The Philosophical Approach: Enduring Questions

"What is Matter?—Never mind."
"What is Mind?—No matter."

—Anonymous, 1855

What Is Philosophy?

Philosophy in its broadest sense is the search for wisdom and knowledge. It is the first approach we will tackle in our voyage through the different disciplines of cognitive science. There are good reasons for beginning here. Philosophy plays a vital participatory role in cognitive science. It does this not by generating results, since it is a theoretical rather than experimental discipline, but by "defining problems, criticizing models, and suggesting avenues for future research" (Garfield, 1995, p. 374). More than any other discipline in cognitive science, philosophy is not limited by its subject matter or a particular theoretical stance. It is therefore free to evaluate and contribute to the remaining disciplines in a way the others cannot. This approach is also the oldest of the different approaches, tracing its origins back to the ancient Greeks. It is thus fitting that we begin our tour here.
The translation of the word philosophy yields “love of wisdom,” indicating the philosopher’s concern with knowledge and with understanding the universe. Philosophy as a formal discipline studies a wide range of topics. In fact, there is no topic that is not fair game for a philosopher; he or she may examine politics, ethics, esthetics, and other subjects. We concern ourselves here with two branches of philosophy. Metaphysics examines the nature of reality. The mind-body problem is a metaphysical one at heart, because it seeks to understand whether the mental world is part of the physical world. Epistemology is the study of knowledge and asks such questions as: What is knowledge? How is knowledge represented in the mind? How do we come to acquire knowledge?

In this chapter we will survey philosophic thoughts that center on four vexing issues, most of which are summed up in terms of “this” versus “that.” This terminology suggests that the debates that have arisen from these issues have polarized the arguments and that there are only two possible answers to a problem. We will see that this is actually not the case and that there are multiple ways to conceptualize the issues. These issues are the mind-body, free will-determinism, and nature-nurture debates. In addition, we discuss the question of consciousness and its relation to cognitive science.

The Mind-Body Problem

The mind-body problem addresses how psychological or mental properties are related to physical properties. The debate stems from a fundamental conception about what the mind is. On the one hand we have the brain that is material and physical. It is made up of substances that we can measure and understand. The mind could be thought of in the same way, as simply a physical thing. On the other hand, there are those who argue that the mind is something more. They say we can’t equate our subjective conscious experiences, such as beliefs, desires, and thoughts, with something as mundane as the brain. They say the mind is nonphysical and consists of something resembling a soul or spirit. The mind as a nonphysical entity inhabiting the brain or other physical entity is sometimes called “the ghost in the machine.”

The first question of the mind-body problem refers to the nature of what mind is. Is the mind physical or something else? A second and more specific question concerns the relationship between these two entities. If we assume that there are two such entities, then what is the causal relationship between them? Does the mind control the mind or does the body control the mind? Table 2.1 shows the possible relationships between mind and body and the labels that go with each.
### Table 2.1
Different interpretations of the mind-body debate and the schools of thought associated with each

<table>
<thead>
<tr>
<th>Class of Theory</th>
<th>Name of Theory</th>
<th>Physical Universe</th>
<th>Causal Direction</th>
<th>Mental Universe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monism</td>
<td>Idealism/Solipsism</td>
<td>None</td>
<td>No causality</td>
<td>Mind</td>
</tr>
<tr>
<td></td>
<td>Physicalism</td>
<td>Body</td>
<td>No causality</td>
<td>None</td>
</tr>
<tr>
<td>Dualism</td>
<td>Classical Dualism</td>
<td>Body</td>
<td></td>
<td>Mind</td>
</tr>
<tr>
<td></td>
<td>Parallelism</td>
<td>Body</td>
<td>No causality</td>
<td>Mind</td>
</tr>
<tr>
<td></td>
<td>Epiphenomenalism</td>
<td>Body</td>
<td></td>
<td>Mind</td>
</tr>
<tr>
<td></td>
<td>Interactionism</td>
<td>Body</td>
<td></td>
<td>Mind</td>
</tr>
</tbody>
</table>


Our discussion in this section will be structured around basic conceptions of the nature of mind. According to monism, there is only one kind of state or substance in the universe. The ancient Greek philosopher Aristotle (384–322 B.C.E.) was a monist (Figure 2.1). He characterized the difference between mind and body as the difference between form and matter. One way to think of his notion is to consider a lump of clay. It is made up of physical matter and we can think of it as corresponding to the brain. We can shape the clay with our hands into different forms—for example, we can roll it into a ball or flatten it out into a pancake. The shapes the clay can assume, Aristotle implied, are like the different thoughts the mind can take on when it undergoes different patterns of activity. These shapes are just different physical states and do not constitute any nonphysical or spiritual substance. There are two classes of monist. Monists who believe only in mental substance are either idealists or solipsists. Monists who believe exclusively in physical substance, like Aristotle, are physicalists.

