GOVERNMENT ARTS COLLEGE (AUTONOMOUS)
COIMBATORE-18

DEPARTMENT OF PSYCHOLOGY

STUDY MATERIALS

18MPS11C-HISTORY OF PSYCHOLOGY

UNIT – III : COGNITION AND LEARNING

The philosophical period – The early scientific period – The modern scientific period.

UNIT -3

COGNITION AND LEARNING

Trying to understand the nature of cognition is the oldest psychological enterprise, having its beginnings in ancient Greek philosophy. Because the study of cognition began in philosophy, it has a somewhat different character than other topics in the history of psychology. Cognition is traditionally (I deliberately chose an old dictionary) defined as follows: "Action or faculty of knowing, perceiving, conceiving, as opposed to emotion and volition" (*Concise Oxford Dictionary,* 1911/1964, p. 233). This definition has two noteworthy features. First, it reflflects the traditional philosophical division of psychology into three fields: cognition (thinking), emotion (feelings), and conation, or will (leading to actions). Second, and more important in the present context, is the definition of cognition as *knowing*. Knowing, at least to a philosopher, is a success word, indicating possession of a justifiably true belief, as opposed to mere opinion, a belief that may or may not be correct or that is a matter of taste. From a philosophical perspective, the study of cognition has a normative aspect, because its aim is to determine what we *ought* to believe, namely, that which is true.

The study of cognition therefore has two facets. The first is philosophical, lying in the field of epistemology, which inquires into the nature of truth. The second is psychological, lying in the field of cognitive psychology or cognitive science, which inquires into the psychological mechanisms by which people acquire, store, and evaluate beliefs about the world. These two facets are almost literally two sides of acoin that cannot be pried apart. Once philosophers distinguished truth from opinion (epistemology), the question immediately arose as to how (psychology) one is to acquire the former and avoid the latter. At the same time, any inquiry into how the mind works (psychology) necessarily shapes investigations into the nature of truth (philosophy). The philosophers whose work is summarized below shuttled back and forth between inquiries into the nature of truth— epistemology—and inquiries into how humans come to possess knowledge.

This joint philosophical-psychological enterprise was profoundly and permanently altered by evolution. Prior to Darwin, philosophers dwelt on the human capacity for knowledge. Their standard for belief was Truth: People ought to believe what is true. Evolution, however, suggested a different standard, workability or adaptive value: People ought to believe what works in conducting their lives, what it is *adaptive* to believe. From the evolutionary perspective, there is little difference between the adaptive nature of physical traits and the adaptive nature of belief formation. It makes no sense to ask if the human opposable thumb is "true": It works for us humans, though lions get along quite well without them. Similarly, it may make no sense to ask if the belief "Lions are dangerous" is metaphysically true; what counts is whether it's more adaptive than the belief "Lions are friendly." After Darwin, the study of cognition drifted away from philosophy (though it never completely lost its connection) and Trying to understand the nature of cognition is the oldest psychological enterprise, having its beginnings in ancient Greek philosophy. Because the study of cognition began in philosophy, it has a somewhat different character than other topics in the history of psychology. Cognition is traditionally (I deliberately chose an old dictionary) defined as follows: "Action or faculty of knowing, perceiving, conceiving, as opposed to emotion and volition" (*Concise Oxford Dictionary*, 1911/1964, p. 233). This definition has two noteworthy features. First, it reflects the traditional philosophical division of psychology into three fields: cognition (thinking), emotion (feelings), and conation, or will (leading to actions). Second, and more important in the present context, is the definition of cognition as *knowing*. Knowing, at least to a philosopher, is a success word, indicating possession of a justifiably true belief, as opposed to mere opinion, a belief that may or may not be correct or that is a matter of taste. From a philosophical perspective, the study of cognition has a normative aspect, because its aim is to determine what we *ought* to believe,

namely, that which is true.

The study of cognition therefore has two facets. The first is philosophical, lying in the field of epistemology, which inquires into the nature of truth. The second is psychological, lying in the field of cognitive psychology or cognitive science, which inquires into the psychological mechanisms by which people acquire, store, and evaluate beliefs about the world. These two facets are almost literally two sides of a became the study of learning, inquiring into how people and animals—another effect of evolution—acquire adaptive beliefs and behaviors. I divide my history of cognition and learning into three eras. The fifirst is the Philosophical Era, from Classical Greece up to the impact of evolution. The second is the Early Scientific Era, from the impact of evolution through behaviorism.

The third is the Modern Scientific Era, when the psychological study of learning and cognition resumed its alliance with philosophy in the new interdisciplinary endeavor of cognitive science.

THE PHILOSOPHICAL PERIOD

During the Premodern period, inquiries into cognition focused on philosophical rather than psychological issues. The chief concerns of those who studied cognition were determining how to separate truth from falsity and building systems of epistemology that would provide sure and solid foundations for other human activities from science to politics.

THE PREMODERN PERIOD:COGNITION BEFORE THE SCIENTIFIC REVOLUTION

Thinking about cognition began with the ancient Greeks. As Greek thought took flight beyond the bounds of religion, philosophers began to speculate about the nature of the physical world. Political disputes within the *poleis* and encounters with non-western societies provoked debates about the best human way of life. These social, ethical, and protoscientific inquiries in turn raised questions about the scope and limits of human knowledge, and how one could decide between rival theories of the world, morality, and the best social order. The epistemological questions the ancient philosophers posed are perennial, and they proposed the first—though highly speculative—accounts of how cognition works psychologically.

The Classical World before Plato

By distinguishing between Appearance and Reality, the Greeks of the fifth century B.C.E. inaugurated philosophical and psychological inquiries into cognition. Various pre Socratic philosophers argued that the way the world seems to us—Appearance—is, or may be, different from the way the world is in Reality. Parmenides argued that there is a fixed reality (Being) enduring behind the changing appearances of the world of experience. Against Parmenides, Heraclitus argued that Reality is even more fluid than our experience suggests. This pre-Socratic distinction between Appearance and Reality was metaphysical and ontological, not psychological. Parmenides and Heraclitus argued about the nature of a "realer," "truer" world existing in some sense apart from the one we live in. However, drawing the distinction shocked Greeks into the realization that our knowledge of the world— whether of the world we live in or of the transcendental one beyond it—might be flawed, and Greek thinkers added epistemology to their work, beginning to examine the processes of cognition (Irwin, 1989).

One of the most durable philosophical and psychological theories of cognition, the *representational theory,* was first advanced by the Greek philosopher-psychologists Alcmaeon and Empedocles. They said that objects emit little copies of themselves that get into our bloodstreams and travel to our hearts, where they result in perception of the object. The famous atomist Democritus picked up this theory, saying that the little copies were special sorts of atoms called *eidola*.

Philosophically, the key feature of representational theories of cognition is the claim that we do not know the external world directly, but only indirectly, via the copies of the object that we internalize. Representational theories of cognition invite investigation of the psychological mechanisms by which representations are created, processed, and stored. The representational theory of cognition is the foundation stone of Simon and Newell's symbol-system architecture of cognition.

Once one admits the distinction between Appearance and Reality, the question of whether humans can know Reality— Truth—arises. Epistemologies can be then divided into two camps: those who hold that we are confined to dealing with shifting appearances, and those who hold that we can achieve genuine knowledge. group the Relativists: For them, truth is ever changing because appearances are ever changing. I will call the second group the Party of Truth: They propose that humans can in

Path Metaphysics	RATIONALISM (typically linked to IDEALISM)	EMPIRICISM
Party of TRUTH	Socrates Plato Stoics Descartes Kant	Alcmaeon Empedocles Locke Positivism
Party of RELATIVISM	Hegel Nietzsche	Sophists Hume Pragmatism

some way get beyond appearances to an enduring realm of Truth. The first relativists were the Greek Sophists. They treated the distinction between Appearance and Reality as insurmountable, concluding that what people call truth necessarily depends on their own personal and social circumstances.

Thus, the Greek way of life seems best to Greeks, while the Egyptian way of life seems best to Egyptians. Because there is no fixed, transcendental Reality, or, more modestly, no transcendental Reality accessible to us, we must learn to live with Appearances, taking things as they seem to be, abandoning the goal of perfect Knowledge. The Sophists' relaxed relativism has the virtue of encouraging toleration: Other people are not wicked or deluded because they adhere to different gods than we do, they simply have different opinions than we do. On the other hand, such relativism can lead to anarchy or tyranny by suggesting that because no belief is better than any other, disputes can be settled only by the exercise of power.

Socrates, who refused to abandon truth as his and humanity's proper goal, roundly attacked the Sophists. Socrates believed the Sophists were morally dangerous. According to their relativism, Truth could not speak to power because there are no Truths except what people think is true, and human thought is ordinarily biased by unexamined presuppositions that he aimed to reveal. Socrates spent his life searching for compelling and universal moral truths. His method was to searchingly examine the prevailing moral beliefs of young Athenians, especially beliefs held by Sophists and their aristocratic students. He was easily able to show that conventional moral beliefs were wanting, but he did not offer any replacements, leaving his students in his own mental state of *aporia*, or enlightened ignorance. Socrates taught that there are moral truths transcending personal opinion and social convention and that it is possible for us to know them because they were innate in every human being and could be made conscious by his innovative philosophical dialogue, the *elenchus*. He rightly called himself truth's midwife, not its expositor. Ironically, in the end Socrates' social impact was the same as the Sophists'. Because he taught no explicit moral code, many Athenians thought Socrates was a Sophist, and they convicted him for corrupting the youth of Athens, prompting his suicide.

For us, two features of Socrates' quest are important. PreSocratic inquiry into cognition had centered on how we perceive and know particular objects, such as cats and dogs or trees and rocks. Socrates shifted the inquiry to a higher plane, onto the search for general, universal truths that collect many individual things under one concept. Thus, while we readily see that returning a borrowed pencil and founding a democracy are just acts, Socrates wanted to know what Justice *itself* is. Plato extended Socrates' quest for universal moral truths to encompass all universal concepts. Thus, we apply the term "cat" to all cats, no two of which are identical; how and why do we do this? Answering this question became a central preoccupation of the philosophy and psychology of cognition.

The second important feature of Socrates' philosophy was the demand that for a belief to count as real knowledge, it had to be justifiable. A soldier might do many acts of heroic bravery but be unable to explain what bravery is; a judge might be esteemed wise and fair but be unable to explain what justice is; an art collector might have impeccable taste but be unable to say what beauty is. Socrates regarded such cases as lying awkwardly between opinion and Truth. The soldier, judge, and connoisseur intuitively embrace bravery, justice, and beauty, but they do not possess knowledge of bravery, justice, and beauty unless and until they can articulate and defend it.

For Socrates, unconscious intuition, even if faultless in application, was not real knowledge.

Plato and Aristotle

Of all Socrates' many students, the most important was Plato. Before him, philosophy—at least as far as the historical record goes—was a hit or miss affair of thinkers offering occasional insights and ideas. With Plato, philosophy became more self-conscious and systematic, developing theories about its varied topics. For present purposes, Plato's importance lies in the influential framework he created for thinking about cognition and in creating one of the two basic philosophical approaches to understanding cognition.

Plato formally drew the hard and bright line between opinions—beliefs that might or might not be true—and knowledge, beliefs that were demonstrably true. With regard to perception, Plato followed the Sophists, arguing that perceptions were relative to the perceiver. What seemed true to one person might seem false to another, but because each sees the world differently, there is no way to resolve the difference between them. For Plato, then, experience of the physical world was no path to truth, because it yielded only opinions. He found his path to truth in logic as embodied in Pythagorean geometry. A proposition such as the Pythagorean theorem could be *proved,* compelling assent from anyone capable of following the argument. Plato was thus the fifirst philosophical *rationalist,* rooting knowledge in reason rather than in perception. Moreover, Plato said, provable truths such as the Pythagorean theorem do not apply to the physical world of the senses and opinion but to a transcendental realm of pure Forms (in Greek) of which worldly objects are imperfect copies. In summary, Platotaught that there is a transcendental and unchanging realm of Truth and that we can know it by the right use of reason.

