views are behaviorist is much debated. The philosopher most often associated with behaviorism as a philosophical doctrine is Gilbert Ryle. Ryle's position is developed in The Concept of Mind (1949). Readers of The Concept of Mind, however, might doubt that Ryle's position is accurately described as behaviorist. Some of the same ambivalence extends to the work of Wittgenstein's students and followers (see, for instance, Norman Malcolm's Dreaming, 1955).

Reductionist programs in the philosophy of science of the kind advanced by Rudolph Carnap, 'Logical Foundations of the Unity of Science' (1938), and Carl Hempel, "The Logical Analysis of Psychology" (1949), were more explicitly and enthusiastically behaviorist. Hilary Putnam's 'Brains and Behaviour' (1966) comprises a withering attack on these and other strains of behaviorism. Behaviorism's association with verificationism probably accounts for its lingering well past its heyday. (Verificationists, who trace their ancestry to the British empiricists, hold that the meaning of claims purporting to be about the world must be analyzable into sentences concerning actual or possible observations.) W. O. Quine's Word and Object (1960) expresses strong behaviorist sympathies, and Daniel Dennett (The Intentional Science, 1987), a student of Ryle's, could be read as advancing a nuanced brand of behaviorism.


5 The Identity Theory

5.1 From Correlation to Identification

Let us, at least for the time being, banish thoughts of behaviorism — philosophical and psychological — and revert to our Cartesian starting point. Let us suppose that states of mind are states of something — some thing, a substance — the mind. Descartes argues that minds are distinct from bodies. And if states of mind are not states of the body, they are not states of some part of the body, states of the brain, for instance.

What makes Cartesianism so unappealing to the modern temperament is not Descartes's contention that minds are entities, but that minds are nonmaterial entities. The more we learn about the nervous system, the more we discover intimate connections, or at least correlations, between mental occurrences and neurological goings-on in the brain. (I follow custom and speak of goings-on in the brain. This should, however, be understood as shorthand for goings-on in the central nervous system. Alternatively, you could think of the brain as distributed throughout the body.) Suppose these correlations were perfect: every kind of mental state or process could be matched to a definite kind of neural state or process. Your undergoing conscious experiences of a particular kind — you seeing a particular shade of red, for instance — might invariably be accompanied by brain processes of a particular kind, perhaps the firing of a particular collection of neurons. Finally, suppose the brains of others undergoing similar experiences exhibit similar processes.

What are we to make of this? Cartesianism explain the correlations as resulting from causal interactions between minds and brains. Such correlations would resemble correlations between falling barometers and the advent of rain. Another possibility is epiphenomenalism: the correlations are the result of mental goings-on being systematically produced as by-products of neurological activity. Occasionalism offers a third possibility: every mental and material event is willed by God in such a way that occurs in orderly patterns.

Each of these ways of accounting for one— one mental—material correlations is founded on the assumption that mental states or events are, and
must be, distinct from physical states or events. Suppose you doubted this assumption, however; suppose you identified conscious states with states of brains; suppose you thought mental occurrences were at bottom nothing more than goings-on in the brain? This is the mind–brain identity theory. Identity theorists hold that mental goings-on are not merely correlated with material goings-on in the brain. Indeed talk of correlation here is misleading. Mental goings-on are brain processes, and you cannot correlate something with itself. Brain events are correlated with subjects’ reports of mental events (recall the discussion in § 4.3). The ‘correlations’ are analogous to correlations in the whereabouts of the butler and the murderer, when the butler is the murderer.

5.2 Parsimony

Its proponents argue that, other things equal, the identity theory is preferable to its dualist rivals for two reasons. First, and most obviously, the identity theory provides a straightforward solution to the mind–body problem. If mental events are neurological events, then there is no special difficulty in understanding how mental events might cause or be caused by material events: a mental event’s causing a material event (or vice versa) is simply a matter of one material event’s causing another.

A second consideration favoring the identity theory is parsimony. Both the identity theory and dualism grant the existence of brains and neurological goings-on. But dualism must suppose that, in addition to – over and above – brains and neurological goings-on, there are minds and mental goings-on. Why posit these additional items unless you are forced to? If you can account for mental phenomena solely by reference to brains and their properties, why follow the dualist in envisaging an independent realm of minds and mental properties?

You have already encountered appeals to parsimony – Ockham’s Razor – within the dualist camp (§ 3.6). Epiphenomenalists hold that epiphenomenalism provides a simpler, hence preferable account of the mind and its relation to material bodies than do competing dualist theories. In assessing this line, I noted that appeals to simplicity ought not to be understood as based on the assumption that the universe must be simple – whatever that might mean. Parsimony figures in the comparative evaluation of theories or explanations.

Imagine two theories, $T_1$ and $T_2$, both of which provide explanations for some particular phenomenon. Suppose, further, that $T_1$ and $T_2$ both appeal to familiar kinds of entity and process, but that $T_2$ appeals, in addition, to a kind of entity not invoked by $T_1$. In that case, assuming the theories are otherwise on a par, parsimony advises that you choose the simpler (‘more parsimonious’) theory, $T_1$. The extra entities appealed to by $T_2$ are not, it seems, needed to explain the phenomenon. Simpler theories are preferred by default.

