

Dualism

Common sense may regard as unusual and eccentric Descartes's dreaming and evil spirit scenarios, but it is not unfamiliar with the distinction between appearance and reality - or, more to the present point, with the distinction between mind and matter. Indeed, if his indirect realist account of perception goes against the grain of everyday thinking, Descartes's dualism - his claim that there is a "real distinction" between the mind and the body, that they are fundamentally different kinds of thing -- is quite in line with it. We reflexively distinguish between mind and body in ordinary contexts as often as in philosophical ones, and in a way that implies that the difference between them goes deeper than a mere difference between part and whole: we do not, after all, distinguish equally naturally between "hand and body" or even "brain and body." Moreover, the metaphysical content of most religions has historically included some version of the idea that a human being has a soul, regarded as the seat of our mental lives, as spiritual rather than physical, and as surviving the death of the body.

Descartes's position is intended rationally to systematize and justify this commonsense view of human nature. It is, naturally enough, referred to as *Cartesian dualism* ("Cartesian" meaning "pertaining to the thought of Descartes"), though some version of it goes back in philosophy at least to Plato. In Descartes's view, the reason mind and body *seem* different in the ways sketched in the last chapter is that they *are* different, and radically so. The body is, in its intrinsic nature, exactly like every other material object, being an essentially extended thing (in Latin, *res extensa*): extended in space, that is to say, and defined by such properties as length, depth, height, mass, motion, and spatial location. Together with other material or extended objects, it is composed of purely physical parts - molecules, atoms, and subatomic particles - and governed entirely by the causal processes enshrined in the laws of physics. The body, and the vast physical universe of which it is a part, are best thought of through the model of a machine, their operations being as mechanically automatic as those of a watch and their elements as brute and unthinking as a watch's gears and mainspring. The mind, by contrast, is essentially a thinking thing (or *res cogitans*), devoid of shape, mass, location in space, or any other

physical property, and governed by reason rather than mechanical causation. It is as utterly distinct from its associated human body as it is from the material world in general, though it does interact with it: changes in the body bringing about changes in the mind (as when the body's sensory organs detect a cheeseburger in the vicinity and produce, in the mind, hunger and an intention to eat) and changes in the mind bringing about changes in the body (as when the mind's intention to eat the burger causes the body to salivate and proceed to eat it).

Since there is a clear sense in which Descartes took mind and body to be distinct substances – a "substance" being something that exists on its own, as opposed to an "attribute" or "property", (like redness, tallness, or heaviness) which cannot exist apart from the substance which has it - his view is often described as *substance dualism*, and he is widely interpreted as regarding the non-physical substance of the mind to be what a person essentially is, the body being a mere excrescence, no more necessary to a human being per se than the clothes he or she wears. on this understanding of Descartes' view, the real you is something outside the material world altogether, an immaterial substance or soul temporarily inhabiting your body like a "ghost in the machine," as Gilbert Ryle (1900-1976; famously and derisively put it. But this interpretation, however common, is at best a caricature. In fact, Descartes took the interaction between mind and body to be so close that the two together constituted a third, unique substance, with its own distinctive properties: while shape, mass, and the like are confined to the body, and pure intellectual activity confined to the mind, sensation- pains, itches, feelings of thirst or hunger - is a feature strictly attributable only to the substance comprised of mind and body interacting together. Moreover, it is this composite substance, rather than the mind alone, with which a person or human being is to be identified.

Nevertheless, however close its connection to the body, the mind is still, in Descartes's view, distinct from it- and that means distinct from the brain, which is no less physical or extended an object than the rest of the human body. But doesn't Descartes thereby contradict common sense after all? Don't we normally use the terms "mind" and "brain" interchangeably, so that they must be regarded as the same thing . in which case the mind really is just part of the body?

Minds and brains, apples and oranges

No doubt people these days often do use these words interchangeably, but this by itself doesn't prove anything. Certainly the two words don't mean the same thing. In Aristotle's day, people knew about the brain, but did not take it to have anything to do with thinking, intelligence, or the mind in general - they thought its function was to cool the body. It is only because we now know that the brain has an intimate relationship to the mind that we so easily (and, from a philosophical point of view, carelessly) shift from talk about the one to talk about the other. Descartes himself was well aware of this connection, and nevertheless took mind and brain to be distinct. The brain was in his view the conduit through which the mind interacted with the body, but nevertheless as distinct from the mind as the wire that connects your television set to the cable company's local relay station is distinct from the television itself.