Plato also taught that some truths are innate. Affected by Eastern religions, Plato believed in reincarnation and proposed that between incarnations our soul dwells in the region of the Forms, carrying this knowledge with them into their next rebirth. Overcome by bodily senses and desires, the soul loses its knowledge of the Forms. However, because worldly objects resemble the Forms of which they are copies, experiencing them reactivates the innate knowledge the soul acquired in heaven. In this way, universal concepts such as *cat* or *tree* are formed out of perceptions of individual cats or trees. Thus, logic, experience, and most importantly Socrates' *elenchus* draw out Truths potentially present from birth. Between them, Socrates and Plato began to investigate a problem in the study of cognition that would vex later philosophers and that is now of great importance in the study of cognitive development. Some beliefs are clearly matters of local, personal experience, capturing facts that are not universal. An American child learns the list of Presidents, while a Japanese child learns the list of Emperors.

Another set of beliefs is held pretty universally but seems to be rooted in experience. American and Japanese children both know that fifire is hot. There are other universal beliefs, however, whose source is harder to pin down. Socrates observed that people tended to share intuitions about what actions are just and which are unjust. Everyone agrees that theft and murder are wrong; disagreement tends to begin when we try to say why. Plato argued that the truth of the Pythagorean theorem is universal, but belief in it derives not from experience—we don't measure the squares on 100 right-angled triangles and conclude that *a*2

$b2 \;\; c2, p \;\;$ .0001—but from universal logic and universal innate ideas. Jean Piaget would later show that children acquire basic beliefs about physical reality, such as conservation of physical properties, without being tutored. The source and manner of acquisition of these kinds of beliefs divided philosophers and divide cognitive scientists. Plato's great student was Aristotle, but he differed sharply from his teacher. For present purposes, two differences were paramount. The first was a difference of temperament and cast of mind. Plato's philosophy had a religious cast to it, with its soul–body dualism, reincarnation, and positing of heavenly Forms. Aristotle was basically a scientist, his specialty being marine biology. Aristotle rejected the transcendental world of the Forms, although he did not give up on universal truths. Second, and in part a consequence of the first, Aristotle was an empiricist. He believed universal concepts were built up by noting similarities and differences between the objects of one's experience. Thus, the concept of *cat* would consist of the features observably shared by all cats. Postulating Forms and innate ideas of them was unnecessary, said Aristotle. Nevertheless, Aristotle retained Plato's idea that there is a universal and eternal essence of *catness,* or of any other universal concept. He did not believe, as later empiricists would, that concepts are human constructions. Aristotle was arguably the first cognitive scientist (Nussbaum & Rorty, 1992). Socrates was interested in teaching compelling moral truths and said little about the psychology involved. With his distrust of the senses and otherworldly orientation, Plato, too, said little about the mechanisms of perception or thought. Aristotle, the scientist, who believed all truths begin with sensations of the external world, proposed sophisticated theories of the psychology of cognition. His treatment of the animal and human mind may be cast, somewhat anachronistically, of course, in the form of an information-processing diagram (Figure 6.2).

Cognitive processing begins with sensation of the outside world by the *special senses,* each of which registers one type of sensory information. Aristotle recognized the existence of what would later be called the problem of sensory integration, or the binding problem. Experience starts out with the discrete and qualitatively very different sensations of sight, sound, and so forth. Yet we experience not a whirl of unattached sensations (William James's famous "blooming, buzzing, confusion") but coherent objects possessing multiple sensory features. Aristotle posited a mental faculty—today cognitive

scientists might call it a mental module—to handle the problem. *Common sense* integrated the separate streams of sensation into perception of a whole object.

This problem of object perception or pattern recognition remains a source of controversy in cognitive psychology and artificial intelligence. Images of objects could be held before the mind's eye by *imagination* and stored away in, and retrieved from, *memory*. So far, we have remained within the mind of animals, Aristotle's *sensitive soul*. Clearly, animals perceive the world of objects and can learn, storing experiences in memory. Humans are unique in being able to form universal concepts; dogs store memories of particular cats they have encountered but do not form the abstract concept *cat*. This is the function of the human soul, or *mind*. Aristotle drew a difficult distinction between active and passive mind. Roughly speaking, *passive mind* is the store of universal concepts, while *active mind* consists in the cognitive processes that build up knowledge of universals. Aristotle's system anticipates Tulving's (1972) influential positing of episodic and semantic memory. Aristotle's memory is Tulving's episodic memory, the storehouse of personal experiences. Aristotle's passive mind is Tulving's semantic memory, the storehouse of universal concepts.

The Hellenistic, Roman, and Medieval Periods

The death of Aristotle's famous pupil Alexander the Great in 323 B.C.E. marked an important shift in the nature of society and of philosophy. The era of the autonomous city-state was over; the era of great empires began. In consequence, philosophy moved in a more practical, almost psychotherapeutic (Nussbaum, 1994) direction. Contending schools of philosophy claimed to teach recipes for attaining happiness in a suddenly changed world. Considerations of epistemology and cognition faded into the background.

Nevertheless, the orientations to cognition laid down earlier remained and were developed. Those of Socrates' students who gave up on his and Plato's ambition to find transcendental truths developed the philosophy of skepticism. They held that no belief should be regarded as certain but held only provisionally and as subject to abandonment or revision. The Cynics turned Socrates' attack on social convention into a lifestyle. They deliberately flouted Greek traditions and sought to live as much like animals as possible. While cynicism looks much like skepticism—both attack cultural conventions as mere opinions—it did not reject Socrates' quest for moral truth. The Cynics lived what they believed was the correct human way of life free of conventional falsehoods. The Neoplatonists pushed Plato's faith in heavenly truth in a more religious direction, ultimately merging with certain strands of Christian philosophy in the work of Augustine and others. Of all the schools, the most important was Stoicism, taught widely throughout the Roman Empire. Like Plato, the Stoics believed that there was a realm of Transcendental Being beyond our world of appearances, although they regarded it as like a living and evolving organism, transcendent but not fifixed eternally like the Forms. Also like Plato, they taught that logic—reason—was the path to transcendental knowledge.

Hellenistic and medieval physician-philosophers continued to develop Aristotle's cognitive psychology. They elaborated on his list of faculties, adding new ones such as *estimation*, the faculty by which animals and humans intuit whether a perceived object is beneficial or harmful. Moreover, they sought to give faculty psychology a physiological basis. From the medical writings of antiquity, they believed that mental processes are carried out within the various ventricles of the brain containing cerebrospinal fluid. They proposed that each mental faculty was housed in a distinct ventricle of the brain and that the movement of the cerebrospinal fluid through each ventricle in turn was the physical basis of information processing through the faculties. Here is the beginning of cognitive neuroscience and the idea of localization of cerebral function.

Summary: Premodern Realism

Although during the premodern period competing theories of cognition were offered, virtually all the premodern thinkers shared one assumption I will call *cognitive realism*. Cognitive realism is the claim that when we perceive an object under normal conditions, we accurately grasp all of its various sensory features.

Classical cognitive realism took two forms. One, *perceptual realism*, may be illustrated by Aristotle's theory of perception. Consider my perception of a person some meters distant. His or her appearance comprises a number of distinct sensory features: a certain height, hair color, cut and color of clothing, gait, timber of voice, and so on. Aristotle held that each of these features was picked up by the corresponding special sense. For example, the blue of a shirt caused the fluid in the eye to become blue; I see the shirt as blue because it is blue. At the level of the special senses, perception reveals the world as it *really* is. Of course, we sometimes make mistakes about the object of perception, but Aristotle attributed such mistakes to common sense, when we integrate the information from the special senses. Thus, I may mistakenly think that I'm approaching my daughter on campus, only to find that it's a similar-looking young woman. The important point is that for Aristotle my error is one of judgment, not of sensation: I really did see a slender young woman about 59 tall in a leopard-print dress and hair dyed black; my mistake came in thinking it was Elizabeth.

Plato said little about perception because he distrusted it, but his *metaphysical realism* endorsed conclusions similar to, and even stronger than, Aristotle's. Plato said that we identify an individual cat as a cat because it resembles the Form of the Cat in heaven and lodged innately in our soul. If I say that a small fluffy dog is a cat, I am in error, because the dog really resembles the Form of the Dog. Moreover, Plato posited the existence of higher-level forms such as the Form of Beauty or the Form of the Good. Thus, not only is a cat a cat because it resembles the Form of the Cat, but a sculpture or painting is objectively beautiful because it resembles the Form of Beauty, and an action is objectively moral because it resembles the Form of the Good. For Plato, if I say that justice is the rule of the strong, I am in error, for tyranny does not resemble the Form of the Good. We act unjustly only to the extent our knowledge of the Good is imperfect.

Premodern relativism and skepticism were not inconsistent with cognitive realism, because they rested on distrust of human thought, not sensation or perception. One might believe in the world of the Forms but despair of our ability to know them, at least while embodied in physical bodies. This was the message of Neoplatonism and the Christian thought it influenced. Sophists liked to argue both sides of an issue to show that human reason could not grasp enduring truth, but they did not distrust their senses. Likewise, the skeptics were wary of the human tendency to jump to conclusions and taught that to be happy one should not commit oneself wholeheartedly to any belief, but they did not doubt the truth of individual sensations.

The Scientific Revolution and a New Understanding of Cognition

The Scientific Revolution marked a sharp, almost absolute, break in theories of cognition. It presented a new conception of the world: the world as a machine (Henry, 1997). Platonic metaphysical realism died. There were no external, transcendental standards by which to judge what was beautiful or just, or even what was a dog and what was a cat. The only reality was the material reality of particular things, and as a result the key cognitive relationship became the relationship between a perceiver and the objects in the material world he perceives and classifies, not the relationship between the object perceived and the Form it resembles. Aristotle's perceptual realism died, too, as scientists and philosophers imposed a veil of ideas between the perceiver and the world perceived. This veil of ideas was consciousness, and it created psychology as a discipline as well as a new set of problems in the philosophy and psychology of cognition.

The Way of Ideas: Rejecting Realism

Beginning with Galileo Galilei (1564–1642), scientists distinguished between *primary* and *secondary* sense properties (the terms are John Locke's). Primary sense properties are those that actually belong to the physical world-machine; they are objective. Secondary properties are those added to experience by our sensory apparatus; they are subjective. Galileo wrote in his book *The Assayer:* Whenever I conceive any material or corporeal substance I immediately . . . think of it as bounded, and as having this or that shape; as being large or small [and] as being in motion or at rest.... From these conditions I cannot separate such a substance by any stretch of my imagination. But that it must be white or red, bitter or sweet, noisy or silent, and of sweet or foul odor, my mind does not feel compelled to bring in as necessary accompaniments. . . . Hence, I think that tastes, odors, colors, and so on . . . reside only in the consciousness [so that] if the living creature were removed all these qualities would be wiped away and annihilated.

The key word in this passage is *consciousness.* For ancient philosophers, there was only one world, the real physical world with which we are in direct touch, though the Platonists added the transcendental world of the Forms, but it, too, was external to us. But the concept of secondary sense properties created a New World, the inner world of consciousness, populated by mental objects—*ideas*—possessing sensory properties not found in objects themselves. In this new representational view of cognition—the Way of Ideas—we perceive objects not directly but indirectly via representations— ideas—found in consciousness. Some secondary properties correspond to physical features objects actually possess. For example, color corresponds to different wavelengths of light to which retinal receptors respond. That color is not a primary property, however, is demonstrated by the existence of colorblind individuals, whose color perception is limited or absent. Objects are not colored, only ideas are colored. Other secondary properties, such as being beautiful or good, are even more troublesome, because they seem to correspond to no physical facts but appear to reside only in consciousness.