An example should make this clear. One morning you discover that the milk in the refrigerator has curdled. You recall that you left the milk out the previous night, but it seemed fine when you returned it to the refrigerator. An acquaintance explains that milk’s curdling is a matter of its undergoing a particular kind of chemical process. This process is inhibited, but not completely blocked, by refrigeration. The milk began to curdle when it was left unrefrigerated, a process subsequent refrigeration was inadequate to subdue.

Meanwhile, a second acquaintance tells you that the milk curdled because it was left out during a full moon. In this case, a chemical explanation is a more parsimonious explanation than one invoking the moon. Appeals to well-understood chemical mechanisms of the sort involved in milk’s curdling are widespread. Appeal to moon effects in this case involves what appears to be a gratuitous – unparsimonious, unmotivated – complication of the chemical picture.

In the same vein, identity theorists contend that, provided the identity theory and dualism both account for the phenomena, the identity theory wins by default. Why appeal to nonmaterial entities, states, or processes, to explain the activities of sentient creatures when material entities, states, or processes will do the job? The next question is, does the identity theory do the job, does the identity theory account for the phenomena?

5.3 Self-Conscious Thought

You might begin by asking, as Descartes does, whether states of mind could be states of the body, more specifically, states of the brain or central nervous system. The chief reason for thinking that states of mind not merely are not, but could not be brain states is that mental and material states appear to be radically different in kind. If $A$’s are different in kind from $B$’s, then there is no chance that $A$’s could turn out to be $B$’s. In the case of states of mind and material states of the nervous system, these differences are both epistemological and ontological. (Here it might be helpful to review Figure 2.1.)

On the epistemological front, you need only recall that the ‘access’ you and others enjoy to your states of mind is notably asymmetrical. Your mental life is ‘private’ in a way that no material object or state ever is. You are aware of your own states of mind and various qualities of your own conscious experiences directly and without evidence or observation. I, in contrast, have access to your mental life, at best, only indirectly. I infer your thoughts and feelings by observing your circumstances, demeanor, and behavior, verbal or otherwise. Suppose I observe goings-on in your brain, and suppose these goings-on are known to be reliable indicators of your mental condition. In that case I would be in an epistemologically strong position to know what you are thinking and feeling. Even so, my access to your thoughts and feelings differs from yours. I must infer what you experience directly.
A Cartesian would explain this epistemological asymmetry by noting that others' knowledge of your states of mind depends on observations, not of the states themselves, but only of their effects on material bodies. My knowledge of your states of mind is epistemologically on a par with a doctor's knowledge that you have chicken pox based on the doctor's observation of a rash on your stomach. Your knowledge of your own mental life is unmediated, however. Indeed it is at the very least misleading to imagine that you observe your own thoughts and feelings.

This last point deserves elaboration. Reflect on your thoughts. Each thought carries with it the potential for self-awareness — an idea implicit in Descartes and explicitly emphasized by Immanuel Kant (1724–1804). Thinking is not something that occurs to you like the beating of your heart, something concerning which you are a mere spectator. Thinking is something you do. And, like anything you consciously do, you need not observe yourself in the act to recognize that you are doing it. (I leave aside for the moment consideration of nonconscious states of mind.)

When you entertain the thought that it is raining and consciously recognize that this is what you are thinking, your conscious recognition is not based on some further act of inward observation of the original thought. That thought, rather, is thought self-consciously. If every thought is at least potentially self-conscious, then this self-conscious thought could itself be entertained self-consciously: you could become aware that you are thinking that it is raining. Try it! And note that this thought is itself perfectly self-contained; it is a single thought, not a sequence of distinct thoughts.

Any account of the mind must, it would seem, accommodate the possibility of self-consciousness. The Cartesian view does so by building it into the nature of the mental: that states of mind are 'self-revealing' is just one way minds differ from material bodies. States of mind, as it were, reflect themselves. Brain states, in contrast, are like any other publicly observable material states. Our access to them is grounded in ordinary observation. How could something epistemologically private possibly be identified with something publicly observable? Anyone who hopes to assimilate minds to bodies — or, more particularly, to brains — must be prepared to answer the Cartesian on this score.

5.4 Locating Mental Qualities

A second hurdle facing anyone aiming to substitute brains for Cartesian minds is ontological. Mental events, states, and properties appear to be utterly different in kind from material events, states, and properties. The difference is striking when you consider qualities of your own conscious states of mind and compare these with the qualities of material bodies, including the qualities of brains. Your visual experience of a ripe tomato in bright sunlight seems wholly different qualitatively from going on in your nervous system.

Neurological occurrences can be observed and described in great detail. But observe the brain as you will, you seem never to observe anything remotely resembling a conscious experience. Indeed, the very idea that you might observe someone else's thoughts or experiences appears unintelligible.