But why take them to be distinct? Why not conclude from the close connection existing between them that the mind and brain are the same?

Why do we believe that apples are different from oranges? The answer, of course, is that they just obviously are different. Oranges are orange, spherical, and have a distinct flavor very different from that of apples, which are typically red, yellow, or green and apple-ish in shape. Anyone who has observed them knows they're different; no fancy argument is needed to prove it. But the same holds true of the mind and the body, or the mind and their brain for that matter, in Descartes's view. The difference between them is "clear and distinct," as obvious as the difference between apples and oranges, and as little in need of complicated philosophical demonstration.

As we know from modern physics, a material thing is ultimately nothing more than a collection of elementary particles. That includes the cheeseburger whose appearance and aroma makes your mouth salivate and your stomach grumble in hunger, and whose flavor and texture, vividly experienced by you as you eat it, brings satisfaction. The particles comprising the cheeseburger have themselves none of these features: no color, odor, taste, or texture. Moreover, they have none of the solidity of the cheeseburger that you feel as you hold it in your hands; there is more room between the particles than is occupied

by the particles themselves, so that the cheeseburger is mostly empty space. It just happens that the particles comprising the cheeseburger are so arranged that they affect your sensory organs in such a way that you experience it as a solid, textured, colorful, aromatic, and flavorful object. Intrinsically, though, it is none of these things, and neither is any other physical object - including your brain, which is constituted of physical particles just as much as the cheeseburger. Yet these features do in some sense exist in your mind, in your experiences of the cheeseburger. But then the mind, the dualist concludes, is just obviously different from the brain, for it has qualities that the brain does not have.

Consider further the nature of experiences in general, and of their qualia. when you see Fred get his hand slammed in a car door, you have no doubt that he is in pain. But this is not because you experience or observe the pain itself; you cannot peer inside the wound and see the pain the way you might see a splinter. you might observe the behavior typical of pain - screaming, crying, swearing, writhing - as well as the damage to the injured part of Fred's body - torn skin, crushed bone, blood and the like. If you happen to have the requisite equipment at hand - such as an fMRI scanner - you might even be able to observe the relevant goings-on in Fred's central nervous system. All of this is as directly accessible to you as it is to Fred. But Fred's sensation of pain the experience of it, the feel of it - is something only he knows directly, from the inside. If you know it is there, it is-only because you infer, from your own experience of what happens when you get your hand caught in a door, that Fred must be in pain. It is even possible that Fred doesn't really feel any pain at all: perhaps he is just an extremely eccentric prankster willing to break a hand in order to raise a laugh, and had earlier injected it with Novocain and is now only acting as if he feels pain. This is unlikely, but the fact that it is at least possible underlines the point that the pain itself - as distinct from its causes and effects, and the bodily damage associated with it - is not directly knowable to anyone but the person experiencing it.

what is true of pain is true also of other experiences. If someone flashes a camera bulb in your face, others might see you blink, wince and throw your arms up reflexively in response, but they will not, and cannot, see the after-image that subsequently occupies your visual field for a few moments. If you form a mental image of the Eiffel Tower, or

think of the way your favorite- song goes, others will be utterly unable to see that image or hear that song, however vivid the images are and however close they get their eyes and ears to your skull. Performing brain surgery on you won't give them access either - it's not as if they'll see a little picture of the Eiffel Tower inscribed in your grey matter or hear music coming from your hypothalamus. Nor can others directly experience what you experience as you eat a cheeseburger. Your sensations of the taste, texture, smell and look of the thing are available only to you; they can have similar experiences, should they eat their own burgers, but their experiences would then be theirs, not yours.

The feeling of pain, the look of an after-image, the taste of a cheese burger, and so on - those aspects of experience we've labeled qualia -thus exhibit a feature that philosophers call privacy, a feature that seems to set them apart from physical reality. Physical objects and properties are "public," in the sense that they can, in principle, be directly accessed, via perception, by any observer. This is as true of the brain and body as of any other physical phenomenon: just as anyone is as capable of peering inside and examining the workings of your car as you are, so too is anyone capable of opening up your body or brain and examining their workings. But your qualia are directly accessible only to you, via your introspection of your mind's contents - you have "privileged access" to them, that no one else has or can have. Everything else in the world is objective, knowable "from the outside" or from the "third-person" point of view; qualia - indeed, mental states and processes in general- are subjective, knowable "from the inside," from the "first-person" point of view. But then it seems that these mental states and processes must be different from anything occurring in the brain, body, or any other physical thing.