Our modern opinion that beauty and goodness are subjective judgments informed by cultural norms is one consequence of the transformation of experience wrought by the Scientific Revolution.

Cartesian Dualism and the Veil of Ideas

For psychology, the most important modern thinker was René Descartes (1596–1650), who created an influential framework for thinking about cognition that was fundamental to the history of psychology for the next 350 years. Descartes' dualism of body and soul is well known, but it also included the new scientifific distinction of physical and mental worlds. Descartes assumed living bodies were complex machines no different from the world-machine. Animals lacked soul and consciousness and were therefore incapable of cognition. As machines, they responded to the world, but they could not think about it. Human beings were animals, too, but inside their mechanical body dwelled the soul, possessor of consciousness. Consciousness was the New World of ideas, indirectly representing the material objects encountered by the senses of the body. Descartes' picture has been aptly called the Cartesian Theater (Dennett, 1991): The soul sits inside the body and views the world as on a theater screen, a veil of ideas interposed between knowing self and known world.

Within the Cartesian framework, one could adopt two attitudes toward experience. The first attitude was that of natural science. Scientists continued to think of ideas as partial reflections of the physical world. Primary properties corresponded to reality; secondary ones did not, and science dealt only with the former. However, the existence of a world of ideas separate from the world of things invited exploration of this New World, as explorers were then exploring the New World of the Western Hemisphere. The method of natural science was observation. Exploring the New World of Consciousness demanded a new method, introspection. One could examine ideas as such, not as projections from the world outside, but as objects in the subjective world of consciousness.

Psychology was created by introspection, reflecting on the screen of consciousness. The natural scientist inspects the objective natural world of physical objects; the psychologist introspects the subjective mental world of ideas. To psychologists was given the problem of explaining whence secondary properties come. If color does not exist in the world, why and how do we see color? Descartes also made psychology important for philosophy and science. For them to discover the nature of material reality, it became vital to sort out what parts of experience were objective and what parts were subjective chimeras of consciousness. From now on, the psychology of cognition became the basis for epistemology. In order to know what people can and ought to know, it became important to study how people actually do know. But these investigations issued in a crisis when it became uncertain that people know—in the traditional Classical sense—anything at all.

The Modern Period: Cognition after the Scientific Revolution

Several intertwined questions arose from the new scientific, Cartesian, view of mind and its place in nature. Some are philosophical. If I am locked up in the subjective world of consciousness, how can I know anything about the world with any confidence? Asking this question created a degree of paranoia in subsequent philosophy. Descartes began his quest for a foundation upon which to erect science by suspecting the truth of every belief he had. Eventually he came upon the apparently unassailable assertion that "I think, therefore I am." But Descartes'method placed everything else in doubt, including the existences of God and the world. Related to the philosophical questions are psychological ones. How and why does consciousness work as it does? Why do we experience the world as we do rather than some other way? Because the answers to the philosophical questions depend on the answers to the psychological ones, examining the mind—doing psychology—became the central preoccupation of philosophy before psychology split off as an independent discipline.

Three philosophical-psychological traditions arose out of the new Cartesian questions: the modern *empiricist*, *realist*, and *idealist* traditions. They have shaped the psychology of cognition ever since.

The Empiricist Tradition

Notwithstanding the subjectivity of consciousness, empiricism began with John Locke (1632–1794), who accepted consciousness at face value, trusting it as a good, if imperfect, reflection of the world. Locke concisely summarized the central thrust of empiricism: "We should not judge of things by men's opinions, but of opinions by things," striving to know "the things themselves." Locke's picture of cognition is essentially Descartes'. We are acquainted not with objects but with the *ideas* that represent them. Locke differed from Descartes in denying that any of the mind's ideas are innate. Descartes had said that some ideas (such as the idea of God) cannot be found in experience but are inborn, awaiting activation by appropriate experiences. Locke said that the mind was empty of ideas at birth, being a *tabula rasa,* or blank slate, upon which experience writes. However, Locke's view is not too different from Descartes', because he held that the mind is furnished with numerous mental abilities, or faculties, that tend automatically to produce certain universally held ideas (such as the idea of God) out of the raw material of experience. Locke distinguished two sources of experience, sensation and reflection. Sensation reveals the outside world, while reflection reveals the operations of our minds.

Later empiricists took the Way of Ideas further, creating deep and unresolved questions about human knowledge. The Irish Anglican bishop and philosopher George Berkeley (1685–1753) began to reveal the startling implications of the Way of Ideas. Berkeley's work is an outstanding example of how the new Cartesian conception of consciousness invited psychological investigation of beliefs heretofore taken for granted. The Way of Ideas assumes with common sense that there is a world outside consciousness. However, through a penetrating analysis of visual perception, Berkeley challenged that assumption. The world of consciousness is three dimensional, possessing height, width, and depth. However, Berkeley pointed out, visual perception begins with a flat, two-dimensional image on the retina, having only height and width. Thus, as someone leaves us, we *experience* her as getting farther away, while *on the retina* there is only an image getting smaller and smaller.

Berkeley argued that the third dimension of depth was a secondary sense property, a subjective construction of the Cartesian Theater. We infer the distance of objects from information on the retina (such as linear perspective) and from bodily feedback about the operations of our eyes. Painters use the first kind of cues on canvases to create illusions of depth. So far, Berkeley acted as a psychologist proposing a theory about visual perception. However, he went on to develop a striking philosophical position called immaterialism.

Depth is not only an illusion when it's on canvas, it's an illusion on the retina, too. Visual experience is, in fact, two dimensional, and the third dimension is a psychological construction out of bits and pieces of experience assembled by us into the familiar three-dimensional world of consciousness. Belief in an external world depends upon belief in three dimensional space, and Berkeley reached the breath taking conclusion that there is no world of physical objects at all, only the world of ideas. Breath taking Berkeley's conclusion may be, but it rests on hard headed reasoning. Our belief that objects exist independently of our experience of them—that my car continues to exist when I'm indoors—is an act of faith. Jean Piaget and other cognitive developmentalists later extensively studied how children develop belief in the permanence of physical objects. This act of faith is regularly confirmed, but Berkeley said we have no knockdown *proof* that the world exists outside the Cartesian Theater. We see here the paranoid tendency of modern thought, the tendency to be skeptical about every belief, no matter how innocent— true—it may seem, and in Berkeley we see how this tendency depends upon psychological notions about the mind.

Skepticism was developed further by David Hume (1711–1776), one of the most important modern thinkers, and his skeptical philosophy began with psychology: "[A]ll the sciences have a relation . . . to human nature," and the only foundation "upon which they can stand" is the "science of human nature." Hume drew out the skeptical implications of the Way of Ideas by relentlessly applying empiricism to every commonsense belief. The world with which we are acquainted is world of ideas, and the mental force of association holds ideas together. In the world of ideas, we may conceive of things that do not actually exist but are combinations of simpler ideas that the mind combines on its own. Thus, the chimerical unicorn is only an idea, being a combination of two other ideas that do correspond to objects, the idea of a horse and the idea of a horn. Likewise, God is a chimerical idea, composed out of ideas about omniscience, omnipotence, and paternal love. The self, too, dissolves in Hume's inquiry. He went looking for the self and could find in consciousness nothing that was not a sensation of the world or the body. A good empiricist, Hume thus concluded that because it cannot be observed, the self is a sort of psychological chimera, though he remained uncertain how it was constructed. Hume expunged the soul in the Cartesian Theater, leaving its screen as the only psychological reality.

Hume built up a powerful theory of the mechanics of cognition based on association of ideas. The notion that the mind has a natural tendency to link certain ideas together is a very old one, dating back to Aristotle's speculations about human memory. The term "association of ideas" was coined by Locke, who recognized its existence but viewed it as a baleful force that threatened to replace rational, logical, trains of thought with nonrational ones. Hume, however, made association into the "gravity" of the mind, as supreme in the mental world as Newton's gravity was in the physical one. Hume proposed three laws that governed how associations formed: the law of similarity (an idea presented to the mind automatically conjures up ideas that resemble it); the law of contiguity (ideas presented to the mind together become linked, so that if one is presented later, the other will automatically be brought to consciousness), and the law of causality (causes make us automatically think of their effects; effects make us automatically think of their causes). After Hume, the concept of association of ideas would gain ground, becoming a dominant force in much of philosophy and psychology until the last quarter of the twentieth century. Various philosophers, especially in Britain, developed rival theories of association, adumbrating various different laws of associative learning.

The physician David Hartley (1705–1757) speculated about the possible neural substrates of association formation. Associative theory entered psychology with the work of Ebbinghaus. Human psychology seemed to make scientifific knowledge unjustifiable. Our idea of causality—a basic tenet of science— is chimerical. We do not see causes themselves, only regular sequences of events, to which we add a subjective feeling, the feeling of a necessary connection between an effect and its cause. More generally, any universal assertion such as "All swans are white" cannot be proved, because they have onlybeen confirmed by experience so far. We might one day find that some swans are black (they live in New Zealand). To critics, Hume had reached the alarming conclusion that we can know nothing for certain beyond the immediate content of our conscious sensations. Science, religion, and morality were all thrown in doubt, because all assert theses or depend on assumptions going beyond experience and which may therefore some day prove erroneous. Hume was untroubled by this conclusion, anticipating later post evolutionary pragmatism.

Beliefs formed by the human mind are not provable by rational argument, Hume said, but they are reasonable and useful, aiding us mightily in everyday life. Other thinkers, however, were convinced that philosophy had taken a wrong turn.

The Realist Tradition

Hume's fellow Scottish philosophers, led by Thomas Reid (1710–1796), offered one diagnosis and remedy. Berkeley and Hume challenged common sense, suggesting that external objects do not exist, or, if they do, we cannot know them or causal relationships among them with any certainty. Reid defended common sense against philosophy, arguing that the Way of Ideas had led philosophers into a sort of madness.

Reid reasserted and reworked the older realist tradition. We see objects themselves, not inner representations of them. Because we perceive the world directly, we may dismiss Berkeley's immaterialism and Hume's skepticism as absurd consequences of a mistaken notion, the Way of Ideas. Reid also defended a form of nativism. God made us, endowing us with mental powers—faculties—upon which we can rely to deliver accurate information about the outside world and its operations.

The Idealist Tradition

Another diagnosis and remedy for skepticism was offered in Germany by Immanuel Kant (1724–1804), who, like Reid, found Hume's ideas intolerable because they made genuine knowledge unreachable. Reid located Hume's error in the Way of Ideas, abandoning it for a realist analysis of cognition. Kant, on the other hand, located Hume's error in empiricism and elaborated a new version of the Way of Ideas that located truth inside the mind. Empiricists taught that ideas reflflect, in Locke's phrase, "things themselves," the mind conforming itself to objects that impress (Hume's term) themselves upon it. But for Kant, skepticism deconstructed empiricism. The assumption that mind reflflects reality is but an assumption, and once this assumption is revealed—by Berkeley and Hume— the ground of true knowledge disappears. Kant upended the empiricist assumption that the mind conforms itself to objects, declaring that objects conform themselves to the mind, which imposes a universal, logically necessary structure upon experience. Things in themselves— *noumena*—are unknowable, but things as they appear in consciousness—*phenomena*—are organized by mind in such a way that we can make absolutely true statements about them. Take, for example, the problem addressed by Berkeley, the perception of depth. Things in themselves may or may not be arranged in Euclidean three-dimensional space; indeed, modern physics says that space is non-Euclidean. However, the human mind imposes Euclidean three-dimensional space on its experience of the world, so we can say truly that phenomena are necessarily arrayed in three-dimensional space.