You might hope to sidestep this problem by appealing to the fact that science regularly tells us that things are not as they seem. Take the ripe tomato. You experience the tomato as possessing a particular color. But physicists like to tell us that the experienced color is in a certain sense an illusion. The tomato's surface exhibits a particular microscopic texture. Surfaces with this texture reflect light in a particular way. And reflected light of this sort, when analyzed by the human visual system, gives rise to an experience of red. It would be a mistake to locate a feature of our experience in the tomato. Considered in its own right, the material universe is colorless.

Similar arguments can be concocted to show that sounds, tastes, smells, and the way things feel are, in the sense described, absent from the material universe. A long tradition, going back at least to Galileo, Descartes, and Locke, classifies colors, sounds, and the like, 'secondary qualities'. Secondary qualities could be thought of as powers possessed by material bodies to produce experiences of familiar kinds in conscious observers.

Whatever its merits, this train of thought does little to advance the case of the anti-Cartesian. If characteristics of conscious experiences of colors are not characteristics of material bodies, then what are they characteristics of? A physicist can banish them to the mind. But this move implies that minds do not themselves belong to the material universe! If you distinguish appearance from material reality by relegating appearances to the mind, you thereby place minds outside the material realm. Assuming that science is devoted to the investigation of the material universe, this line of reasoning would land you back with a Cartesian conception: minds are distinct from material bodies.

Notwithstanding these difficulties, many philosophers (and many more scientists) have been attracted to the view that, at bottom, minds are in fact material bodies. When a material body is organized in a particular way, organized in the way a human brain is organized, for instance, you have a mind. In the end, mental characteristics are, despite appearances, material characteristics.

As you have seen, the impetus for such a theory is twofold. First, the more scientists investigate the brain, the more they uncover intimate, fine-grained connections between neurological and mental goings-on.

Second, by far the simplest, most straightforward explanation of these connections is that minds are brains, mental states and processes are neurological states and processes. The idea that complex actions and reactions among the ultimate components of the universe are uniformly physically explicable — except for those occurring in the brains of sentient creatures — seems incredible. Thus, a view according to which there is at most one kind of substance, material substance, appears preferable to a dualistic view, on
grounds of simplicity—assuming, of course, that it is possible somehow to account for central features of our mental lives without having to introduce nonmaterial substances. But is it possible?

5.5 Substances, Properties, States, and Events

What has come to be called the identity theory of mind emerged simultaneously in the United States and Australia in the 1950s in papers published by Herbert Feigl, U. T. Place, and J. J. C. Smart. According to the identity theory, minds are material entities—brains—and mental properties are, as a matter of discoverable empirical fact, material properties of brains and nervous systems. In claiming that mental properties are material properties, Feigl, Place, and Smart were not claiming merely that mental properties were properties of material bodies. You might think this, and yet imagine that mental properties were quite different from nonmaterial mental properties. The result would be a substance monism coupled with a dualism of properties (as discussed in Chapter 11). Identity theorists, however, hold that every mental property, every property you might regard as mental, is in reality a material property, that is, a property of the kind appealed to in the physical sciences.

Earlier, I spoke of mental (and material) characteristics, states, and the like. Identity theorists talk of identifying mental processes with brain processes. Now I am formulating the identity theory as a theory about properties. These terminological vicissitudes deserve comment.

In Chapter 2, I followed Descartes in distinguishing attributes and modes from substances. Descartes's modes are what are more familiarly called properties; Cartesian attributes are kinds of property. To put it somewhat unhelpfully, a substance is an individual possessor of properties; and properties are what substances possess: ways those substances are. Substances themselves differ from one another both numerically, and with respect to their properties. Here is one billiard ball, and here is another, distinct, billiard ball. The balls differ numerically. But they also differ in color: one is red, the other white.

You will recall that Descartes distinguishes material and mental substances by reference to their fundamental attributes: a material substance is extended; a mental substance thinks. Properties possessed by material bodies are, one and all, modes of extension—ways of being extended; and particular thoughts, feelings, or sensory experiences possessed by nonmaterial substances are modes of thought—ways of thinking. No substance that possesses properties of the one sort could possess properties of the other sort; and the possession of one kind of property precludes possession of the other.

You might balk at a substance dualism of this sort, yet join the dualist in insisting on a distinction between mental and material properties. You might think, for instance, that there is a single substance, the body, or maybe the brain, and this substance possesses both mental and material properties. This is not what identity theorists have in mind, however. Their contention is that every mental property just is—its identical with—some material property. (I shall return to the notion of identity in play here presently.)

So much for properties. What of states, events, processes, and the like? Think of a state as a substance's possessing a property. There is the billiard ball, a substance, and the ball's redness, a property of the ball. The ball's being red is a state of the ball. Suppose that being angry is a mental property. Then your possessing this property, your being angry, is for you to be in a state of anger. If the state of anger turns out to be some neurological state, then this is so because the property of being angry is identical with—its a certain neurological property.

You could think of events and processes as state transitions. An object's coming to be in a particular state, the object's coming to possess a certain property, would be an event. Finally, a process would be a sequence of events. Now it is possible to formulate an identity principle.