In dualism, one believes that both mental and physical substances are possible. Plato, another Greek philosopher (427–347 B.C.E.), was a dualist. Plato was Aristotle's teacher, but the two held quite different views. Plato believed that the mind and the body exist in two separate worlds. Knowledge of the mind, he thought, exists in an ideal world of forms, which is immaterial, non-extended, and eternal. The body resides in a world that is material, extended,
and perishable. There are crucial differences between the objects of one world versus those of the other. Mental ideas such as “circle” that reside in the ideal world of forms are perfect, according to Plato: the circles of this world are always perfectly round. Concrete examples of circles that we find in the real world are always imperfect. If we examine an actual circle, at some level of magnification the circle’s edge will lose its roundness. There are a variety of different modern dualist schools, each based on a particular view of the relationship between the mental and the physical. These include classic dualism, parallelism, epiphenomenalism, interactionism, and functionalism.
Flavors of Monism

Idealism has been attributed to the Irish philosopher George Berkeley (1685–1783). In this philosophy, there is only the mental realm. There is no physical realm. Our brains, bodies, and everything else in the universe exist only as concepts in God’s mind. All our experiences are the same as God’s experiences and we have them because God has them. Problems with the idealist position should become immediately clear. It resorts to a religious and mystical form of explanation and must be accepted on faith. Although it is logically consistent, this view cannot be tested and so is not considered scientific.

Solipsism also falls into the category of “mental only” theories of mind. According to this perspective, the universe exists only in one’s mind. In other words, each person, by virtue of having a mind, creates his or her own world. This brings up a host of questions. If the universe exists only in each individual mind, then there must be as many universes as there are individuals. If this were the case, then which universe is the right one? Do they all exist at the same time? This view is inherently subjective and, like idealism, is not subject to experimental scrutiny.

On the flip side of our metaphysical coin is physicalism or materialism. The origins of this view go back to the Greek philosopher Democritus (ca. 460–370 B.C.), who believed all things were composed of atoms. The attributes and behaviors of the atoms, he said, can explain the differences between things, including the differences between mind and body. Physicists, like idealists and solipsists, are also monistic, and believe that the universe is composed of a single substance. They however regard this substance as physical and material rather than spiritual or ethereal. Physicalism is thus the doctrine that everything that exists is physical. The operations of the mind are seen here as simply the operations of the brain.

A school of physicalism known as reductive physicalism is an example of reductionism, where one theory or view is used to completely account for another. As an example, it is often believed that neuroscience, which is concerned with the study of the brain’s physical make up and processes, will ultimately be able to account for all psychological or mentalistic levels of description. To illustrate, a psychologist’s explanation of an anxiety disorder that uses a mentalistic term such as “fear” may someday—this class of physicists hopes—be reduced to a description of neurotransmitter changes in brain structures, such as the amygdala. A second breed of physicalists go by the name of nonreductive physicalists. They believe physical processes can give rise to emergent and irreducible mental phenomena. This school believes that one cannot completely do away with mentalistic description. We will discuss the property of emergence in the context of consciousness.
Physicalism has received its share of critique as well. Some allow that physical processes can determine mental ones but deny that they can explain them. So, they argue, changes in the amygdala may very well correlate with and produce fear, but they do not explain different kinds of fear, how a person becomes fearful, and so on. These critics acknowledge the world is physical but indicate that for many phenomena there is no physical explanation. In these cases, they believe, it is perhaps better to explain using mental terms.

**Flavors of Dualism**

Now that we have reviewed the various forms of monism, let us turn our attention to its logical alternative. Dualists believe both mental and physical realms are possible, but differ in the way they think these two interact. Classical dualism originated with the French philosopher René Descartes (1596–1650). Descartes was a revolutionary philosopher for his time and introduced theories on many of the ideas that underlie cognitive science. He believed in a one-way causal link, with the mind controlling the body but not vice versa. Descartes thought the mind exerted its control on the body through the pineal gland, perhaps because it is one of the few anatomical structures not duplicated on either side of the brain (see Figure 2.2). In this view, the mind is like a puppet master, the body is like a puppet, and the pineal gland is like the puppet strings, by which the former controls the latter. Classical dualism conforms to most people’s common-sense notion of the mind-brain relationship, which is that our thoughts control our actions. For instance, when we feel hungry, we get up and eat a snack. It certainly seems as if the feeling of hunger comes first and causes the action of getting up to eat.

There are three other dualist schools of thought. In parallelism, the mind and body are distinct and isolated from each other. One cannot have an effect on the other. But there certainly appears to be some sort of causal link between them. Otherwise, how is it that we can move our legs when we decide to walk? A parallelist response would be that God or some other unknown force has synchronized the workings of the mind and body so that they work seamlessly together. It is like having two clocks, a physical and a mental one, that were both started at the same time and run alongside one another, keeping the same time. The critique here is the same as for idealism. This view resorts to mystical explanation and so is not accepted within mainstream cognitive science.

The epiphenomenalist school allows the physical to cause the mental, but prohibits causation in the other direction. In this account, the mind is like a side effect of the brain’s operations, but cannot in any way feed backward to
influence the brain. As an analogy, imagine the exhaust coming out of a car engine. The engine is the brain, while the exhaust it creates would be the mind. The engine produces the exhaust, but the exhaust in no way affects the operation of the engine or car. A difficulty with this view is that it runs counter to our introspections, many of which place thoughts before actions.

**Interactionism** allows causality to travel both ways. The body can affect the mind and the mind can also affect the body. Here we have a two-way street that allows each realm to influence the other. One can have a thought that produces an action, but also a physical brain activity that produces a thought. The approach has its problems: how these two mutually influence one another as an integrated whole is not specified. Of particular concern is how the mental can affect the physical. It is easier to conceptualize how physical causes have

---

**Figure 2.2** Descartes believed the pineal gland to be the location where the mind influenced the body. This belief was itself influenced by the fact that the pineal gland is located in the center of the brain.
effects. We already have a good account of this in Newtonian physics. Immaterial causation is more of a mystery.