Similarly, the mind imposes other Categories of experience on noumena to construct the phenomenal world of human experience. A science fifiction example may clarify Kant's point. Imagine the citizens of Oz, the Emerald City, in whose eyes are implanted at birth contact lenses making everything a shade of green. Ozzites will make the natural assumption that things *seem* green because things *are* green. However, Ozzites' phenomena are green because of the contact lenses, not because things in themselves are green. Nevertheless, the Ozzites can assert as an absolute and irrefutable truth, "Every phenomenon is green." Kant argued that the Categories of experience are logically necessary preconditions of any experience whatsoever by all sentient beings. Therefore, since science is about the world of phenomena, we can have genuine, irrefutable, absolute knowledge of that world and should give up inquiries into Locke's "things themselves." Kantian idealism produced a radically expansive view of the self. Instead of concluding with Hume that it is a construction out of bits and pieces of experience, Kant said that it exists prior to experience and imposes order on experience. Kant distinguished between the Empirical Ego—the fleeting contents of consciousness—and the Transcendental Ego. The Transcendental Ego is the same in all minds and imposes the Categories of understanding on experience. The self is not a construction out of experience; it is the active constructor of experience. In empiricism the self vanished; in idealism it became the only reality.

Summary: Psychology Takes Center Stage

Nineteenth-century philosophers elaborated the empiricist, realist, and idealist philosophical theories of cognition, but their essential claims remained unchanged. The stage was set for psychologists to investigate cognition empirically.

THE EARLY SCIENTIFIC PERIOD

Contemporary cognitive scientists distinguish between *procedural* and *declarative* learning, sometimes known as *knowing how* and *knowing that* (Squire, 1994). Although the distinction was drawn only recently, it will be useful for understanding the study of cognition and learning in the Early Scientifific Period. A paradigmatic illustration of the two forms of learning or knowing is bicycle riding. Most of us know *how* to ride a bicycle (procedural learning), but few of us know the physical and physiological principles*that* are involved (declarative learning).

The Psychology of Consciousness

With the exception of comparative psychologists (see following), the founding generation of scientifific psychologists studied human consciousness via introspection (Leahey, 2000). They were thus primarily concerned with the processes of sensation and perception, which are discussed in another chapter of this handbook. Research and theory continued to be guided by the positions already developed by philosophers. Most psychologists, including Wilhelm Wundt, the traditional founder of psychology, adopted one form or another of the Way of Ideas, although it was vehemently rejected by the gestalt psychologists, who adopted a form of realism proposed by the philosopher Franz Brentano (1838–1917; Leahey, 2000).

The Verbal Learning Tradition

One psychologist of the era, however, Hermann Ebbinghaus (1850–1909), was an exception to the focus on conscious experience, creating the experimental study of learning with his *On Memory* (1885). Ebbinghaus worked within the associative tradition, turning philosophical speculation about association formation into a scientifific research program, the verbal learning tradition. Right at the outset, he faced to a problem that has bedeviled the scientifific study of human cognition, making a methodological decision of great longterm importance. One might study learning by giving subjects things such as poems to learn by heart. Ebbinghaus reasoned, however, that learning a poem involves two mental processes, comprehension of the meaning of the poem and learning the words in the right order. He wanted to study the latter process, association formation in its pure state. So he made up nonsense syllables, which, he thought, had no meaning. Observe that by excluding meaning from his research program, Ebbinghaus studied procedural learning exclusively, as would the behaviorists of the twentieth century.

Ebbinghaus's nonsense syllables were typically consonantvowel-consonant (CVC) trigrams (to make them pronounceable), and for decades to come, thousands of subjects would learn hundreds of thousands of CVC lists in serial or paired associate form. Using his lists, Ebbinghaus could empirically investigate traditional questions philosophers had asked about associative learning. How long are associations maintained? Are associations formed only between CVCs that are adjacent, or are associations formed between remote syllables? Questions like these dominated the study of human learning until about 1970. The verbal learning tradition died for internal and external reasons. Internally, it turned out that nonsense syllables were not really meaningless, undermining their raison d'etre. Subjects privately turned nonsense into meaning by various strategies. For example, RIS looks meaningless, but could be reversed to mean SIR, or interpreted as the French word for rice. Externally, the cognitive psychologists of the so-called cognitive revolution (Leahey, 2000) wanted to study complex mental processes, including meaning, and rejected Ebbinghaus's procedures as simplistic.

The Impact of Evolution

From the time of the Greeks, philosophers were concerned exclusively with declarative cognition. Recall the warrior, jurist, and connoisseur discussed in connection with Socrates. Each was flflawless in his arena of competence, the battlefield, the courtroom, and the art gallery, knowing how to fifight, judge, and appreciate. Yet Socrates denied that they possessed real knowledge, because they could not state the principles guiding their actions. Exclusive concern with declarative cognition was codified in its modern form by Descartes, for whom knowledge was the preserve of human beings, who uniquely possessed language in which knowledge was formulated and communicated. Action was the realm of the beast-machine, not the human, knowing soul.

Evolution challenged philosophers' preoccupation with declarative knowledge. To begin with, evolution erased the huge and absolute gap Descartes had erected between human mind and animal mindlessness. Perhaps animals possessed simpler forms of human cognitive processes; this was the thesis of the fifirst comparative psychologists and of today's students of animal cognition (Vauclair, 1996). On the other hand, perhaps humans were no more than complex animals, priding themselves on cognitive powers they did not really possess; this was the thesis of many behaviorists (see below).

Second, evolution forced the recognition that thought and behavior were inextricably linked. What counted in Darwin's struggle for existence was survival and reproduction, not thinking True thoughts. The American movement of pragmatism assimilated evolution into philosophy, recognizing the necessary connection between thought and behavior and formulating evolution's new criterion of truth, usefulness. The fifirst pragmatist paper, "How to Make Our Ideas Clear," made the fifirst point. C. S. Peirce (1838–1914) (1878) wrote that "the whole function of thought is to produce habits of action," and that what we call beliefs are "a rule of action, or, say for short, a habit." "The essence of belief," Peirce argued, "is the establishment of a habit, and different beliefs are distinguished by the different modes of action to which they give rise." Habits must have a practical significance if they are to be meaningful, Peirce went on: "Now the identity of a habit depends on how it might lead us to act. . . . Thus we come down to what is tangible and conceivably practical as the root of every real distinction of thought... there is no distinction so fine as to consist in anything but a possible difference in practice." In conclusion, "the rule for attaining [clear ideas] is as follows: consider what effects, which might conceivably have practical bearings, we conceive the object of our conceptions to have.

Then, our conception of these effects is the whole of our conception of the object" (Peirce, 1878/1966, p. 162). William James (1842–1910) made the second point in *Pragmatism* (1905, p. 133): True ideas are those that we can assimilate, validate, corroborate and verify. False ideas are those that we can not. That is the practical difference it makes for us to have true ideas. . . . The truth of an idea is not a stagnant property inherent in it. Truth happens to an idea. It becomes true, is made true by events. Its verity is in fact an event, a process. Peirce and James rejected the philosophical search for transcendental Truth that had developed after Plato. For pragmatism there is no permanent truth, only a set of beliefs that change as circumstances demand.

With James, philosophy became psychology, and scientifific psychology began to pursue its own independent agenda. Philosophers continued to struggle with metaphysics and epistemology—as James himself did when he returned to philosophy to develop his radical empiricism—but psychologists concerned themselves with effective behavior instead of truth.

Animal Psychology and the Coming of Behaviorism

In terms of psychological theory and research, the impact of evolution manifested itself fifirst in the study of animal mind and behavior. As indicated earlier, erasing the line between humans and animals could shift psychological thinking in either of two ways. First, one might regard animals as more humanlike than Descartes had, and therefore as capable of some forms of cognition. This was the approach taken by the fifirst generation of animal psychologists beginning with George John Romanes (1848–1894). They sought to detect signs of mental life and consciousness in animals, attributing consciousness, cognition, and problem solving abilities to even very simple creatures (Romanes, 1883). While experiments on animal behavior were not eschewed, most of the data Romanes and others used were anecdotal in nature. Theoretically, inferring mental processes from behavior presented difficulties. It is tempting to attribute to animals complex mental processes they may not possess, as we imagine ourselves in some animal's predicament and think our way out. Moreover, attribution of mental states to animals was complicated by the prevailing Cartesian equation of mentality with consciousness. The idea of unconscious mental states, so widely accepted today, was just beginning to develop, primarily in German post-Kantian idealism, but it was rejected by psychologists, who were followers of empiricism or realism (Ash, 1995). In the Cartesian framework, to attribute complex mental states to animals was to attribute to them *conscious* thoughts and beliefs, and critics pointed out that such inferences could not be checked by introspection, as they could be in humans. (At this same time, the validity of human introspective reports was becoming suspect, as well, strengthening critics' case again the validity of mentalist animal psychology; see Leahey, 2000.) C. Lloyd Morgan

(1852–1936) tried to cope with these problems with his famous canon of simplicity and by an innovative attempt to pry apart the identification of mentality with consciousness. Morgan (1886) distinguished objective inferences from projective—or, as he called them in the philosophical jargon of his time, ejective—inferences from animal behavior to animal mind. Imagine watching a dog sitting at a street corner at 3:30 one afternoon. As a school bus approaches, the dog gets up, wags its tail, and watches the bus slow down and then stop. The dog looks at the children getting off the bus and, when one boy gets off, it jumps on him, licks his face, and together the boy and the dog walk off down the street. Objectively, Morgan would say, we may infer certain mental powers possessed by the dog. It must possess sufficient perceptual skills to pick out one child from the crowd getting off the bus, and it must possess at least recognition memory, for it responds differently to one child among all the others. Such inferences are objective because they do not involve analogy to our own thought processes. When we see an old friend, we do not consciously match up the face we see with a stored set of remembered faces, though it is plain that such a recognition process must occur. In making an objective inference, there is no difference between our viewpoint with respect to our own behavior and with respect to the dog's, because in each case the inference that humans and dogs possess recognition memory is based on observations of behavior, not on introspective access to consciousness.

Projective inferences, however, are based on drawing unprovable analogies between our own consciousness and putative animal consciousness. We are tempted to attribute a subjective mental state, happiness, to the watchful dog by analogy with our own happiness when we greet a loved one who has been absent. Objective inferences are legitimate in science, Morgan held, because they do not depend on analogy, are not emotional, and are susceptible to later verifification by experiment. Projective inferences are not scientifically legitimate because they result from attributing our own feelings to animals and may not be more objectively assessed. Morgan's distinction is important, and although it is now the basis of cognitive science, it had no contemporary impact.

In the event, skepticism about mentalistic animal psychology mounted, especially as human psychology became more objective. Romanes (1883, pp. 5–6) attempted to deflflect his critics by appealing to our everyday attribution of mentality to other people without demanding introspective verification: "Skepticism of this kind is logically bound to deny evidence of mind, not only in the case of lower animals, but also in that of the higher, and even in that of men other than the skeptichim self. For all objections which could apply to the use of [inference] . . . would apply with equal force to the evidence of any mind other than that of the individual objector" (pp. 4–5).

Two paths to the study of animal and human cognition became clearly defifined. One could continue with Romanes and Morgan to treat animals and humans as creatures with minds; or one could accept the logic of Romanes's rebuttal and treat humans and animals alike as creatures without minds. Refusing to anthropomorphize humans was the beginning of behaviorism, the study of learning without cognition.