\( \text{(i) a state, event, or process, } \alpha \text{, is identical with state, event, or process } \beta \text{, only if the properties involved in } \alpha \text{ and } \beta \text{ are identical.} \)

Trying to keep all this straight could induce a certain amount of giddiness. Imagine that it does, and that your feeling giddy is a matter of your being in a particular state of mind. Your being in this state is your possessing a certain, possibly elusive, mental property. Now suppose the question arises whether your feeling giddy—your being in this mental state—is your being in a particular brain state (and nothing more). If you agree with the identity theorist that this is what your feeling giddy is—your brain's being in a particular state—then you will accept the identification of the mental property, being giddy, giddiness, with some definite neurological property.

The moral: so long as you bear in mind that state, event, and process identity requires property identity, you can speak indifferently of the identity theory as identifying mental and material states, processes, events, or properties.

5.6 Predicates and Properties

Getting clear on what it means to identify properties \( \alpha \) and \( \beta \), requires distinguishing properties and predicates. Sidestepping assorted complications, you could think of predicates as linguistic devices, terms used to ascribe properties to objects. The English expression 'is round' is a predicate used to characterize objects: coins, rings, manhole covers, dinner plates, automobile tires. This predicate holds of a given coin, in virtue of that coin's being a certain way: being round. The predicate 'is shiny' holds of the same coin in virtue of the coin's being another way: being shiny. These distinct 'ways' answer to distinct predicates.
One reason to make the predicate-property distinction explicit is that, as you will see below, it would be impossible to understand what claims about 'properties being identical' amount to otherwise. Another reason is that philosophers all too often assume without argument that every predicate capable of meaningful application to an object designates a property of that object. Such a view is ill-advised. What the properties are is largely a question for science.

Predicates can be constructed ad lib to suit all manner of parochial needs and interests, including those of philosophers writing textbooks. Consider the predicate 'is a left ear or made of copper.' This predicate holds of many objects. It holds of anything that is a left ear (so it holds of your left ear) and anything that is made of copper (so it holds of the penny on your dresser). The predicate holds of objects in virtue of properties they possess, ways those objects are. It does not follow, however, that the predicate designates a property in the sense that objects that it applies to possess the very same property. These are deep waters. I shall return to them in later chapters. In the meantime, you need only bear in mind that properties and predicates, even those predicates that uncontroversially designate properties, belong to different orders: predicates are linguistic, representational; properties are nonlinguistic features of objects.

5.7 Strict Identity

Identity theorists contend that mental properties are identical with material, physical properties. It is time to say what exactly this means. Consider, first, how identity applies to objects or substances as distinct from properties.

Our concept of identity, of self-sameness, is indispensable. This is due in part to the fact that it is common to speak or think of a single object in different ways. You need some way of registering that distinct ways of speaking about something are ways of speaking about one and the same something. John le Carré is an author. Saying that le Carré is an author, amounts to saying something about le Carré, attributing a characteristic to him. As it happens, John le Carré is David Cornwell. Saying that le Carré is Cornwell is to say that the man called 'le Carré' and the man named 'Cornwell' are the selfsame individual.

1 Le Carré is an author ('is' of predication)
2 Le Carré is Cornwell ('is' of identity: le Carré, is identical with Cornwell)

Any object could be given multiple names; any object could be described in different ways. You could know an object under one name or description, but not under another. Imagine that you are traveling in Australia, intending to visit Uluru. En route, you hear talk of an impressive rock edifice, Ayers Rock, and regret not having enough time to visit it as well. Much later you discover that Uluru is Ayers Rock. In hiking around Uluru, you hiked around Ayers Rock without knowing it.

If α and β are strictly identical, α is β. But debutantes who wear identical gowns to the ball do not occupy the selfsame gown. Wayne and Dwayne are identical twins, but Wayne isn't Dwayne. In these cases 'identical' means, not 'one and the same', but 'exactly similar'. The distinction between strict identity and identity-as-exact-similarity is signaled in English by the use of different prepositions: le Carré is identical with Cornwell; Wayne is identical to Dwayne. A word of warning: this grammatical subtlety is often flouted by careless writers, including philosophers writing on identity.

The identity theory extends the notion of identity — strict identity — to properties. Like objects, properties can be the subject of identity claims.

3 Red is the color of ripe tomatoes.

In this case, a single property, a color, is designated by distinct predicates: 'is red' and 'is the color of ripe tomatoes.' And just as you might be familiar with Ayers Rock and with Uluru without realizing that Ayers Rock and Uluru are identical, without realizing that Ayers Rock is Uluru, so you might fail to realize that two predicates in fact designate the selfsame property. You might know that a particular color is red without knowing that it is the color of ripe tomatoes — if, for instance, you were ignorant of tomatoes.