Evaluating the Dualist Perspective

One critique of dualism comes from the philosopher Gilbert Ryle. Ryle's argument (1949) centers on our conceptualization of mind and its relation to the body. He believes that the mind is not any particular component of the brain, but all the parts working together as a coordinated, organized whole. He illustrates with a story. Imagine a visitor from a foreign country arriving at a large university. He is shown around the campus and the various parts of the school are pointed out to him, including the dormitories, departments, and lawns. The visitor, who has never seen any of this before, is puzzled. He says; "Well, I've seen all of this, but I haven't yet seen the university." We would have to explain to him that the university is not any of the individual sites he has viewed, but all the sites together and the interconnections among them (see Figure 2.3). Ryle thinks philosophers fall into the same trap as the visitor, mistaking the part or parts for the whole. He argues that the mind belongs in a conceptual category different from that of the body, just as the university is in a category different from those of the things that make it up.

Andy Clark (2001) summarizes several other critiques of dualism. These would apply to Descartes' conception as well as other perspectives. Clark says that dualism is uninformative and tells us what the mind isn't, rather than what it is. If the mind isn't the brain and isn't physical, then what is it? Dualists are remarkably silent on this matter, often conceding that it is something nonphysical that we can't understand yet. As a theory, dualism is also inelegant, because it postulates two worlds that must be coordinated. An explanation that does not violate the principle of Occam's razor (that the simpler explanation is usually the correct one) would involve a single type of world, not requiring coordination.

There are further problems with dualism. One has to do with the dependence of the mental on the physical. Factors that affect the brain such as head trauma or drug use have direct and dramatic mental effects. We can see that damage to a certain part of the brain, say, from a motorcycle accident, results in specific forms of mental disruption, for example, language deficits. Taking a drug like marijuana, which alters brain chemistry, results in altered mental states. In addition, the evolutionary approach shows us there is a general positive correlation between brain size and intelligence across species, with larger
brain sizes linked to increased cognitive capacity. It is obvious from these observations that the mental is integrated with the physical, that the mind depends on the brain.

Some dualists in response to attacks on their positions have stated that the mind exhibits extraordinary abilities and that it is or will be impossible for a physical system to duplicate such abilities. For instance, how can a physical system, be it a brain or a computer, write a novel or negotiate a peace treaty? The truth is that as our technological sophistication increases, many of these abilities are becoming better understood and implemented computationally. There are now computers that can beat the best chess champions and successfully diagnose medical disorders. These are capacities once thought to be the exclusive domain of humans.

Dualists and other philosophers also argue that our subjective experiences—things like thoughts, beliefs, and desires—are not equivalent to physical brain
states. They base this conclusion primarily on introspection. When we examine what is inside our heads, they say, these subjective experiences seem to be something more than just physical. The problem with this argument is that introspection is a weak form of evidence and can be wrong (as are many of our ideas). What is required is objective proof that such experiential states are not physical.

**Functionalism**

The most influential philosophical theory of mind in cognitive science is functionalism. For this reason we will discuss it in considerably more detail than any of the theories we’ve already discussed. To get an idea of what functionalism is about, we need to make a distinction between two ways of classifying things. **Physical kinds** are identified by their material composition only. In this view, jellyfish and carpets are different because they are made up of fundamentally different physical substances. **Functional kinds** however are distinguished by their actions or tendencies. Here, we could say that all automobiles fall under the same functional category because they do the same things, namely, transport goods and people, even though they may be made up of different elements.

So far so good, but things get interesting when we extend these ways of classifying to the idea of mind. If we think of mind as a physical kind, then minds must be the same things as brains, since, as far as we know, minds cannot exist apart from physical brains. To many, this seems too exclusive. It is possible, they argue, that computers might develop minds and that there might be alien species with minds (see Figure 2.4). Neither computers nor aliens need have brains in the sense that we know them. It is more fruitful, they say, to identify minds as functional kinds and to define them by the sorts of processes they carry out rather than the stuff they’re made of. According to functionalism, mental states are not just physical states, but also the functioning or operation of those physical states. According to this view, a mind could conceivably be implemented in any physical system, artificial or natural, capable of supporting the appropriate computation.

Functionalism has several significant implications (Garfield, 1995). One is that the same mental state could be realized in a quite different way in two separate physical systems. This can be illustrated with computing devices. Two such different devices, say a desktop computer and a palm-sized personal data assistant, can both compute the same result, such as displaying a page of text, but in entirely different ways. The same might also be true for human
computation. If we examined the brains of two people thinking exactly the same thought, we would in all likelihood not find exactly the same processes at work.

There are several schools of thought in functionalism. These range from conservative views that advocate direct connections between physical and computational states, to more liberal ones that emphasize computation over physicality. The liberal schools give two reasons for their stance. They say that for both computers and thinking organisms, the number of possible computational states always exceeds the number of possible physical states. Take for example all the different possible beliefs one could hold concerning politics, the environment, one's friends, and so on. Mathematically, the number of such beliefs is infinite (Garfield, 1995). The number of possible physical states the brain can assume, though, is finite. A computational level of description thus becomes a richer and more diverse way of describing the mind and should be the preferred level. Second, liberal functionalists argue that psychological states such as beliefs are defined more by their relations to other such states, to inputs from the environment, and to behaviors than their relations to physical states. A belief such as “patriotism” usually manifests itself in other beliefs, for
example, in flag-waving. It will elicit predictable reactions to environmental stimuli, for example, feeling upset when one's country has been criticized, and will produce external behaviors such as marching or protesting.