Behaviorism: The Golden Age of Learning Theory

With a single exception, E. C. Tolman (see following), behaviorism fifirmly grasped the second of the two choices possible within the Cartesian framework. They chose to treat humans and animals as Cartesian beast-machines whose behavior could be fully explained in mechanistic causal terms without reference to mental states or consciousness. They thus dispensed with cognition altogether and studied procedural learning alone, examining how behavior is changed by exposure to physical stimuli and material rewards and punishments. Behaviorists divided on how to treat the stubborn fact of consciousness. Methodological behaviorists admitted the existence of consciousness but said that its private, subjective nature excluded it from scientifific study; they left it the arts to express, not explain, subjectivity. Metaphysical behaviorists had more imperial aims. They wanted to explain consciousness scientifically, ceding nothing to the humanities (Lashley, 1923).

Methodological Behaviorism

Although methodological behaviorists agreed that consciousness stood outside scientifific psychology, they disagreed about how to explain behavior. The dominant tradition was the stimulus-response tradition originating with Thorndike, and carried along with modifification by Watson, Hull, and his colleagues, and the mediational behaviorists of the 1950s. They all regarded learning as a matter of strengthening or weakening connections between environmental stimuli and the behavioral response they evoked in organisms. The most important rival form of methodological behaviorism was the cognitive-purposive psychology of Tolman and his followers, who kept alive representational theories of learning. In short, the stimulus-response tradition studied how organisms react to the world; the cognitive tradition studied how organisms learn about the world. Unfortunately, for decades it was not realized that these were complementary rather than competing lines of investigation.

Stimulus-Response Theories.

By far the most influential learning theories of the Golden Age of Theory were stimulus-response (S-R) theories. S-R theorizing began with Edward Lee Thorndike's (1874–1949) connectionism. Thorndike studied animal learning for his 1898 dissertation, published as *Animal Learning* in 1911. He began as a conventional associationist studying association of ideas in animals. However, as a result of his studies he concluded that while animals make associations, they do not associate ideas: "The effective part of the association [is] a direct bond between the situation and the impulse [to behavior]" (Thorndike, 1911, p. 98). Thorndike constructed a number of puzzle boxes in which he placed one of his subjects, typically a young cat. The puzzle box was a sort of cage so constructed that the animal could open the door by operating a manipulandum that typically operated a string dangling in the box, which in turn ran over a pulley and opened the door, releasing the animal, who was then fed before being placed back in the box. Thorndike wanted to discover how the subject learns the correct response. He described what happens in a box in which the cat must pull a loop or button on the end of the string:

The cat that is clawing all over the box in her impulsive struggle will probably claw the string or loop or button so as to open the door. And gradually all the other nonsuccessful impulses will be stamped out and the particular impulse leading to the successful act will be stamped in by the resulting pleasure, until, after many trials, the cat will, when put in the box, immediately claw the button or loop in a defifinite way. (Thorndike, 1911, p. 36)

Thorndike conceived his study as one of associationformation, and interpreted his animals' behaviors in terms of associationism: Starting, then, with its store of instinctive impulses, the cat hits upon the successful movement, and gradually associates it with the sense-impression of the interior of the box until the connection is perfect, so that it performs the act as soon as confronted with the sense-impression. (Thorndike, 1911, p. 38) The phrase *trial-and-error*—or perhaps more exactly trialand-success—learning aptly describes what these animals did in the puzzle boxes. Placed inside, they try out (or, as Skinner called it later, emit) a variety of familiar behaviors.

In cats, it was likely to try squeezing through the bars, clawing at the cage, and sticking its paws between the bars. Eventually, the cat is likely to scratch at the loop of string and so pull on it, fifinding its efforts rewarded: The door opens and it escapes, only to be caught by Thorndike and placed back in the box. As these events are repeated, the useless behaviors die away, or extinguish, and the correct behavior is done soon after entering the cage; the cat has learned the correct response needed to escape. Thorndike proposed three laws of learning. One was the *law of exercise,* which stated that use of a response strengthens its connection to the stimuli controlling it, while disuse weakens them. Another was the *law of readiness*, having to do with the physiological basis of the law of effect.

Thorndike proposed that if the neurons connected to a given action are prepared to fifire (and cause the action), their neural fifiring will be experienced as pleasure, but that if they are inhibited from fifiring, displeasure will be felt. The most famous and debated of Thorndike's laws was the

law of effect:

The Law of Effect is that: Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will, other things being equal, be more fifirmly connected with the situation, so that, when it recurs, they will be more likely to recur; those which are accompanied or closely followed by discomfort to the animal will, other things being equal, have their connections with that situation weakened, so that, when it recurs, they will be less likely to occur. The greater the satisfaction or discomfort, the greater the strengthening or weakening of the bond. (Thorndike, 1911, p. 244)

Thorndike seems here to state a truism not in need of scientifific elaboration, that organisms learn how to get pleasurable things and learn how to avoid painful things. However, questions surround the law of effect. Is reward *necessary* for learning? Reward and punishment surely affect behavior, but must they be present for learning to occur? What about a reward or punishment makes it change behavior? Is it the pleasure and pain they bring, as Thorndike said, or the fact that they inform us that we have just done the right or wrong action? Are associations formed gradually or all at once? Thorndike laid out the core of stimulus-response learning theory. It was developed by several generations of psychologists, including E. R. Guthrie (1886–1959) and most notably by Clark Hull (1884–1952), his collaborator Kenneth Spence (1907–1967), and their legions of students and grandstudents. Hull and Spence turned S-R theory into a formidably complex logico-mathematical structure capable of terrifying students, but they did not change anything essential in Thorndike's ideas. Extensive debate took place on the questions listed above (and others). For example, Hull said reward was necessary for learning, that it operated by drive reduction, and that many trials were needed for an association to reach full strength. Guthrie, on the other hand, said that mere contiguity between S and R was suffificient to form an association between them and that associative bonds reach full strength on a single trial. These theoretical issues, plus those raised by Tolman, drove the copious research of the Golden Age of Theory (Leahey, 2000; Leahey & Harris, 2001). When S-R theorists turned to human behavior, they developed the concept of mediation (Osgood, 1956). Humans, they conceded, had symbolic processes that animals lacked, and they proposed to handle them by invoking covert stimuli and responses. Mediational theories were often quite complex, but the basic idea was simple. A rat learning to distinguish a square-shaped stimulus from a triangular one responds only to the physical properties of each stimulus. An adult human, on the other hand, will privately label each stimulus as "square" or "triangle," and it is this mediating covert labeling response that controls the subject's observable behavior. In this view, animals learned simple one-stage SR connections, while humans learned more sophisticated S-r-s-R connections (where *s* and *r* refer to the covert responses and the stimuli theycause). The great attraction of mediational theory was that it gave behaviorists interested in human cognitive processes a theoretical language shorn of mentalistic connotations (Osgood, 1956), and during the 1950s and early 1960s mediational theories dominated the study of human cognition. However, once the concept of information became available, mediational theorists—and certainly their students—became information processing theorists (Leahey, 2000).

Edward Chace Tolman's Cognitive Behaviorism. E. C.

Tolman (1886–1959) consistently maintained that he was a behaviorist, and in fact wrote a classic statement of methodological behaviorism as a psychological program (Tolman, 1935). However, he was a behaviorist of an odd sort, as he (Tolman, 1959) and S-R psychologists (Spence, 1948) recognized, being inflfluenced by gestalt psychology and the neorealists (see below). Although it is anachronistic to do so, the best way to understand Tolman's awkward position in the Golden Age is through the distinction between procedural and declarative learning. Ebbinghaus, Thorndike, Hull, Guthrie, Spence, and the entire S-R establishment studied only procedural learning. They did not have the procedural/declarative distinction available to them, and in any case thought that consciousness—which formulates and states declarative knowledge—was irrelevant to the causal explanation of behavior. S-R theories said learning came about through the manipulation of physical stimuli and material rewards and punishments. Animals learn, and can, of course, never say why. Even if humans might occasionally figure out the contingencies of reinforcement in a situation, S-R theory said that they were simply describing the causes of their own behavior the way an outside observer does (Skinner, 1957). As Thorndike had said, reward and punishment stamp in or stamp out S-R connections; consciousness had nothing to do with it. Tolman, on the other hand, wanted to study cognition—declarative knowledge in the traditional sense—but was straitjacketed by the philosophical commitments of behaviorism and the limited conceptual tools of the 1930s and 1940s.

Tolman anticipated, but could never quite articulate, the ideas of later cognitive psychology. Tolman's theory and predicament are revealed by his "Disproof of the Law of Effect" (Tolman, Hall, & Bretnall, 1932). In this experiment, human subjects navigated a pegboard maze, placing a metal stylus in the left or right of a series of holes modeling the left-right choices of an animal in a multiple T-maze. There were a variety of conditions, but the most revealing was the "bell-right-shock" group, whose subjects received an electric shock when they put the stylus in the correct holes. According to the Law of Effect these subjects should not learn the maze because correct choices were followed by pain, but they learned at the same rate as other groups. While this result seemed to disprove the law of effect, its real signifificance was unappreciated because the concept of information had not yet been formulated (see below). In Tolman's time, reinforcers (and punishers) were thought of only in terms of their drive-reducing or affective properties. However, they possess informational properties, too. A reward is pleasant and may reduce hunger or thirst, but rewards typically provide information that one has made the correct choice, while punishers are unpleasant and ordinarily convey that one has made the wrong choice. Tolman's "bell-right shock" group pried apart the affective and informational qualities of pain by making pain carry the information that the subject had made the right choice. Tolman showed—but could not articulate—that it's the informational value of behavioral consequences that cause learning, not their affective value.

Nevertheless, Tolman tried to offer a cognitive theory of learning with his concept of cognitive maps (Tolman, 1948). S-R theorists viewed maze learning as acquiring a series of left-right responses triggered by the stimuli at the various choice points in the maze. Against this, Tolman proposed that animals and humans acquire a representation—a mental map—of the maze that guides their behavior. Tolman and his followers battled Hullians through the 1930s, 1940s, and into the 1950s, generating a mass of research fifindings and theoretical argument. Although Tolman's predictions were often vindicated by experimental results, the vague nature of his theory and his attribution of thought to animals limited his theory's impact (Estes et al., 1954).

Metaphysical Behaviorism

Metaphysical behaviorists took a more aggressive stance toward consciousness than methodological behaviorists. They believed that scientifific psychology should explain, not shun, consciousness. Two reasons guided them. First, they wanted to achieve a comprehensive scientifific account of everything human, and since consciousness is undoubtedly something humans have, it should not be ceded to the humanities (Lashley, 1923). Second, stimuli registered only privately in a person's experience sometimes affects behavior (Skinner, 1957). If I have a headache, it exists only in my private consciousness, but it alters my behavior: I take aspirin, become irritable, and tell people I have a headache. Excluding private stimuli from psychology by methodological fifiat would produce incomplete theories of behavior. (This is not the place to discuss the various and subtle ways metaphysical behaviorists had of explaining or dissolving consciousness. I will focus only on how such behaviorists approached learning and cognition.) Metaphysical behaviorism came in two forms, physiological behaviorism and radical behaviorism.

Physiological Behaviorism.

The source of physiological behaviorism was Russian objective psychology, and its greatest American exponent was Karl Lashley, who coined the term "methodological behaviorism," only to reject it (Lashley, 1923, pp. 243–244): Let me cast off the lion's skin. My quarrel with [methodological] behaviorism is not that it has gone too far, but that it has hesitated... that it has failed to develop its premises to their logical conclusion. To me the essence of behaviorism is the belief that the study of man will reveal nothing except what is adequately describable in the concepts of mechanics and chemistry. . . . I believe that it is possible to construct a physiological psychology which will meet the dualist on his own ground . . . and show that [his] data can be embodied in a mechanistic system. . . . Its physiological account of behavior will also be a complete and adequate account of all the phenomena of consciousness... demanding that all psychological data, however obtained, shall be subjected to physical or physiological interpretation.