Identity theorists focus on what Smart calls theoretical identities. Such identities are uncovered by scientists exploring the way the universe is put together. Lightning, scientists discovered, is an electrical discharge; water is H₂O; temperature is mean molecular kinetic energy; liquidity is a particular kind of molecular arrangement. An identity theorist holds that it is a good bet that research on the brain will lead to the discovery that certain properties we now designate using mental terms are in fact properties of brains. Pain, for instance, might turn out to be the firing of C-fibers in the brain. (This, a standard example, has been empirically discredited, but it will do to illustrate the point.) Were that so, the property of being in pain would be identified with the neurological property of being a C-fiber firing. Believing this is to believe that pain is C-fiber firing.

Identity theorists do not set out to advance particular identity claims. The establishment of these would be the job of brain researchers who uncover correlations between goings-on in the brain and subjects' reports of experiences. Rather, the identity theory offers an interpretation of these results: in reporting conscious experiences, you are reporting goings-on in your brain. Details will be revealed as the neurosciences move forward.

5.8 Leibniz's Law

Strict identity is self-sameness. If object α and object β are strictly identical (if α = β), then any property of α must be a property of β, and vice versa. This principle — called 'Leibniz's Law' in honor of the philosopher who first articulated it explicitly — provides a test for strict identity. You can decisively establish that
some $\alpha$ is not identical with some $\beta$, if you could show that $\alpha$ possesses some property that $\beta$ lacks, or that $\beta$ possesses some property lacked by $\alpha$. The butler could not be the murderer if the murderer has a limp and the butler does not.

Applying Leibniz's Law to the case at hand, minds could not be identified with brains if minds had properties lacked by brains or brains possessed properties not possessed by minds. Similarly, brain states, events, or processes would fail to be identical with mental states, events, or processes, if brain states, events, or processes involved properties absent from mental states, events, or processes; or if mental states, events, and processes involved properties absent from brain states, events, or processes.

What of properties themselves? Suppose property $\psi$ is strictly identical with property $\phi$. Does this mean that every property of $\psi$ must be a property of $\phi$ and vice versa? That might be so if $\psi$ and $\phi$ were complex properties, properties that had properties as parts. In that case $\psi = \phi$ only if every constituent property of $\psi$ is a constituent of $\phi$ and vice versa. If $\psi$ and $\phi$ are simple, you reach the limit: $\psi$ is identical with $\phi$ tout court.

Suppose $\psi$ and $\phi$ are properties, you are familiar with both $\psi$ and $\phi$, and $\psi$ and $\phi$ are, as it happens, identical. Could you fail to know that $\psi$ and $\phi$ are identical? You could fail to know this, presumably, just as a detective could fail to know that the butler is the murderer despite knowing a good deal about both. This is how it is with mental and material properties according to proponents of the identity theory. Mental properties are material properties, although this is not something you could discover without expending considerable effort – just as considerable effort was required to discover that lightning is an electrical discharge and that water is $H_2O$.

5.9 The $64$ Question

The question to be faced now, the $64$ question, is whether it is even remotely plausible to suppose that mental properties, the kinds of property the having of which might constitute your undergoing a particular conscious experience, for instance, could turn out to be properties of your brain? There appear to be powerful reasons to doubt this possibility. As noted earlier, the qualities you encounter when you undergo a conscious experience seem nothing at all like the qualities scientists find when they inspect brains. (For a convenient listing, see Figure 2.1.)

Imagine that it is a sunny day and you are standing in Trafalgar Square watching a red double-decker bus rumble past. You have a visual experience of the red bus, you hear it, very probably smell it and, through the soles of your feet, feel its passing. The qualities of your conscious experience are vivid and memorable. But now, could anyone seriously think that, were neuroscientists to open your skull and observe the operation of your brain while you were undergoing this experience, the neuroscientists would encounter those qualities? And if this is implausible, how could anyone seriously hope to identify experiences with brain processes?

Pretend for a moment that the identity theory is correct, pretend that states of mind are brain states. Your undergoing an experience — seeing, hearing, feeling, smelling the passing bus — would presumably be a matter of your brain's undergoing a complex sequence of processes (Figure 5.1).

Next imagine that a neuroscientist is observing your brain's undergoing this sequence of processes. The neuroscientist's conscious experiences of your brain would themselves be a sequence of processes in the neuroscientist's own brain (Figure 5.2). (Remember, we are assuming for the sake of argument that the identity theory is correct.)

Now ask yourself, is it really so obvious that the qualities of your experience differ from qualities the scientist observes when the scientist probes your brain?
In considering this question, it is vital not to confuse qualities of the scientist’s experience of your brain with qualities of your experience of the passing bus or, what comes to the same thing if the identity theory is true, qualities of your brain. The qualities to consider in this case are qualities of processes in the scientist’s brain that (we are assuming) coincide with the scientist’s observation of your brain. What needs to be compared is, to put it crudely, how a conscious experience looks to an observer and what the experience is like for someone undergoing it. This means that you must compare qualities of the observing scientist’s conscious experiences of your brain (which, by hypothesis, are themselves neurological goings-on) with qualities of your conscious experiences (also, presumably, neurological events). And, although these will be different — observing a brain differs qualitatively from observing a passing bus — there is no reason to think that they must be dramatically different in kind.