To summarize, functionalism implies that mental states might not be reduced to any particular physical state. This argument does not require us to be dualists. It is not saying that mental states don't conform to physical ones, only that there may be a wide variety of possible physical states capable of producing any given mental state.

**Evaluating the Functionalist Perspective**

Although functionalism has been the dominant view in cognitive science since the 1970s, it is not without its deficiencies (Maloney, 1999). Remember that a tenet of functionalism is that minds that are not based on brains can exist. They can exist in things like computers as long as the physical substrates of those objects allow for the relevant computations. Critics have argued that, although it is possible that minds can exist in the absence of brains, this does not make it plausible. There is no current empirical evidence to justify this claim. We have yet to see something mental in the absence of a brain. Also, some have argued that the failure to identify mind with a physical kind can itself be considered reason to do away with the concept of mind—rather than give it special status as a functional kind.

An additional problem with functionalism is that it cannot account for the felt or experienced character of mental states—a phenomenon known as qualia (quale, singular). Examples of qualia include the subjective experience of what it is like to feel "hungry," to be "angry," or to see the color "red." It would seem that these kinds of experiences cannot be replicated as purely functional processes. A machine could be programmed to "see" the color red, even mimicking the same human functional process, but this machine could not have the same experience of what it is like to see red that a person has.

What is more, two individuals having the same conscious experience often do not experience it subjectively in the same way. A number of experiments have shown this to be the case with color perception. Participants looking at the same color will describe it differently (Chapanis, 1965). If asked to point out on a color spectrum what pure green looks like, one person may select a yellow-green, another a blue-green. This is the case even though the functional operations of their respective brains as they view the color are approximately equivalent. In this case, the neurophysiological operations behind color perception tend to be the same across individuals.
The Free Will–Determinism Debate

The Issue of Determinism

Think for a minute about all the decisions you’ve made in your life. Some of these were important, for example, deciding which college or university to attend, or perhaps deciding whether or not to pursue a romantic relationship. Others may have been less essential, for example, whether to get pork with fried rice or Szechuan chicken for lunch. The free will–determinism debate is about whether these behaviors are within our control. Did you consciously ponder the consequences of attending one school versus another, thinking of the possible pros and cons of each, or did forces beyond your control push you to attend one school over the others? Those who take the free will side of this debate argue that individuals independently initiate their own actions. Those favoring determinism argue that actions can be explained in terms of the initiating causes that precede them, which implies that individuals act only dependently, as a consequence of these causes.

Let us start with determinism. It is the view that all physical events are caused or determined by the sum total of all prior events. Our actions, which are physical events, must therefore also be determined. Suppose you get up to grab a snack from the refrigerator. According to determinism, your getting up was caused by another physical event (or events) that immediately preceded it. This event may have been a stomach pang or hearing a restaurant advertisement on the radio. Furthermore, the determinist view is that it was inevitable that you would have undertaken this action because of the preceding event. In other words, there was no other action you could have taken, given the nature of who you are and the set of events that preceded the act of getting up.

The philosopher David Hume (1711–1776) suggests we conceive of a deterministic universe in terms of billiard balls. To get a sense of this notion, imagine that, initially, a set of billiard balls is scattered across a table and the balls occupy random positions. We then come along with a cue stick and knock one ball into another. Each moving ball, in this Humean model of causation, is an event that was caused by a preceding event and in turn causes another (see Figure 2.5).

This model has several implications. The first, as mentioned before, is determinism. If ball A strikes ball B, then there is only one “action” that ball B can take, that action being entirely caused by ball A’s action. Of course, this is a simplified model. It is possible to have multiple forces acting on ball B, in which case their conjoint influence determines ball B’s action. The point is that ball B’s action is determined completely by the forces acting on it in the
Figure 2.5  In the Humean model of causation, action in the physical universe can be likened to billiard balls on a billiard ball table, where balls bump into one another. In this instance the cue ball initiates two causal paths, one involving balls having a stripe, the other solid-colored balls.

temporally preceding moment. A second implication is replication. If we were to return all of the balls to their original positions and strike ball A with the cue stick again, in exactly the same way, then all of the balls would move once more in exactly the same way. The third implication is prediction. If we know the positions of all of the balls and the conditions under which the first ball will be hit, such as the force to be applied and angle of the cue stick, then we can know ahead of time what all of the balls on the table will do. Many of these properties of a causal system also happen to underlie scientists’ understanding of the way that variables in a controlled scientific study interact. The billiard balls’ actions on each other describe the way experimental variables are
hypothesized to influence each other. Indeed, replicability and predictability are two of the cornerstones of the scientific method. We will talk more about the scientific method in the psychology chapter.

If we translate this deterministic model of causality to human behavior, it means we can also predict any action a person might undertake for the rest of his or her life. All that is needed is an understanding of the system and the forces acting on it. In this case the system is the brain. The forces that act on the brain can include influences such as a stomach pang or radio advertisement. The behavior of the system would be the overt behavior the brain gives rise to, such as getting up to get a snack. This notion is not too far removed from what the behaviorist psychologists in the early 20th century were proposing. In their case they didn’t purport to understand the mind or how it operated; that is, they ignored the system itself. They instead concentrated on understanding the causal relationship between stimulus inputs from the environment and the consequent behavioral output. Some behaviorists felt they could entirely predict a person’s actions on the basis of that person’s conditioning history, which is the schedule of rewards and punishments he or she has undergone so far in his or her life. We will talk more about behaviorism in the following chapter.