Ultimately, Lashley said, the choice between behaviorism and traditional psychology came down to a choice between two “incompatible” worldviews, “scientifific versus humanistic.” It had been demanded of psychology heretofore that “it must leave room for human ideals and aspirations.” But “other sciences have escaped this thralldom,” and so must psychology escape from “metaphysics and values” and “mystical obscurantism” by turning to physiology.

For the study of learning, the most important physiological behaviorist was Ivan Petrovich Pavlov (1849–1936). Although Pavlov is mostly thought of as the discoverer of classical or Pavlovian conditioning, he was fifirst and foremost a physiologist in the tradition of Sechenov. For him, the phenomena of Pavlovian conditioning were of interest because they might reveal the neural processes underlying associative learning—he viewed all behavior as explicable via association—and his own theories about conditioning were couched in neurophysiological terms.

The differences between Pavlov’s and Thorndike’s procedures for studying learning posed two questions for the associative tradition they both represented. Pavlov delivered an unconditional stimulus (food) that elicited the behavior, or unconditional response (salivation), that he wished to study. He paired presentation of the US with an unrelated conditional stimulus (only in one obscure study did he use a bell); finding that gradually the CS came to elicit salivation (now called the conditional response), too. Thorndike had to await the cat’s first working of the manipulandum before rewarding it with food. In Pavlov’s setup, the food came first and caused the unconditional response; in Thorndike’s, no obvious stimulus caused the fifirst correct response, and the food followed its execution.

Were Pavlov and Thorndike studying two distinct forms of learning, or were they merely using different methodologies to study the same phenomenon? Some psychologists, including Skinner, believed the former, either on the operationist grounds that the procedures themselves defifined different forms of learning, or because different nervous systems were involved in the two cases (Hearst, 1975). Although this distinction between instrumental (or operant) and classical, or Pavlovian (or respondent) conditioning has become enshrined in textbooks, psychologists in the S-R tradition believed S-R learning took place in both procedures. The debate was never resolved but has been effaced by the return of cognitive theories of animal learning, for which the distinction is not important.

The second question raised by Pavlov's methods was intimately connected to the fifirst. Exactly what was being associated as learning proceeded? In philosophical theory, association took place between ideas, but this mentalistic formulation was, of course, anathema to behaviorists. Thorndike began the S-R tradition by asserting that the learned connection (his preferred term) was directly between stimulus and response, not between mental ideas of the two. Pavlovian conditioning could be interpreted in the same way, saying that the animal began with an innate association between US and UR and created a new association between CS and CR. Indeed, this was for years the dominant behaviorist interpretation of Pavlovian conditioning, the stimulus substitution theory (Leahey & Harris, 2001), because it was consistent with the thesis that all learning was S-R learning.

However, Pavlovian conditioning was open to an alternative interpretation closer to the philosophical notion of association of ideas, which said that ideas that occur together in experience become linked (see above). Thus, one could say that as US and CS were paired, they became associated, so that when presented alone, the CS evoked the US, which in turn caused the CR to occur. Pavlov's own theory of conditioning was a materialistic version of this account, proposing that the brain center activated by the US became neurally linked to the brain center activated by the CS, so when the latter occurred, it activated the US's brain center, causing the CR. American behaviorists who believed in two kinds of learning never adopted Pavlov's physiologizing and avoided mentalism by talking about S-S associations. It was sometimes said that Tolman was an S-S theorist, but this distorted the holistic nature of his cognitive maps. As truly cognitive theories of learning returned in the 1970s, Pavlovian and even instrumental learning were increasingly interpreted involving associations between ideas—now called "representations" (Leahey & Harris, 2001), as in the pioneering cognitive theory of Robert Rescorla (1988).

Radical Behaviorism.

A completely different form of metaphysical behaviorism was developed by B. F. Skinner (1904–1990). Skinner extended to psychology the philosophy of neorealism propounded by a number of American philosophers after 1910 (Smith, 1986). The neorealists revived the old realist claim that the Way of Ideas was mis taken, that perception of objects was direct and not mediated by intervening ideas. Tolman, too, built his early theories on neorealism but later returned to the Way of Ideas with the concept of the cognitive map (Smith, 1986). Skinner never wavered from realism, working out the radical implication that if there are no ideas, there is no private world of consciousness or mind to be populated by them. Introspective psychology was thus an illusion, and psychology should be redefined as studying the interactive relationship between an organism and the environment in which it behaves. The past and present environments provide the stimuli that set the occasion for behavior, and the organism's actions operate (hence the term *operant*) on the environment. Actions have consequences, and these consequences shape the behavior of the organism.

Skinner's thinking is often misrepresented as a S-R psychology in the mechanistic tradition of Thorndike, John B. Watson (1878–1958), or Clark Hull. In fact, Skinner rejected—or, more precisely, stood apart from—the mechanistic way of thinking about living organisms that had begun with Descartes. For a variety of reasons, including its successes, its prestige, and the influence of positivism, physics has been treated as the queen of the sciences, and scientists in other fields, including psychology, have almost uniformly envied it, seeking to explain their phenomena of interest in mechanical causal terms. A paradigmatic case in point was Clark Hull, who acquired a bad case of physics-envy from reading Newton's *Principia*, and his logico-mathematical theory of learning was an attempt to emulate his master. Skinner renounced physics as the model science for the study of behavior, replacing it with Darwinian evolution and selection by consequences (Skinner, 1969). In physical-model thinking, behaviors are caused by stimuli that mechanically provoke them. In evolution, the appearance of new traits is unpredictable, and their fate is determined by the consequences they bring. Traits that favour survival and reproduction increase in frequency over the generations; traits that hamper survival and reproduction decrease in frequency. Similarly, behaviors are emitted, and whether they are retained (learned) or lost (extinguished) depends on the consequences of reinforcement or nonreinforcement.

As a scientist, Skinner, like Thorndike, Hull, and Tolman, studied animals almost exclusively. However, unlike them Skinner wrote extensively about human behavior in a speculative way he called interpretation. His most important such work was *Verbal Behavior* (1957), in which he offered a theory of human cognition. Beginning with Socrates, the central quest of epistemology was understanding the uniquely human ability to form universal concepts, such as *cat, dog,* or *Truth*. From Descartes onward, this ability was linked to language, the unique possession of humans, in which we can state universal defifinitions. In either case, universal concepts were the possession of the human mind, whether as abstract images (Aristotle) or as sentences (Descartes). Skinner, of course, rejected the existence of mind, and therefore of any difference between explaining animal and human behavior. Mediational theorists allowed for an attenuated difference, but Skinner would have none of it. He wrote that although "most of the experimental work responsible for the advance of the experimental analysis of behavior has been carried out on other species . . . the results have proved to be surprisingly free of species restrictions . . . and its methods can be extended to human behavior without serious modification" (Skinner, 1957, p. 3). The final goal of the experimental analysis of behavior is a science of human behavior using the same principles first applied to animals. In *Verbal Behavior*, Skinner offered a behavioristic analysis of universal concepts with the technical term*tact*, and drew out its implications for other aspects of mind and cognition. A tact is a verbal operant under the stimulus control of some part of the physical environment, and the verbal community reinforces correct use of tacts. So a child is reinforced by parents for emitting the sound "dog" in the presence of a dog (Skinner, 1957). Such an operant is called a tact because it "makes con*tact* with" the physical environment. Tacts presumably begin as names (e.g., for the fifirst dog a child learns to label "dog"), but as the verbal community reinforces the emission of the term to similar animals, the tact becomes generalized. Of course, discrimination learning is also involved, as the child will not be reinforced for calling cats "dog." Eventually, through behavior shaping, the child's "dog" response will occur only in the presence of dogs and not in their absence. For Skinner, the situation is no different from that of a pigeon reinforced for pecking keys only when they are illuminated any shade of green and not otherwise. Skinner reduced the traditional notion of reference to a functional relationship among a response, its discriminative stimuli, and its reinforcer.

Skinner's radical analysis of tacting raises an important general point about his treatment of human consciousness, his notion of private stimuli. Skinner believed that earlier methodological behaviorists such as Tolman and Hull were wrong to exclude private events (such as mental images or toothaches) from behaviorism simply because such events are private. Skinner held that part of each person's environment includes the world inside her or his skin, those stimuli to which the person has privileged access. Such stimuli may be unknown to an external observer, but they are experienced by the person who has them, can control behavior, and so must be included in any behaviorist analysis of human behavior. Many verbal statements are under such control, including complex tacts. For example: “My tooth aches” is a kind of tacting response controlled by a certain kind of painful inner stimulation.

This simple analysis implies a momentous conclusion. How do we come to be able to make correct private tacts? Skinner’s answer was that the verbal community has trained us to observe our private stimuli by reinforcing utterances that refer to them. It is useful for parents to know what is distressing a child, so they attempt to teach a child self-reporting verbal behaviors. “My tooth aches” indicates a visit to the dentist, not the podiatrist. Such responses thus have Darwinian survival value. It is these self-observed private stimuli that constitute consciousness. It therefore follows that human consciousness is a product of the reinforcing practices of a verbal community. A person raised by a community that did not reinforce self-description would not be conscious in anything but the sense of being awake. That person would have no selfconsciousness.

Self-description also allowed Skinner to explain apparently purposive verbal behaviors without reference to intention or purpose. For example, “I am looking for my glasses” seems to describe my intentions, but Skinner (1957) argued: “Such behavior must be regarded as equivalent to *When I have behaved in this way in the past, I have found my glasses and have then stopped behaving in this way*” (p. 145). Intention is a mentalistic term Skinner has reduced to the physicalistic description of one’s bodily state. Skinner fifinally attacked the citadel of the Cartesian soul, thinking. Skinner continued to exorcise Cartesian mentalism by arguing that “thought is simply *behavior*.” Skinner rejected Watson’s view that thinking is subvocal behavior, for much covert behavior is not verbal yet can still control overt behavior in a way characteristic of “thinking”: “*I think I shall be going* can be translated *I fifind myself going*” (p. 449), a reference to self-observed, but nonverbal, stimuli.

Skinner’s radical behaviorism was certainly unique, breaking with all other ways of explaining mind and behavior. Its impact, however, has been limited (Leahey, 2000). At the dawn of the new cognitive era, *Verbal Behavior* received a severe drubbing from linguist Noam Chomsky (1959) from which its theses never recovered. The computer model of mind replaced the mediational model and isolated the radical behaviorists. Radical behaviorism carries on after Skinner’s death, but it is little mentioned elsewhere in psychology.

THE MODERN SCIENTIFIC PERIOD

The modern era in the study of cognition opened with the invention of the digital electronic computer during World War II. The engineers, logicians, and mathematicians who created the first computers developed key notions that eventually gave rise to contemporary cognitive psychology.

The Three Key Ideas of Computing *Feedback*

One of the standard objections to seeing living beings as machines was that behavior is purposive and goal-directed, flexibly striving for something not yet in hand (or paw). James (1890) pointed to purposive striving for survival when he called mechanism an "impertinence," and Tolman's retention of purpose as a basic feature of behavior set his behaviorism sharply apart from S-R theories, which treated purpose as something to be explained away (Hull, 1937). Feedback reconciles mechanism and goal-oriented behavior.