Finally, physical objects and processes are not only "public" rather than "private", and intrinsically devoid of color, odor, taste, and the like, but they are also intrinsically without meaning or intentionality. Even the words you're now reading are in themselves just meaningless squiggles of ink on paper; what meaning they have is meaning we give them, by interpreting them as having meaning. The same goes for the noises made by a tape recorder or the electronic impulses generating images on a computer screen. Intrinsically there is nothing there but sound-waves and electrical current, as devoid of significance as the sound-waves generated by a fan or the electrical

current passing through the fan's motor. The reason the former have any meaning at all is, again, that we interpret them as having it - we interpret the sounds made by the recorder and the images on the screen as words rather than merely noises and shapes. So, it seems that physical objects and processes have meaning only when they derive it from minds, which have it intrinsically. This is as true of brain processes as of any other physical process - in themselves, the electrochemical signals passing between neurons surely have no more meaning or intentionality than the electrical current passing through the wires and motor of an electric fan. So, again, the mind seems just obviously different from the brain.

The Indivisibility argument

A further difference between mind and matter, which Descartes took to have considerable significance, concerns the notion of divisibility into parts. A physical object is divisible - into halves, quarters, and so on, ultimately into its constituent molecules, atoms, and subatomic particles - and the smaller objects that remain after each division are themselves physical. As with the other features of physical objects we've noted, this is no less true of the human body and brain. But a mind is simple, not composed of parts and thus not divisible into further, smaller units. By this Descartes doesn't mean that we can't distinguish various aspects of the mind - its distinct capacities for reason, sensation, emotion and so forth - but rather that these aspects are, unlike the aspects of a physical object, aspects of a kind of thing that cannot be divided into further things of the same kind. You can divide a material thing into parts which are still themselves material, but you cannot divide a mind into parts which are still themselves minds. In that case, Descartes argues, the mind cannot be identified with any material thing, including the body or brain. Furthermore, it seems to follow that the immaterial substance of the mind is, unlike the body, immortal. Physical things can perish precisely because they are composite, and can thus be broken down into their constituent Parts. The mind, being simple, has no parts to be broken down into.

Descartes's conviction that the mind is a simple substance no doubt stems in part from the cogito argument described in chapter 1. In knowing for certain that "I think", what I know to exist is precisely a single thinking thing - after all, "I think" not "we think." I do not know

for certain, at least not initially, that there is any other thinking thing in the world; I can certainly coherently imagine that there isn't, that I alone exist, as in solipsism. But this thinking "I" just is my mind; in imagining it alone existing, I am imagining that a single mind exists, not a composite of smaller minds. Surely, then, I am imagining something simple. Consider further that when I wonder whether my body exists, I can do so in stages - I can imagine first that my torso and head are real, but my limbs a mere hallucination, and then imagine that my torso too is hallucinatory and so forth. I can inquire into the existence of my body part by part. But the same isn't true of my mind, the 'I' that thinks about its own existence. I either exist or I don't: it's all or nothing, not a matter of degree. Thus, the thing whose existence I'm concerned with seems clearly to be a simple, non-composite entity.

It is, nevertheless, sometimes suggested that modern psychological and neurological research have demonstrated that Descartes was wrong about the mind's simplicity. There are famous cases of "multiple personality disorder" (MPD), wherein a single mind seems to have fragmented into several personalities. Wouldn't this involve a mind being divided into smaller minds? There is also the odd behavior of "split-brain" patients, in whom the corpus callosum - the thick bundle of neurons connecting the two halves of the brain - has been severed. Such patients are claimed by some researchers to behave as if there were two people living in the same body, each controlling one half of it: for instance, one of a patient's hands will attempt slowly to stack blocks while the other moves in, as if impatiently, to stack them more quickly, only to be pushed aside by the first hand. Again, it would appear that what was once a single mind has divided into two.

But appearances, as we've seen, can be deceiving. In MPD, we have a phenomenon that was traditionally categorized as demonic possession. Accordingly, people exhibiting the behavior now associated with MPD described it, not as a fragmentation of a single mind into multiple ones, but as the entrance from without of a distinct and alien mind. If anything like this sort of description is correct, these cases would not count as evidence against Descartes's view at all, for they would not involve the division of a mind into smaller units, but rather the control over a single body of two distinct and otherwise unrelated minds. of course, few philosophers these days would take seriously the suggestion that demonic possession is the best explanation

of cases of so-called MPD (though this is largely because of the materialist worldview most of them presuppose, which is itself precisely what is in question in arguments for dualism). In any case, the possibility does at least show that MPD cases by themselves do not entail that the mind is divisible. Such cases need interpretation, and interpretations can reflect philosophical biases as much as philosophical conclusions.