The Issue of Free Will

Most people find the Humean billiard ball model of causality unsettling. This is because it turns us into automatons reacting in known ways to the forces impinging on us. We would like to believe that we instead choose our own course of action. The free will perspective regards behavior as stemming from a decision or act of will. The decision is made autonomously and is itself not under the influence of any preceding causal factors. In other words, the will of the individual is the sole determining cause of the action. The will itself is not the product of any other cause, and is considered to be its own cause. People, according to this view, are thus not at the whim of forces beyond their control, but have the power to independently initiate their own actions.

The 20th century philosopher Ayn Rand (1963) has formulated an entity model of causation that, according to Rand, underlies free will. In this model, entities with specific identities are the cause of actions. An entity is defined roughly as an object capable of independent action. A person is an example of an entity. The actions an entity undertakes are determined not by some antecedent factor that acts on the entity, but instead by the nature of that entity. If an entity is a certain way, then it will act only in accordance with that way.
In fact, it is not possible for an entity to act in a way that is contradictory to its nature. By making entities rather than actions the cause of actions, Rand shifts the force of causality away from the environment and toward individuals.

Rand further argues that humans are beings of volitional consciousness. By this she means that although people are capable of thinking, they must make the decision to do so. By this account, both thought and the decision to think or not are part of human nature. If we decide to think, then we consciously control our actions. If we fail to think, then we are at the whim of our mind’s subconscious associational processes. This notion of voluntarily directing one’s thought processes Rand calls focus. Cognitive psychologists refer to it more broadly as attention. We discuss several models of attention in the cognitive approach chapter.

**Evaluating the Free Will–Determinism Debate**

A problem with free will is that it violates one key assumption of causality, which is that all events must have a cause. In this assumption, actions can’t cause themselves. The causal universe is seen here as a network of dependencies, with actions existing only as possibilities until they are actualized by a triggering event or events. If this is true, then how can a person initiate an action “out of nowhere”? The strict version of free will implies that the decision or will of an individual to act is a cause that is itself uncaused. This cannot be accounted for in the scientific and causal view of the universe.

According to the cognitive science view, a decision is itself a mental process. Like other such processes, it should have informational inputs, perform a computation, and have corresponding informational or behavioral outputs. The decision to decide which school to attend must have as a starting point a mental list of possible schools along with data about each. The information, in this view, is then passed through some sort of decision-making algorithm. This algorithm might calculate a winning school by adding up scores pertaining to certain features such as prestige and geographic location. The output of the computation would be the single school to which one would then apply.

Using the above model, the decision is not uncaused, at least in the sense of having no preceding events. These events are the list of schools and farther back in time the decision to go to school in the first place. From a scientific and causal perspective, if the decision-making algorithm and the data fed to it are completely known, then the decision in principle should follow deterministic consequences, that is, the decision should be determined, replicable, and predictable. However, empirical evidence in the cognitive sciences rarely produces models that adhere perfectly to all three of these criteria. Part of the reason is
incomplete knowledge. If the algorithm or data is only partially understood, then the decision can be only partially determined, replicated, or predicted.

One way to reconcile the free will–determinism dilemma is to allow both to be true. This is what compatibilism offers. Compatibilists allow that free will and determinism can be reconciled and/or compatible with each other. It allows these two schools of thought to coexist. Compatibilists believe we have free will in the broad sense of moral responsibility, meaning that even though we may be constrained to act in a certain way, we always have the freedom to choose otherwise. For instance, imagine a thief pointing a gun at your head and asking for your wallet. You are certainly constrained under these circumstances to fork over your cash. But in a compatibilist world you don’t have to do so, you are perfectly free to refuse and accept the consequences. Just as compatibilists do not deny causality, they believe every event has a cause and that our actions are always preceded by a cause. However, these antecedent events do not fix our actions. In any past circumstance, we would have been free to embark on some other alternate course of action.

In contrast to this is the theory of incompatibilism. Incompatibilists see free will and determinism as irreconcilable, meaning they are not both simultaneously true. This school argues that a person cannot be truly free in the moral sense if preceding causal events impact on his or her actions. Causal laws are seen here as preventing us from being true “free agents,” from having absolute moral responsibility for our actions. One group of incompatibilists, called libertarians, believes we do have free will, that free will is not compatible with determinism, and that determinism must therefore be false. The problem with this view is that its adherents must show how we can be morally responsible in an indeterminist world (one where determinism is false). A second school, whose adherents are called pessimists or no-freedom theorists, argue that moral free will is impossible to prove. The pessimists say there are circumstances in which we can be free agents—but those circumstances occur only when we are unconstrained, when circumstances allow us to choose to do what we want. This however is not enough to explain free will in the larger sense, which is that we always have moral responsibility for our actions, regardless of the constraining factors.

The Knowledge Acquisition Problem

A fundamental question asked by even the earliest of philosophers was: How do we acquire knowledge? Clearly, you are not born knowing everything, otherwise you would not need to go to school, or wouldn’t be reading this book.
But are we born knowing anything at all? Is the mind completely blank or do we start with some rudimentary understanding of the world? One way to frame these questions is within the nature-nurture debate. This debate centers on the relative contributions of biology and experience in determining any particular capacity. The term nature, in this context, refers to traits that are genetically or biologically determined. These are coded for in our genes and so are “hardwired,” meaning they are present at birth or appear at a specific time during development. The term nurture refers to traits that are learned through experience and interaction with the environment. We will examine theories of knowledge acquisition that argue for the greater influence of one or the other.