As a practical matter, feedback had been employed since the Industrial Revolution. For example, a "governor" typically regulated the temperature of steam engines. This was a rotating shaft whose speed increased as pressure in the engine's boiler increased. Brass balls on hinges were fifitted to the shaft so that as its speed increased, centrifugal force caused the balls to swing away from the shaft. Things were arranged so that when the balls reached a critical distance from the shaft— that is, when the boiler's top safe pressure was reached—heat to the boiler was reduced, the pressure dropped, the balls descended, and heat could return. The system had a purpose— maintain the correct temperature in the boiler—and responded flexibly to relevant changes in the environment—changes of temperature in the boiler. But it was not until World War II that feedback was formulated as an explicit concept by scientists working on the problem of guidance (e.g., building missiles capable of tracking a moving target; Rosenblueth, Wiener, & Bigelow, 1943/1966). The standard example of feedback today is a thermostat. A feedback system has two key components, a sensor and a controller. The sensor detects the state of a relevant variable in the environment. One sets the thermostat tothe critical value of the variable of interest, the temperature of a building. A sensor in the thermostat monitors the temperature, and when it falls below or above critical value, the controller activates the heating or cooling system. When the temperature moves back to its critical value, the sensor detects this and the controller turns off the heat pump. The notion of feedback is that a system, whether living or mechanical, detects a state of the world, acts to alter the state of the world, which alteration is detected, changing the behavior of the system, in a complete feedback loop. A thermostat plus heat pump is thus a purposive system, acting flexibly to pursue a simple goal. It is, of course at the same time a machine whose behavior could be explained in purely causal, physical, terms. Teleology and mechanism are not incompatible.

Information

The concept of information is now so familiar to us that we take it for granted. But in fact it is a subtle concept that engineers building the fifirst computers recognized by the middle of the twentieth century (MacKay, 1969). We have already seen how Tolman could have used it to better understand the nature of reward and punishment. Before the advent of the computer, information was hard to separate from its physical embodiment in parchment or printed pages. Today, however, the separation of information from physical embodiment is a threat to publishers because the content of a book may be scanned and digitized and then accessed by anyone for free. Of course, I could lend someone a book for free, but then I would no longer have its information, but if I share the information itself on a disk or as a download, I still have it, too. The closest the premodern world came to the concept of information was the *idea*, but looking back from our modern vantage point we can see that philosophers tended to assume ideas had to have some kind of existence, either in a transcendent realm apart from the familiar material world, as in Plato, or in a substantial (though nonphysical) soul, Descartes'*res cogitans*. Realists denied that ideas existed, the upshot being

Skinnerian radical behaviorism, which can tolerate the idea of information no more than the idea of a soul. The concept of information allows us to give a more general formulation of feedback. What's important to a feedback system is its use of information, not its mode of physical operation. The thermostat again provides an example. Most traditional thermostats contain a strip of metal that is really two metals with different coeffificients of expansion. The strip then bends or unbends as the temperature changes, turning the heat pump on or off as it closes or opens an electrical circuit. Modern buildings, on the other hand, often contain sensors in each room that relay information about room temperature to a central computer that actually operates the heat pump. Nevertheless, each system embodies the same informational feedback loop.

This fact seems simple, but it is in fact of extraordinary importance. We can think about *information as such*, completely separately from *any* physical embodiment. My description of a thermostat in the preceding section implicitly depended on the concept of information, as I was able to explain what *any* thermostat does without reference to how any *particular* thermostat works. My description of the older steam engine governor, however, depended critically on its actual physical operation.

In any information system we find a kind of dualism. On the one hand, we have a physical object such as a book or thermostat. On the other hand, we have the information it holds or the information processes that guide its operation. The information in the book can be stored in print, in a computer's RAM, on a hard-drive, in bubble memory, or be floating about the World Wide Web. The information flows of a thermostat can be understood without regard to how the thermostat works. This suggests, then, that mind can be understood as information storage (memory) and processes (memory encoding and retrieval, and thinking). Doing so respects the insight of dualism, that mind is somehow independent of body, without introducing all the problems of a substantial soul. Soul is information.

The concept of information opened the way for a new cognitive psychology. One did not need to avoid the mind, as methodological behaviorists wanted, nor did one have to expunge it, as metaphysical behaviorists wanted. Mind was simply information being processed by a computer we only just learned we had, our brains, and we could theorize about information flflows without worrying about how the brain actually managed them. Broadbent's *Perception and Communication* (1958), Neisser's *Cognitive Psychology* (1967), and Atkinson and Shiffrin's "Human Memory: A Proposed System and Its Control Processes" (1968) were the manifestos of the information-processing movement. Broadbent critically proposed treating stimuli as information, not as physical events. Neisser's chapters described information flows from sensation to thinking. Atkinson and Shiffrin's model of information flflow (Figure 6.3) became so standard that it's still found in textbooks today, despite significant changes in the way cognitive psychologists treat the details of cognition (Izawa, 1999).

Information from the senses is first registered in near physical form by sensory memory. The process of pattern recognition assigns informational meaning to the physical stimuli held in sensory memory. Concomitantly, attention focuses on important streams of information, attenuating or blocking others from access to consciousness. Organized information is stored briefly in working, or short-term, memory, and some manages to get stored in long-term, or permanent, memory. There is, of course, loss and distortion of information along the way, so that what's remembered is very seldom a veridical record of what happened.

Only one aspect of contemporary cognitive psychology was missing from Neisser and Atkinson and Shiffrin, the computational metaphor of mind, then just making headway in psychology.

The Program: Computation

In the information-processing perspective developed by Broadbent, Neisser, and Atkinson and Shiffrin, the notion of *processing* remained vague. Information itself is passive: It has to be transformed and manipulated in order to effect behavior. This problem was solved by the development of another concept that today we take for granted, the computer program. Again, the idea seems obvious, but did not come into existence until the 1930s in the work of Alan Turing (Hodge, 2000) and John von Neumann (MacRae, 1999). Previously, all machines, including the calculators built by Blaise Pascal, Gottfried Leibniz, and Charles Babbage, were dedicated, single-purpose machines whose mechanical workings defined the function they carried out. Computers, however, are general-purpose machines, capable of performing a variety of tasks. Their operations are determined not by their mechanical workings but by their programs, a series of instructions the computer carries out. Because they manipulate information, programs are independent of their physical substrate. A program written in BASIC (or any other computer language) will run on any computer that understands BASIC, whatever its physical makeup, whether it be an Apple, PC, or a mainframe. As Turing (1950) pointed out, a human being following a sequence of steps written on slips of paper is functionally equivalent to a computer. The computational approach to mind was complete and is known in philosophy as *functionalism*. The mind is essentially a computer program implemented in a meat-machine (Clark, 2001) rather than a silicon-and-metal machine. The program of the mind acts on and controls the flow of information through the human information-processing system the way a computer's program controls the flow of information through a computer. The program arrives at decisions and controls the system's—the body's—behavior. The mind is what the brain does (Pinker, 1998). Cognitive psychology becomes a form of reverse engineering. In reverse engineering, computer scientists take a chip and without opening it up, study its input-output functions and try to deduce what program controls the chip's processing. Often this is done to imitate an existing chip without violating the patent holder's rights. In psychology, experiments reveal the human mind's input-output functions, and psychological theories attempt to specify the computational functions that intervene between input and output.

The Fruits of Computation: Cognitive Science

Mind Design and the Architectures of Cognition Ironically, the first application of the computer conception of mind arose not in psychology but in computer science, when Alan Turing (1950) proposed that computer programs might emulate human intelligence. Turing put forward no new analysis of cognition but provided a now famous test by which computer intelligence might be recognized. A person interacts as in a chat room with two entities, one of which is a human being and the other of which is a computer program.

Turing said that the program would have to be called intelligent when the person could not tell if his or her conversational partner was human or computer. As yet, no program has passed the Turing test in the form Turing originally suggested.

Obviously, constructing artificial intelligences has great practical value. For cognitive psychology, the value of mind design (Haugeland, 1981, 1985) is that it forces theorists to think deeply and precisely about the requirements for intelligent cognition. In an inflfluential book, Marr (1982) specified three hierarchically arranged levels at which computational analysis takes place. In the case of artifificial intelligence, the levels define the job of making a mind, while in the case of psychology—which studies an already evolved intelligence— they define three levels of reverse-engineering psychological theory. The levels are most readily described from the standpoint of artifificial intelligence.

• The *cognitive level* specifies the task the AI system is to perform.

• The *algorithm level* specifies the computer programming that effects the task. The *implementation level* specifies how the hardware device is to carry out the program instructions.

The cognitive level is a detailed analysis of what a system must be able to know and do in order to perform a specified job. In certain respects, this is psychologically the most revealing level, because so much of what we know and do involves consciousness not at all. It is easy for me to walk downstairs and retrieve a book, and I can often do it while my conscious mind is engaged in thinking about writing this chapter. However, we find that building a robot to do the same thing reveals deep problems that my mind/brain solves effortlessly. Even recognizing an open doorway requires complexities of scene analysis that no robot can yet carry out.

Once one has specified the cognitive requirements of a task, the next job is writing the program that can get the job done. This is the algorithm level, defining the exact computational steps the system will perform. In psychology, this is the level of psychological theory, as we attempt to describe how our existing human program operates. An artificial system, on the other hand may achieve the same results with a very different program. For example, a human chess master and a chess-playing program such as Deep Blue solve the cognitive level problems of chess very differently. A computational psychological theory of chess playing needs to replicate the mental steps of the human player; the computational AI theory does not.

Finally, one implements the program in a working physical system. In AI, this means building or programming an intelligent system; in psychology it means working out the neuroscience about the workings of the human meat machine. Within Marr's broad framework, two different approaches to mind design—two architectures of cognition— came into existence, the symbol-system hypothesis and connectionism.

The Symbol-System Hypothesis

Herbert Simon and his colleague Allan Newell first drew the connection between human and computer cognition at the RAND Corporation in 1954 (Simon, 1996). Simon was by training an economist (he won the 1981 Nobel Prize in that field). As a graduate student, Simon had been greatly influenced by the writings of E. C. Tolman, and was well schooled in formal logic. Previously, computers had been seen as glorious, if flexible, number crunchers, calculators writ large. Simon saw that computers could be more fruitfully and generally viewed as symbol manipulators.

By the early twentieth century, logicians had established the concept of interpreted formal systems, in which propositions stated in language could be reduced to abstract formal statements and manipulated by formal rules. For example, the statement "If it snows, then school will be closed" could be represented by $p \supset q$, where p "it snows," q "school closes," and $\supset$ the logical relation if... then. If one now learns that it is snowing, one may validly infer that school will be closed. This inference may be represented as the formal argument *modus ponens:*

1. $p \supset q$

2. p

3. therefore, *q*

The significance of the translation into abstract, formal symbols is that we can see that it is possible to reason through a situation without knowledge of the content of the propositions. *Modus ponens* is a valid inference whether the topic is the connection between snow and school closings or whether a pair of gloves fits a murder suspect and the verdict ("If the gloves don't fit, you must acquit.") Mathematics is a formal system in which the variables have quantitative values; logic is a formal system in which the variables have semantic values.

In both systems, valid reasoning is possible without knowledge of the variables'value or meaning. Simon proposed, then, that human minds and computer programs are both *symbol systems* (Simon, 1980). Both receive informational input, represent the information internally as formal symbols, and manipulate them by logical rules to reach valid conclusions. Simon and Newell turned the notion into the pioneering computer simulation of thought, the General Problem Solver (Newell, Shaw, & Simon, 1958). Simon's symbol-system hypothesis established the fifirst of the two architectures of cognition inspired by the analogy between human being and computer, and it was firmly ensconced in psychology and artificial intelligence by the late 1970s. It gave rise to the creation of a new discipline, cognitive science, devoted to the study of *informavores,* creatures that consume information (Pylyshyn, 1984). It brought together cognitive psychologists, computer scientists, philosophers, and—especially in the 1990s, the decade of the brain— neuroscientists. (Space precludes a treatment of cognitive neuroscience. See Gazzinaga, Ivry, and Mangun [1998] for an excellent survey.)