The moral is that, if the aim is to compare the qualities of conscious experiences with the qualities of brains, you must be careful to compare the right things. If the identity theory were correct, your undergoing a conscious experience would be a matter of your brain’s undergoing a complex process. If you want to compare qualities of your conscious experience with observations of your brain, then the appropriate target of comparison would be the brain of the observer. Goings-on in the observer’s brain are what would constitute, for the observer, the ‘look and feel’ of your brain.

All this is just to insist on a simple point: undergoing an experience is one thing, observing the undergoing of an experience (a distinct experience) is something else again. The qualities of these will certainly be different. Looking at a brain, after all, is nothing at all like watching a passing bus. But the qualities need not be radically different in kind — radically different in the way harped on by dualists.

You might think that these remarks completely miss the point of the original worry: when you consider ‘what it is like’ to observe the passing bus, when you reflect on the qualities of this experience, those qualities seem not to be remotely plausible candidates for possible qualities of brains. You know what your conscious experiences are like and you know enough about what brains are like to make it seem obvious that conscious experiences are nothing at all like neurological goings-on. If that is so, the identity theory, whatever its scientific appeal, cannot get off the ground.

5.10 The Phenomenological Fallacy

You might be tempted to reason as follows. When you observe the passing bus, you observe something red, loud, and smelling of diesel fumes. But redness, loudness, and that distinctive diesel odor are not found in your brain. If a scientist scrutinizes your brain when you are undergoing this experience, the scientist will not find anything that possesses these qualities. The philosopher Leibniz, who made an appearance earlier as author of Leibniz’s Law, provides an analogy.

Suppose there were a machine so constructed as to produce thought, feeling, and perception, we could imagine it as increased in size and while retaining the same proportions, so that one could enter it as one might a mill. On going inside we should only see the parts impinging on one another; we should not see anything which would explain a perception.

(1787, 181)

(Leibniz goes on to argue that ‘the explanation of perception must be sought in a simple substance, and not in a compound or in a machine’.)

Despite its seeming plausibility, this line of reasoning is seriously flawed. When you undergo a conscious experience — when you observe the passing bus, for instance — your experience is qualitatively saturated. But what exactly are its qualities, what are the qualities of your experience? Whatever they are, they are not to be confused with the qualities of objects and events observed, in this case qualities of a passing bus. Your experiencing a bus is one thing, the bus is another. It is natural, maybe unavoidable, when describing experiences to describe objects experienced and their qualities. What it is like to experience the bus, however, the qualities of your experience of the bus, are not qualities of the bus.

The identity theory identifies your experience of the bus with some occurrence in your brain. As noted, we typically describe our experiences by reference to objects that typically cause them, what they are experiences of. You can convey to me an experience you had at Trafalgar Square by telling me that it was an experience of a passing red double-decker bus. I have a decent idea what it is like to observe passing red double-decker buses, and so I acquire a sense of what you experienced. But, again, the qualities of your experience are not to be confused with the qualities of their objects, qualities of what you experience. An experience of something red, massive, and smelly is not itself red, massive, and smelly (Figure 5.1).

This point was one insisted on by both Place and Smart in their respective defenses of the identity theory, but it has not always been fully appreciated. The rhetorical punch of the dualist’s contention that it is just obvious that qualities of experiences differ from brain qualities relies heavily on your tacitly identifying, as Leibniz apparently does, qualities of experiences with qualities of objects experienced. In his original 1956 paper defending mind–brain identity, Place dubbed the mistake of confusing properties of objects experienced with properties of experiences, the phenomenological fallacy.

Once you distinguish qualities of experiences from qualities of what is experienced (and these must be distinguished on any view), it is much less obvious that experiences could not turn out to be brain processes. Anyone who persists in claiming that experiential qualities differ in kind from
neurological qualities owes us an argument. What exactly are the qualities of experience? And what reasons might there be for thinking that qualities of experiences could not be qualities of brains?

Materialist philosophers sometimes accuse their opponents of mistakenly thinking of conscious experiences as playing out on a kind of internal television screen. Think about what exactly might be wrong with such a conception. Start by considering images on a television screen. These are meant to reproduce important qualities of perceived objects. A televised billiard match results in spherical red and white images on a trapezoidal green background that, on a television screen, correspond to the colors and shapes of billiard balls arranged on a billiard table as you might visually perceive them from a nearby vantage point. The properties on the screen, the colors and shapes, are stand-ins for properties of objects perceived. But it would be crazy to think of these properties as themselves properties of conscious experiences. Whatever conscious experiences are like, whatever their properties, they are not like this. (This point will reassert itself in Chapter 10.)

The suggestion on the table is that the distinction between qualities of experiences and qualities of objects experienced is theory-neutral. The distinction must be made by dualists as well as proponents of the identity theory. It is worth pointing out that, provided you keep the distinction firmly in mind, you can begin to make sense of a range of mental phenomena that might appear baffling otherwise. This is so, whatever you ultimately conclude about the status of mental properties.