According to nativism, a significant body of knowledge is innate or “built into” an organism. In this sense, nativism is a theory of knowledge that favors nature over nurture. Plato was the first to outline a nativist theory of knowledge. He thought learning was a matter of recollecting what is already known—these concepts existing in the ideal world of forms and being part of our immortal soul. Rationalism must be subtly distinguished from nativism. Descartes was the progenitor of this perspective. Rationalists also believe in the existence of innate ideas. These basic concepts include such ideas as “God” and “triangle.” However, they additionally emphasize the existence of innate reasoning powers. These include certain logical propositions, such as knowing that something cannot exist and not exist at the same time. We can use these a priori rational powers to form new ideas that are not given to us innately. Descartes would agree that we are not born with the idea of “table,” but can acquire it given our innate ability to perceive and think about objects.

Empiricism alternatively sees knowledge as acquired through experience: it favors nurture over nature. In this view, knowledge gets into the head through interaction with an environment, meaning it is learned. The senses provide the primary channels via which knowledge of the world is born. Our knowledge of the concept “lemon” in this account begins with looking at a lemon, touching and tasting it. The British philosopher John Locke (1632–1704) is credited as the founder of the empiricist movement. He used the phrase *tabula rasa*, which literally translates as “blank slate.” Locke believed that we are born as blank slates, lacking any knowledge, and that over time experience puts writing onto the slate, filling it up.

Locke had a more fully developed theory of learning. He differentiates between simple ideas and complex ideas. Simple ideas are derived through sensory input or simple processes of reflection. They are received passively by the mind and cannot be reduced to simpler ideas. Looking at a cherry would generate the simple idea of “red.” Tasting a cherry would produce the simple idea of “sweet.” Complex ideas are formed from the active mental combination of simple ideas. They are created through reflection only and can be reduced to
parts, their component simple ideas. The idea of “cherry” would result from
the associative combination of such simple ideas as “red,” “sweet,” and other
commonly occurring sensations derived from one’s experiencing cherries. This
cluster of simple ideas is naturally associated because each time we experience
a cherry it comes to mind. For this reason, Locke and others who have pro-
posed similar notions are sometimes known as the associationists.

Evaluating the Knowledge Acquisition Debate

One might be tempted to immediately dismiss the doctrine of innate ideas
put forth by the nativists and rationalists. After all, it seems absurd that
we should be born knowing factual information such as the content of the
Gettysburg address. But the scope of knowledge is broader than this. Think
back to the previous chapter, in which we defined declarative knowledge for
facts and procedural knowledge for skills. There is quite a bit of research
supporting the notion that some forms of procedural knowledge are innate.
Newborn infants for instance come into this world with a variety of dif-
ferent skills. These skills are universal across the human species and manifest
themselves so soon after birth that they couldn’t possibly have been learned.
They therefore qualify as examples of innate knowledge. Let us examine a
few of these.

All infants demonstrate a set of reflexes. These reflexes include the grasping
reflex, in which the fingers tighten around a touch to the palm, and the root-
ing reflex, in which the infant turns his or her head and begins sucking
an object placed near the mouth. Reflexes serve a clear adaptive function.
Grasping and sucking, along with behaviors generated by other early reflexes,
are important for survival. The physiology behind reflexes is simple and fairly
well understood. A stimulus triggers one or more sensory neurons that then
activate intermediary neurons. These in tum activate motor neurons, causing
the resulting behavior. It is easy to see how such a simple mechanism could be
hardwired at birth to enable the infant to respond effectively to its environ-
ment. Figure 2.6 shows the anatomy of a spinal reflex.

Smell preference is another example of innate behavior. Steiner (1979)
found newborns tend to agree with adults in terms of which odors they con-
sider pleasant or unpleasant. He found that odors such as strawberry and
banana elicited agreeable facial expressions from young infants, for example,
smiling. Unpleasant odors such as fish and rotten eggs elicited expressions of
disgust. As is the case with reflexes, these smell preferences have survival value.
Babies who find the smell of fruit attractive will eat the fruit and thereby gain
nutrients, those that are repulsed by spoiled or unpalatable food will reject the
food and avoid getting sick. The neural mechanisms behind such preferences are probably not too complex either. They need involve little more than a mapping between the odor and the emotional response.

Given the above examples, we see that it is not so far-fetched for us to be born with procedural knowledge. This knowledge is in the form of simple neural circuits that map stimulus inputs to appropriate behavioral outputs. This knowledge can even be represented using the conditional rules we talked about in Chapter 1. A coding of a smell preference might look something like: “If smell is fish then disgust.” The odor, if it satisfies the first part of the conditional, would then trigger the response in the second part.

But how did these circuits get there in the first place? The early nativists and rationalists either did not specify the source of innate knowledge or attributed it to God. Evolutionary psychology offers us another explanation. It attributes such capacities to generations of selection pressures acting on a species. These pressures promote the development of adaptive (survival-related) cognitive abilities. Evolutionary psychologists further argue that these innate abilities are domain-specific, meaning they are attuned to perform special operations only on a certain type of information. Domain-specific mechanisms can be distinguished from content-neutral or domain-general learning mechanisms, such as the processes of associationism proposed by Locke. Evolutionary psychologists can be considered modern-day nativists. See the evolutionary approach chapter for more on their views.
The phrasing of the nature-nurture debate as allowing for only a single alternative, either one or the other, is misleading. Although some traits may indeed be the product entirely of nature or of nurture, there is a large middle ground consisting of traits or cognitive abilities that can result from the complex interaction of the two. In these cases, nature may set constraints or limits on environmental influence. Take memory for example. Tsien et al. (1996) engineered a mutation in a gene that affects a particular type of receptor in the hippocampus, a brain area responsible for the learning of new information. Rats with the mutation did poorly in a memory task as compared to normal rats in a control group. Tang et al. (1999) did something even more remarkable. Through genetic manipulation they increased production of a particular subunit in the hippocampal receptor. This change increased the receptor’s effectiveness. Rats with this “enhanced” version of the gene outperformed rats with normal receptors on a spatial memory test.