The Connectionist, Subsymbolic, Hypothesis

From the dawn of the computer era, there had been two approaches to information processing by machines, serial processing and parallel processing. In a serial processing system, for example in home PCs and Apples, a single central processing unit (CPU) processes the steps of a program one at a time, albeit very quickly. The flow diagrams of information-processing psychology implicitly assumed that the human mind was a serial processor. Figure 6.3, for example, shows that multiple streams of input to sensory memory are reduced to a single stream by attention and pattern recognition. Likewise, the symbol-system hypothesis was predicated on a serial processing architecture, the human CPU executing one logical step at a time.

In parallel processing, multiple data streams are processed simultaneously by multiple processors. In the most interesting of these systems, distributed cognition systems (Rumelhart, McClelland, & PDR Research Group, 1986), there are large numbers of weak processors, in contrast to serial systems' single powerful processor.

Obviously, parallel-processing computers are potentially much more powerful than single CPU machines, but for a long time obstacles stood in the way of constructing them. Parallel machines are more physically complex than sequential machines, and they are vastly more difficult to program, since one must somehow coordinate the work of the multiple processors in order to avoid chaos. With regard to self programming machines, there is the special difficulty of figuring out how to get feedback information about the results of behavior to interior ("hidden") units lying between input and output units. Since sequential machines were great successes very early on, and the power of the parallel architecture seemed unnecessary, work on parallel-processing computers virtually ceased in the 1960s. In the 1980s, however, developments in both computer science and psychology converged to revive the fortunes of parallel-processing architectures. Although serial processors continued to gain speed, designers were pushing up against the limits of how fast electrons could move through silicon. At the same time, computer scientists were tackling jobs demanding ever-greater computing speed, making a change to parallel processing desirable. For example, consider the problem of computer vision, which must be solved if effective robots are to be built. Imagine a computer graphic made up of 256 256 pixels. For a serial computer to recognize such an image, it would have to compute one at a time the value of 256 256 65,536 pixels, which might take more time than allowed for a response to occur. On the other hand, a parallel-processing computer containing 256 256 interconnected processors can assign one to compute the value of a single pixel and so can process the graphic in a tiny fraction of a second.

In psychology, continued failings of the symbolic paradigm made parallel, connectionist processing an attractive alternative to serial symbol systems. Two issues were especially important for the new connectionists. First of all, traditional AI, while it had made advances on tasks humans find intellectually taxing, such as chess playing, was persistently unable to get machines to perform the sorts of tasks that people do without the least thought, such as recognizing patterns. Perhaps most importantly to psychologists, the behavior that they had most intensively studied for decades— learning—remained beyond the reach of programmed computers, and the development of parallel machines that could actually learn was quite exciting. That the brain could solve these problems while supercomputers could not suggested that the brain was not a serial machine.

The other shortcoming of symbolic AI that motivated the new connectionists was the plain fact that the brain is not a sequential computing device. If we regard neurons as small processors, then it becomes obvious that the brain is much more like a massively parallel processor than it is like a PC or an Apple. The brain contains thousands of interconnected neurons, all of which are working at the same time. As Rumelhart et al. (1986) announced, they aimed to replace the computer model in psychology with the brain model. The interconnected processors of connectionist models function like neurons: Each one is activated by input and then "fires," or produces output, depending on the summed strengths of its input. Assembled properly, such a network will learn to respond in stable ways to different inputs just as organisms do: Neural nets, as such processor assemblages are often called, learn. Connectionism suggested a new strategy for explaining cognition. The symbol-system approach depends, as we have seen, on the idea that intelligence consists in the manipulation of symbols by formal computational rules. Like the symbol-system approach, connectionism is computational, because connectionists try to write computer models that emulate human behavior. But connectionist systems use very different rules and representations (Dreyfus & Dreyfus, 1986; Smolensky, 1988): weighted mathematical connections between neuronlike units rather that logical manipulation of symbols that map on to propositions.

Connectionist systems differ critically from symbolic systems at Marr's implementation and algorithm levels. Analysis at the cognitive level is indifferent between the two architectures. However, at the implementation level, the nature of the hardware (or wetware, in the case of the brain) becomes crucial, because the implementation consists in executing a program with a real machine or real person, and different computers implement the same cognitive task in different ways. One of the two main issues that separate the symbol-system architecture of cognition from its connectionist rival concerns whether or not psychological theories of learning and cognition need be concerned with the implementation level. According to the symbol-system view, theimplementation of programs in a brain or a computer may be safely ignored at the cognitive and algorithm levels, while, according to the connectionist view, theorizing at higher levels must be constrained by the nature of the machine that will carry out the computations.

The second main issue concerns the algorithmic level of intelligence. William James (1890) first addressed the fundamental problem. James observed that when we first learn a skill, we must consciously think about what to do; as we become more experienced, consciousness deserts the task and we carry it out automatically, without conscious thought. One of the attractions of the symbolic paradigm is that it fits our conscious experience of thought: We think one thought at a time to the solution of a problem. The symbolic paradigm assumes that once a task becomes mastered and unconscious, we continue to think one thought at a time with consciousness subtracted. On the other hand, connectionism suggests that nonconscious thought may be very different from conscious thought. Smolensky (1988) analyzed the architecture of cognition from the perspective of how thoughtful processes become intuitive actions. Smolensky's framework distinguishes two levels, the conscious processor and the intuitive processor.

The conscious processor is engaged when we consciously think about a task or problem. However, as a skill becomes mastered, it moves into the intuitive processor; we just "do it" without conscious thought. Driving an automobile over a familiar route requires little if any conscious attention, which we turn over to listening to the radio or having a conversation with a passenger. Moreover, not everything the intuitive processor performs was once conscious. Many of the functions of the intuitive processor are innate, such as recognizing faces or simple patterns, while some abilities can be learned without ever becoming conscious, such as pure procedural learning in the absence of declarative learning, such as bicycle riding.

When it becomes automatic, driving or bicycling is performed by the intuitive processor, but what happens during the transition from conscious thought to intuition is a difficult issue to resolve. To see why, we must distinguish between *rule-following* and *rule-governed* behavior. Physical systems illustrate how rule-governed behavior need not be rule-following behavior. The earth revolves around the sun in an elliptical path governed by Newton's laws of motion and gravity. However, the earth does not follow these laws in the sense that it computes them and adjusts its course to comply with them. The computer guiding a spacecraft does follow Newton's laws, as they are written into its programs, but the motions of natural objects are governed by physical laws without following them by internal processing.

The following example suggests that the same distinction may apply to human behavior. Imagine seeing a cartoon drawing of an unfamiliar animal called a "wug." If I show you two of them, you will say, "There are two wugs." Shown two pictures of a creature called "wuk,'' you will say, "There are two wuks." In saying the plural, your behavior is governed by the rule of English morphology that to make a noun plural, you add an *-s*. Although you probably did not apply the rule consciously, it is not implausible to believe that you did as a child. However, your behavior was also governed by a rule of English phonology that an *-s* following a voiced consonant (e.g., /g/) is also voiced—wugz—while an *-s* following an unvoiced consonant (such as /k/) is also unvoiced—wuks. It is unlikely you ever consciously knew this rule at all.

Having developed the distinction between rule-governed and rule-following behaviors, we can state the algorithm-level distinction between the symbol-system and the connectionist architectures of cognition. All psychologists accept the idea that human behavior is rule governed, because if it were not, there could be no science of human behavior. The issue separating the symbol-system hypothesis from connectionism concerns whether and when human behavior is rule following.

According to the symbol system view, both the conscious processor and the intuitive processor are rule-following and rule-governed systems. When we think or decide consciously, we formulate rules and follow them in behaving. Intuitive thinking is likewise rule following. In the case of behaviors, that were once consciously followed, the procedures of the intuitive processor are the same as the procedures once followed in consciousness, but with awareness subtracted. In the case of intuitive behaviors, the process is truncated, with rules being formulated and followed directly by the intuitive processor. Connectionists hold that human behavior is rule following only at the conscious level. In the intuitive processor, radically different processes are taking place (Smolensky, 1988). Advocates of the symbol-system view are somewhat like Tolman, who believed that unconscious rats use cognitive maps as conscious lost humans do. Connectionists are like Hull, who believed that molar rule-governed behavior is at a lower level, the strengthening and weakening of input-output connections. After all, Thorndike called his theory connectionism 80 years ago. The intuitive processor lies between the conscious mind— the conscious processor—and the brain that implements human intelligence. According to the symbol-system account, the intuitive processor carries out step-by-step unconscious thinking that is essentially identical to the step-by-step conscious thinking of the conscious processor, and so Clark (1989) calls the symbol-system account the *mind's-eye view* of cognition. According to connectionism, the intuitive processor carries out nonsymbolic parallel processing similar to the neural parallel processing of the brain, and Clark calls it the *brain's-eye view* of cognition. Historically, connectionism represents more than simply a new technical approach to cognitive psychology. From the time of the ancient Greeks, Western philosophy assumed that having knowledge is knowing rules and that rational action consists in the following of rules. Human intuition has been deprecated as at best following rules unconsciously, and at worst as based on irrational impulse. Consistent with this view, psychology has been the search for the rule-governed springs of human behavior. But connectionism might vindicate human intuition as the secret of human success and rehabilitate a dissident tradition in philosophy—represented, for example, by Friedrich Nietzsche—that scorns being bound by rules as an inferior way of life (Dreyfus & Dreyfus, 1986). In addition, psychologists and philosophers are coming to believe that thought guided by emotion is wiser than pure logic (Damasio, 1994).

In the late 1980s, connectionism and the symbol-system view of learning and cognition acted as rivals, seemingly recreating the great theoretical battles of behaviorism's GoldenAge. However, around 1990 a *modus vivendi* reunified the field of cognitive science. The two architectures of cognition were reconciled by regarding the human mind as a hybrid of the two (Clark, 1989). At the neural level, learning and cognition must be carried out by connectionist-type processes, since the brain is a collection of simple but massively interconnected units. Yet as we have learned, physically different computational systems may implement the same programs. Therefore, it is possible that, although the brain is a massively parallel computer, the human mind in its rational aspects is a serial processor of representations, especially when thought is conscious. The more automatic and unconscious (intuitive) aspects of the human mind are connectionist in nature. Connectionist theories thus have a valuable role to play in being the vital interface between symbol-system models of rational, rule-following thought, and intuitive, nonlinear, nonsymbolic thought.

Cognitive Psychology Today

The computer metaphor of mind dominates the psychological study of cognition. There are more computational models of information processes than can be briefly summarized. However, four large problems remain outstanding.

• *Consciousness*. The stubborn fact of consciousness remains, and the computer model of mind has been of little help, because computers are not conscious (though see Dennett, 1991). Why are we conscious? Does consciousness play any causal role in our mental economy or behavior? Little real progress has been made since behaviorist days.

• *Meaning*. How do physical symbols get their meaning; why does GIFT mean a present in English but poison in German? Ebbinghaus and S-R behaviorists avoided the question. Mediational behaviorists said meaning was carried by covert r-s connections, and Skinner offered an explanation in terms of tacting. The symbol system hypothesis finesses the issue by saying thinking is governed by formal logical rules (syntax), not meaning (semantics). Connectionism, like S-R psychology, tries to dissolve meanings into nonmeaningful units of response. The problem has not been solved.

• *Development*. Why and how do children throughout the world grow up with similar, if not identical, cognitive processes and a store of common beliefs, despite differences in environment?

• *Evolution.* Given that the human mind was constructed by evolution, are there important limits on human cognition, and certain thoughts it's easy to think while there may be others that are difficult or impossible to think? Space prevents full discussion of these issues, and solving them lies in the future. See Clark (2001), Leahey (2000, 2001), and Leahey and Harris (2001) for more.