clear.

4. Nevertheless, the cardiovascular school had disciples into the 18th century, and its echoes persist in everyday expressions such as "to learn by heart" (Clarke & O'Malley, 1996).

5. In a less usual form, the hypothesis is that the function of the brain is to connect activity in those neurons that receive nonneuronal input (i.e., principally neurons that are excited by sensory receptors) to activity in those neurons that excite nonneuronal structures (i.e., proprioceptive neurons that excite muscles or other effector organs). These two classes of neurons are few in number, comprising in humans less than 0.1% of the total neuronal population (Blinkov & Glezer, 1968; Brodal, 1981; Forger & Breedlove, 1987; Henry & Calaresu, 1972; Kaalen, Jacobsen, & Jakubow, 1988; Kandel, 1991b; Kupfer, Chumbley, & Downer, 1967; McCleod & Wray, 1967; Petras & Cummings, 1972). Nevertheless, they are believed to be, respectively, the origin and focus of the activity of the remaining 99.9+%.

The sensorimotor hypothesis as described here does not ignore or discount intrinsic nervous system activity, such as that produced by continuously active cells or internal circuits of cells, in determining behavior. In terms of the hypothesis, this activity is comparable to the neuronal structure that comprises the nervous system, in that it is among the factors that determine the relationship between input and output. Stated in another way: To the extent that intrinsic activity does not change in response to sensory input, it is part of the structure of the nervous system; and to the extent that it does change in response to sensory input, it is part of the sequence of events leading from input to output.

It should be noted that this new definition of memory does not encompass one of the traditional categories of memory (Section 2.1)—memories that are not records of experience but rather records of thoughts, records "of the operations of our own mind within us" (Locke, 1825). This issue is addressed in Section 4.4.1.

7. The use of the term compensatory plasticity to describe a portion of the effects of normal learning might seem to risk confusion with compensatory plasticity occurring in response to nervous system trauma. But this is in fact the central point: There really are no grounds for distinguishing between the two forms of compensation. They have the same objective—the preservation or restoration of behavior—and they appear to involve similar mechanisms (see next section). The lack of any realistic distinction between them illustrates further the inadequacy of the concept of memory as representation of experience.

8. From an experimental point of view, the distinction between explicit and implicit memory is essentially an operational one. They are distinguished by the information that an investigator must have to recognize them. To recognize implicit memory, the behavior alone is required. For example, the behavior itself reveals whether a subject has learned a skill or has been classically conditioned to a particular conditioned stimulus: If the skill or the conditioned response is displayed, implicit memory is present. In contrast, to recognize explicit memory, both the behavior and the subject's remote experience are required. For example, knowledge of a subject's training history is needed to decide whether his account of it is correct, and knowledge of the symbol an animal was last exposed to is needed to decide whether its delayed nonmatching to sample is correct (e.g., Mishkin, 1978). Just as the subject without knowledge of remote experience lacks explicit memory, the investigator without knowledge of that remote experience cannot detect the presence or absence of explicit memory.

While synaptic plasticity has traditionally received the most attention as the probable basis of learning, the possibility that learning can also involve changes in neuronal voltage-gated ion channels has recently drawn interest (Carr & Catterall, 2001; Spitzer, 1999). The shift in mean neuron firing threshold produced by down conditioning of the H-reflex appears to be an example of such neuronally based learning.

REFERENCES