This research is exciting because it shows that memory in these animals is at least partially under genetic control. However, it is also well documented that human memory capability can be improved through organization and the use of memory strategies (Roediger, 1980). The way in which these genetic and environmental factors interact to determine memory in any given individual is complex. It could be that any amount of organization could produce no more than a small increase in memory performance if the aforementioned gene were lacking. Alternatively, those with the enhanced version of the gene who also employ memory strategies could perhaps acquire “supermemories,” and then no longer need to employ memory strategies in order to remember effectively.

---

**The Mystery of Consciousness**

Consciousness is a complex concept and has no single agreed-upon definition. In its broadest sense, we can think of it as the subjective quality of experience (Chalmers, 1996). It may be thought of as our individual subjective awareness of mental states. These states include sensation, perception, visual images, conscious thought processes, emotions, and sense of self, just to name a few. But these states assume a person is in a normal, awake, and alert frame of mind. The issue becomes more complex when we think of other types of consciousness; for example, being unconscious, asleep, in a drug-induced state, hypnotized, or meditating. There are clinical cases representing other states of consciousness as well. In dissociative identity disorder, a person can alternate between separate personalities. Each personality can possess unique skills and may or may not be aware of the others. In split-brain patients, one half of the
brain can possess an awareness of an object the other half does not possess. For simplicity, we do not consider these alternate states of mind.

An interesting aspect of consciousness is whether it is unitary or divided. Subjectively, our consciousness seems to be unitary. That is, one recognizes himself or herself to be one person, experiencing things in the present moment. When one studies the brain, though, one finds that there is no single place or even time where consciousness seems to happen. Instead, the brain in action is a case of activity going on all over the place. Furthermore, the brain may even be processing different aspects of a single experience at different times. How can we reconcile this objective evidence with our subjective experience? See the In Depth section for one theory on this apparent contradiction.

Chalmers (1996) makes a distinction between phenomenal and psychological concepts of mind. The **phenomenal concept of mind** is essentially the idea of mind as a conscious experience. Mental states in this view need to be explained in terms of how they feel. The **psychological concept of mind** sees mental states only in terms of how they cause and explain behavior. Here, mind is characterized by what it does—how it feels is irrelevant. Philosophers have concerned themselves primarily with the former, psychologists and cognitive scientists with the latter. To make this distinction clear, imagine biting into a candy bar. A phenomenal investigation would attempt to explain why you experience the mental states of “sweetness” or “chocolate” and why you might perceive them differently than somebody else. A psychological investigation would concern itself with the neural circuits that become activated during the bite, how they might be represented computationally, and how this explains when you might stop eating. In this section, we concern ourselves with the phenomenal concept of mind and its relation to consciousness, since the psychological view is in most cases the topic of the remainder of this book.

**The What-It's-Like Argument**

Nagel (1974) says there is “something that it is like” to have a conscious mental state. When you bite into a candy bar, you have a subjective conscious experience of tasting it. The candy bar of course has no such experience. There is nothing that “it is like” for the candy bar being bitten. This is one way of describing consciousness—that organisms that possess it can be described as having some sort of experience. Things incapable of supporting consciousness cannot.

But what is this experience like? Nagel asks us to imagine what it must be like for a bat to navigate by echolocation. In echolocation, the bat emits high-pitched sounds. The sound waves bounce off an object in the animal’s path and
the animal uses the reflection time as a measure of the object's distance (see Figure 2.7). We could conceivably build a machine that could compute echolocation the same way a bat does. It might even perform as successfully as the bat. But this would not tell us what it is like for the bat to experience the world in the way it does. We have seen this argument before in our evaluation of functionalism. There we said that a functional description of a cognitive process does not account for the qualia, or subjective experience, of the process.

The problem here is that science can only provide an objective account of a phenomenon and consciousness is an inherently subjective state. As organisms capable of supporting consciousness, we can introspect and analyze what it is like to have or to experience a mental state. Unfortunately for cognitive science, this is not what is needed. Cognitive science must instead have a scientific and objective account of what consciousness is. Frank Jackson (1982) aptly illustrates the difference between objective and subjective accounts of a conscious experience. He asks us to think about a neuroscientist named Mary who is well trained in the physical mechanisms underlying color vision. She understands everything there is to know about how the eye and brain process color information. Mary, however, is colorblind. Imagine now that we take away her
colorblindness and allow her to look at a tomato. Interesting questions arise. Does Mary learn anything new by this experience? Does the scientific community gain anything by Mary's (or anybody else's) description of what it is like to see red? Jackson argues that we do gain something, and that science needs to explain this new information. This gulf between an objective and a subjective description of mental phenomena goes by the name of the explanatory gap.