This brings us to a more fundamental response to the MPD objection (and a more crucial one, since dualists will be much better off if they needn't resort to something as controversial as the notion of demonic possession). The reality is that it simply isn't clear that MPD cases (which are extremely rare and difficult to confirm) really are, in the first place, cases of multiple minds existing in one body. Many well-known cases of alleged MPD - such as that of "Sybil," made famous in the film of that title - have been shown to have been exaggerations or even hoaxes. "Sybil" herself has admitted that her "disorder" was more or less her own invention, that she was coaxed into believing that she had multiple personalities by therapists eager to prove that MPD was real, and that under their encouragement and in an emotionally fragile state she had manufactured and acted out various "personalities" to confirm their diagnosis. Many other MPD patients, emotionally disturbed people to start with, acknowledge that they see themselves less as literally "fragmenting" into different personalities than as fantasizing and acting out different roles again, often under the influence of overzealous therapists.

The behavior of "split-brain" patients is no less subject to interpretation, interpretation that can reflect the enthusiastic theorizing of the researcher as much as the objective facts. To begin with, the two hemispheres of the brains of such patients are not completely disconnected - there are other connections between the halves that remain undisturbed, and thus there are no grounds for insisting that the halves must be associated with different "minds." Furthermore, under ordinary conditions, such patients behave more or less normally, or at least not in any way that suggests that more than a single mind occupies their bodies. It is only in contrived experimental contexts that they can be made to exhibit remarkable behavior, and even then that behavior is by no means obviously best interpreted as involving a "division" of the mind. Many researchers hold instead that such behavior, when

examined carefully, amounts to little more than a variation on the awkwardness, failure of co-ordination, or general cognitive malfunctioning that can result from any serious injury to the brain, or an exaggeration of the absent-mindedness or incoherence that we all exhibit from time to time.

The "indivisibility" argument remains controversial, but since the evidence of the mind's divisibility is inconclusive, it seems the argument hasn't decisively been refuted.

The conceivability argument

We will return to the issue of the mind's simplicity and the plausibility of Descartes's indivisibility argument, when we consider the unity of conscious experience in chapter 5. Let us turn now to what many philosophers regard as the paradigmatic argument for dualism: the "conceivability argument." Dualism says that the mind is a different thing from the body or brain and can, in principle, exist apart from them; the opponent of dualism says otherwise, holding that the mind just is the brain, or at least that it necessarily depends on it for its existence (an alternative way of formulating the opponent's view to which we'll return in chapter 3). But to make such a claim commits the opponent of dualism to certain implications - implications which, the conceivability argument tries to show, are false, so that the claim that the mind and brain are identical must also be false.

Properly to understand the argument, we need first to understand a distinction philosophers make between different kinds of possibility and impossibility. When we say that it is impossible for a human being to run a mile in two minutes or to jump fifty feet, what we mean is that such feats go beyond the limits set by human physiology and the laws of physics. Such things are impossible given the way the world works; they are, we might say, physically impossible (or, what amounts to the same thing, we might say that it is a matter of physical necessity that no one can run a two-minute mile, etc.). But they are not impossible in the same way in which it is impossible for a square to be circular, or for $2 + 2$ to equal 5. Had the muscles of the human body or the gravitational pull of the earth been different, a two-minute mile or fifty-foot high jump may well have been possible. They aren't, given the way the world happens to work, but they would have been, had the world

worked in some other way. But no matter how different the human body, gravity, or the laws of physics may have been, there just couldn't have been such a thing as a circular square, and it couldn't have been true that $2 + 2 = 5$. These things would be impossible no matter how different the world might have been. They are, we might say, not just physically but metaphysically impossible (or, in other words, it is a matter of metaphysical necessity that they cannot obtain). They are impossible not only in the actual, but in any possible world.

How do we know this? In the case of running a two-minute mile, even though we know such things to be impossible in the real world, we can give a coherent description of how things might have been different in such a way that they would be possible. We can, if we care to, describe in detail what the gravitational force of the earth, a human being's musculature and lung capacity, etc. would have to be like in order for one to run a two-minute mile. We can give a description of such a state of affairs in a way that involves no contradiction, and thus what we would be describing is, though not physically possible – not allowed by the laws of nature obtaining in the actual world – nevertheless metaphysically possible allowed by the laws of nature in some other possible world. But we can do no such thing where circular squares and the like are concerned. A world where squares are circular and $2 + 2 = 5$ cannot be coherently described; the very attempt to describe it involves a contradiction. So there can be no such world. We might sum this up by saying that metaphysically impossible worlds, like a world with circular squares, are strictly inconceivable—we cannot even imagine the existence of such a world, for the attempt to do so involves a contradiction. By the same token, though, the fact that we can conceive of worlds where a two-minute mile is possible is reason to believe such worlds are not metaphysically impossible.