Consider dreams, mental images, and hallucinations. Some theorists have wanted to downplay such phenomena, or depict them as purely cognitive processes. The worry is that there is no room in the brain for images, hallucinations, or dreams, the qualities of which appear to differ dramatically from qualities discoverable in brains. Suppose you hallucinate a pink penguin (or dream, or form an image in your mind’s eye of a pink penguin). Nothing in your brain is pink or penguin-shaped. Indeed, it is entirely possible that nothing anywhere in your vicinity (or, for that matter, anywhere at all!) is pink and penguin-shaped.

But if this is supposed to cast doubt on hallucination, dreaming, or imagery, it succeeds only by conflating qualities of objects hallucinated (or dreamed, or imagined) with qualities of the hallucinatory (or dream, or imagistic) experience. Visually hallucinating a pink penguin ought to resemble having a visual experience of a pink penguin, not a pink penguin. Just as the experience is not pink and penguin-shaped, neither is the hallucinating pink or penguin-shaped. Nor need we suppose that hallucinating, or imagining, or dreaming of a pink penguin is a matter of inwardly viewing a picture-like image of a pink penguin.

My hope is that, your appreciating these points might enable you to relax a bit and think more clearly about the character of hallucination, mental imagery, and dreaming. (I shall have more to say about the importance imagery and the qualities of conscious experiences in Chapters 10 and 13.)

5.11 Epistemological Loose Ends

What can an identity theorist say about the asymmetry of ‘access’ to states of mind? You have ‘privileged first-person access’ to your thoughts and sensory experiences. The rest of us have, at best, indirect, ‘third-person access’ to your mental life. Your mental life is private, but your brain is public. A scientist could observe and measure goings-on in your brain, but only you are in a position to be ‘directly aware’ of your thoughts and feelings. If those thoughts and feelings were brain processes, however, they would be open to public scrutiny, an apparent impossibility!

These are difficult issues. Materialists must somehow accommodate asymmetry of access without backsliding into dualism, without, for instance, resorting to the notion that mental occurrences take place in a private interior chamber, visible only to the agent to whom they belong. (Recall Wittgenstein’s beetle in the box example in § 4.4.) But how else could the privileged access each of us enjoys to our own states of mind be explained? How else might we account for the asymmetry?

Consider, first, an observation made earlier concerning conscious thought. Thinking is something you do. Like anything you do, in doing it, you are in a position to appreciate what you are doing. To be sure, you rarely bother to reflect on the fact that you are doing what you are doing. But, when you do reflect, you are not acting as an observer — even an especially well-placed observer — of what you are doing. Your recognition of what you are up to stems from the fact that it is you who are up to it.

Imagine that you draw a diagram on the blackboard to illustrate a point about the economy of pre-Roman Britain. I am in the audience. Compare your understanding of the diagram with mine. I observe what you have drawn and endeavor to interpret it in light of your lecture. You, in contrast, grasp its significance immediately. You are able to do this, not because you have a better, more intimate view of the diagram, not because you are much closer to it, but because it is your diagram: you drew it with this significance.

You bear an analogous relation to your own thoughts. You immediately grasp the significance of those thoughts, not because your view of them is especially acute or unimpeded, but because you think them. Because you do not always do what you set out to do, your capacity to recognize what you are doing is not infallible. You take yourself to be walking west, when you are, in reality, walking east. In the same way, you might take yourself to be thinking of your grandmother, when in reality you are not: the person you have always assumed was your grandmother is an impostor!

What of your ‘access’ to your own conscious sensory episodes? Your recognition that you are suffering a headache is apparently direct and unmediated in a way anyone else’s access to your headache never is — or could be. Does this imply that headaches are states or processes ‘visible’ only to those undergoing them?
Two points bear mention. First, in undergoing a conscious sensory experience, you do not (1) have the experience, and (2) observe — perhaps in an especially intimate way by means of an inward-looking perceptual organ — the experience. Your awareness of the experience is constituted, at least in part, by your undergoing it. This is why talk of ‘access’ to one’s sensory experiences is potentially misleading. Your recognition that you have a headache is constituted, in part, by your having or undergoing the headache. Differently put: your conscious experience of the headache is a matter of your having it. It is not that the headache occurs and, in inwardly observing it, you ‘access’ it.

Second, and to echo a point made earlier, it is important to distinguish a system’s undergoing some process or being in some state, from observations of that system’s undergoing a process or being in a state. Your refrigerator probably defrosts automatically. The refrigerator’s defrosting on an occasion is, in an obvious way, very different from your observing its defrosting. Similarly, your undergoing a pain is very different from someone else’s observing your undergoing it, even if that someone else has an unrestricted view of the operation of your brain. Now, if ‘directly observing a sensation’ just amounts to having that sensation, then there is no puzzle at all in the idea that only you can ‘directly observe’ your sensations. This is just to say that only you can have your sensations. And this is no more mysterious than the thought that only your refrigerator can undergo its defrosting.