In contrast to this position, some state that subjective knowledge is not factual knowledge at all and therefore does not constitute any kind of an explanation. Personally knowing what it is like to taste a candy bar or to see red is not the same thing as objectively and factually knowing it. Adopting this position, we as investigators would be forced to ignore introspection and any other form of subjective description. Our focus would be only on legitimate objective techniques for studying the mind, such as experimentation and brain scanning.

Mind as an Emergent Property

Consciousness is a “hot” topic in contemporary cognitive science. In the past fifteen or so years there have been renewed interdisciplinary efforts to understand it. A number of well-known authors have published books for academic and layperson audiences that outline their definitions and views on the subject. These authors’ theories are too numerous to mention here. We instead describe one popular theory in this section and a more detailed description of another in the In Depth part of the chapter.

John Searle (1992) introduces a new twist on consciousness in his book The Rediscovery of Mind. He argues that consciousness is an emergent property of the brain. An emergent property of a system is realized through the interaction of the system’s parts. He says if we have a given emergent system S, made up of elements a, b, c, and so on, then the features of S may not be the same as the features of a, b, c, and so on. This is because the features of S arise from the causal interactions of the parts. Water for example has the features or properties of “liquidity” and “transparency.” The H₂O molecules that make it up do not share these properties. The causal interactions of these molecules, that is, their bumping into each other, give rise to these properties. In the same way, Searle says, consciousness is a property of the brain but not of its parts. If we take neurons to be the relative parts, then they have their own properties, such as being able to communicate via electrical signals with one another. These properties that are inherent in the way the neurons interact give rise to consciousness, but the properties of individual neurons need not be those of a conscious mind.
Searle is very careful to point out that he is not a reductionist. He does not believe consciousness is reducible to its parts. In fact, emergence is just the opposite of reductionism. In reductionism, explanation goes downward and a phenomenon is directly explainable in terms of what is happening at a smaller scale. In emergence, explanation goes upward. The smaller now gives rise to the larger. The large-scale phenomena are more than just what is happening in and around the parts and cannot be explained solely by an account of what the parts are doing. This idea is similar to the concept of a gestalt in perception. Gestalts are discussed in the next chapter (The Psychological Approach).

Searle seeks to avoid the monism-dualism dichotomy of the mind-body problem. He does this by talking about consciousness as a property rather than a substance. He likens consciousness to an emergent characteristic of what brains do in the same way that digestion is what stomachs do, or photosynthesis is what plants do. He sees consciousness as a natural process and a byproduct of the brain’s nature. However, he does classify conscious mental states as separate from physical ones. He states that they constitute a unique and novel category of phenomena, with an independent reality and a distinct metaphysical status.

Evaluating the Emergent View of Mind

As appealing as this formulation is, it still leaves us with some vexing questions. The reformulation of consciousness as a property, and a nonphysical one at that, still begs the question: what is a property? If a property is not physical, then of what substance is it? Although attempting to avoid the mind-body debate, Searle seemingly ends up as a type of dualist. Restating consciousness as a nonmaterial property of a material brain doesn’t get us any further toward understanding what this type of property is. Also, it is not clear how emergence happens, that is, we do not yet have an understanding of the relationship between microscopic and macroscopic properties. In the case of water, we can say its properties have something to do with the three-dimensional shape of the \( \text{H}_2\text{O} \) molecules and other conditions, such as the surrounding temperature. For consciousness and the brain, this relationship between the microscopic and the macroscopic is far more ambiguous.

Searle’s reason for believing in a nonmaterial consciousness is based on his conception of the difference between physical and mental things. For physical things, we can make a distinction between appearance and reality. A piece of wood may subjectively appear a certain way to us; as brown, as having a certain length and weight, and so on. These characteristics can also be measured
objectively; we can put the wood on a scale, use a ruler to determine its length and a wavelength detector to measure its color. For mental things this distinction between the subjective and objective goes away. As regards mental experience, Searle believes that appearance is the same as reality and that our subjective introspections are objectively correct. But if this were true, we would have to trust our intuitions about the mental world as metaphysically special and nonmaterial.

In opposition, Paul Churchland (1995) points out that this reliance on the infallibility of introspection is an outdated notion. He notes that introspection often does not give us direct and accurate knowledge of the mental. Our inner assessments of mental states can be quite often and notoriously wrong. It is commonplace for us to err in judging our thoughts, feelings, and desires. Many of the early psychologists relied on introspection as a means to study the mind. The next chapter provides a more elaborate discussion of the problems they encountered.

**Consciousness and Neuroscience**

What does the brain have to do with consciousness? Is there some part of the brain or some particular pattern of neural activity that gives rise to consciousness? What is the neural correlate of conscious experience? Although philosophers have been debating the relation between the brain and mental phenomena for millennia, recent advances in neuroscience have yielded more specific insights into these questions. Let’s examine some of them here.

In general, the neuroscience view is that consciousness results from the coordinated activity of a population of neurons in the brain. Popper and Eccles (1981) see consciousness as an emergent property of a large number of interacting neurons. A different idea is that there are neurons specifically devoted to producing consciousness. Crick and Koch (1995) believe these are located throughout the cortex and in other areas associated with the cortex. Activity in at least some subset of these neurons produces conscious experience. They believe that these neurons are special and that they differ from other neurons in terms of their structure and function. A similar but slightly different conception is that any cortical neuron may contribute to a conscious experience; however, different groups of cortical neurons mediate different types of conscious experience.

If there were special consciousness neurons, where might they be located? It has been proposed that one area is the intralaminar nuclei of the thalamus (Purpura, 1997). The thalamus is a relay center for incoming sensory information.