Considerations of this sort tell against the Cartesian picture, not by providing a refutation of that picture, but by offering an alternative depiction of what self-awareness might encompass. On the Cartesian model, self-awareness resembles the awareness of ‘external’ objects and events turned inward. As the foregoing discussion makes clear, however, you need not embrace this way of depicting the situation. And, the Cartesian conception aside, it would seem that you probably ought not to embrace it.

This does not mean that dualists must forthwith abandon dualism and accept the identity theory or some other form of materialism. It does mean that one consideration apparently favoring dualism needs to be reevaluated. It might be possible to accommodate the epistemological asymmetry associated with states of mind without recourse to dualism. (The topic will be revisited in § 13.10.)

5.12 Taking Stock

Materialism, in the guise of the identity theory, offers itself as a replacement for dualism, one that purports to explain mentality, but more parsimoniously, that is, without recourse to nonmaterial substances or properties. I have touched on one respect in which the identity theory is vindicated. Dualists sometimes argue as though it is indisputably obvious that the properties of states of mind could not be properties of brains — or indeed properties of any material entity.

As it happens, the force of this argument depends in large measure on a tacit conflation of the qualities of experiences and the qualities of what is experienced, the objects of experience. This is Place’s phenomenological fallacy. The qualities of objects experienced are indeed very different from the qualities we experience in the course of observing brains. This, however, is not something that need trouble an identity theorist. Colorful experiences are experiences of colorful objects, not experiences that are themselves colored. If dualists continue to insist that qualities of conscious experiences could not be possessed by brains, then the ball is back in the dualists’ court.

I do not mean to leave the impression that this is the end of the matter. I have suggested that it is not obvious that conscious experiences could not be identified material states and processes. Yet neither is it obvious that they could be. I counsel suspicion of anyone who claims that either answer to this question is obvious.

Another worry that I have left untended concerns the unity of experience. On the one hand, the brain is a complex system encompassing endless subsystems. On the other hand, our mental lives are apparently unified. Although we possess various distinguishable mental faculties, at any given time each of us confronts the universe as a single ego with a single point of view or perspective. (This is so, I suspect, even for persons said to possess multiple personalities.) How is this unity of experience to be reconciled with the widely dispersed and fragmented character of neural processing?

In recent years, hopes for finding a neurological ‘central processing unit’, a neurological analogue of a computing machine’s CPU, have receded. Even if we were to locate a neurological CPU, however, it is by no means clear that its operation could account for the unity of experience. A point of view is just that: a point, a position in space—time, from which the universe is apprehended. The relation of this point of view to the experienced universe resembles the relation of the eye to the visual field. The eye is not within the visual field, but stands at its limit (Wittgenstein 1922, § 5.6531; and see § 13.10 below).

Scientists and philosophers traditionally sought to reconcile appearance with reality by banishing appearances to the mind. Many apparent features of the universe — colors, for instance, or sounds, or odors — were taken to belong not to the universe, but to us, to our point of view on the universe. If your hope is to see minds and universe as parts of a single reality, however, you are faced with the task of finding a place for appearance within that reality. And this requires locating points of view on the universe wholly within the universe. The trick, as you have seen in considering the qualities of experience, is to be clear on the nature of appearance, the character of points of view.

Although these strike me as central themes, they have not played an appreciable role in philosophical attacks on the identity theory. Those attacks have centered on the claim that states of mind are functional states of creatures possessing them, not material states. Functionalism takes center stage in Chapter 6.
6 Functionalism

6.1 The Rise of Functionalism

The identity theory enjoyed a surprisingly brief period of favor among philosophers. Its decline in popularity was not the result of dualist counter-attacks, however, but a consequence of the rise of a very different conception of mind: functionalism. Functionalists were not put off by identity theorists' commitment to materialism. Although, as you will see, functionalism is not a materialist theory per se, functionalism can be understood as compatible with the aims of materialism; most functionalists would regard themselves as materialists of one sort or another.

Functionalists would allow that, although immaterial substances — Cartesian souls, for instance — are conceivable, in all probability every substance is a material substance. Were that so, every property possessed by a substance would be possessed by a material substance. Does this imply that every property is a material property? Are mental properties a species of material property? The issues here are tricky. They occupy the sections — and chapters — that follow.

These days functionalism dominates the landscape in the philosophy of mind, in cognitive science, and in psychology. Functionalism offers a perspective on the mind that suits the needs of many empirical scientists, a perspective that promises solutions to a host of long-standing philosophical puzzles about minds and their relation to material bodies. Clearly functionalism — the doctrine, if not the label — has etched its way into the popular imagination by way of the press and television. When basic tenets of functionalism are put to non-philosophers, the response is, often enough, 'Well, that's obvious, isn't it?'

This is not to say that functionalism lacks critics. On the contrary, plenty of philosophers and empirical scientists have found functionalism wanting. There is scant agreement among its opponents, however, concerning where exactly functionalism fails down. Indeed, opponents are typically willing to concede that functionalism is right about some things — although, again, what these things are is something concerning which there is little consensus. In the absence of clear competitors, many theorists have opted to