Suppose that we're considering a claim, not about two-minute miles or circular squares, but about identity. That is, suppose we're considering a claim of the form $A = B$, for instance, the claim that water = H_2O . We know that water is H_2O in the actual world, of course; it is physically impossible for something to be water without being H_2O . But is it metaphysically impossible too? Couldn't there be another possible world where water is not H_2O but something else? It seems that in the nature of the case, this is not possible. Water and H_2O are the same thing, so how could you have the one without the other? If

you could, wouldn't that show that they aren't really the same thing after all? If I could even conceivably have some water without having any H₂O, or you could have some H₂O without having any water, wouldn't this entail that water and H₂O are really just different substances?

This suggests the following principle: for any A and any B, if A = B, it is metaphysically (not just physically) impossible to have A without B (with qualifications I'll explain later on). But then, given what I've said above, it should also be impossible to give a coherent description of a world where A exists without B: A existing without B should be inconceivable. A corollary of this is that if it is metaphysically possible to have A without B, then A and B can't really be identical after all; and this means in turn that if it is conceivable for A to exist apart from B -if we can give a coherent description of A existing apart from B – then A and B just aren't identical. This gives us a way to test identity claims. If someone claims that a certain A is identical with a certain B, then we should see whether we can coherently conceive of A existing apart from B. If we cannot, this would not prove that they are identical maybe we just haven't thought about the matter carefully enough; but if we can conceive of it, this would surely give us reason to believe they are not identical.

Consider the claim that the mind is identical to the brain. If this is true, then it should be, not just physically, but metaphysically impossible for the mind to exist apart from the brain. And thus, if what we've said so far is correct, it should also be inconceivable for the mind to exist apart from the brain: we should be unable to describe coherently, in manner involving no contradiction, a situation where a mind but no brain exists. Can we conceive of such a situation?

We already have, in chapter 1. Descartes argued that it was impossible for him not to exist as long as he was thinking that he did, or thinking anything at all; nevertheless, it was still at least possible that his body, including his brain, did not exist, because those things might just be part of a hallucination put into his mind by an evil spirit. That is to say, it is entirely conceivable that one could exist as a disembodied mind, with one's body and brain, and indeed the entire physical world, being nothing but a figment of one's imagination. But then it is conceivable and therefore at least metaphysically possible for the mind to exist apart from the brain. Therefore, the mind is not identical to the brain.

Lest one think that this crucially depends on the possibility of there being a Cartesian demon - which would itself be a disembodied mind so that the argument might appear to beg the very question at issue – it should be noted that the same point could be made in terms of solipsism, the scenario in which "I alone exist" as a disembodied mind, with nothing, neither a demon nor a physical body, existing apart from my mind and its hallucinations. Or we can appeal to the sort of scenario vividly described by the dualist philosopher W. D. Hart. Imagine waking up one day and staggering groggily to the bathroom sink to splash some water on your face. As you gaze into the mirror, you notice, to your great horror, that where normally there would be two eyes staring back at you, you see instead two dark, and vacant eye sockets - with the eyeballs completely missing! Frantic, you reach into the sockets to verify that they are empty, and, sure enough, feel nothing but the stumps of the optic nerves. This would, of course, be impossible in real life. But you can certainly conceive of it happening, without contradiction - you can vividly imagine having an unsettling experience of this sort, in a way that you cannot conceive of a circular square or $2 + 2$ adding up to 5. If you can conceive of this, you can also conceive that, being intrigued by your ability to see without eyeballs, and wondering if any other vision-related body parts are missing, you get out a hacksaw and carefully remove the top of your skull, only to reveal an empty cavity where your brain should be. Now you've conceived, in a nauseatingly vivid fashion, of seeing without either eyeballs or a brain. And if that's conceivable, you can take the next step and imagine that instead of seeing empty eye sockets staring back at you, what you see is your own headless body - in which case you'd be conceiving of seeing without a head. Finally following this exercise in conception to its logical conclusion, you can imagine that what you see in the mirror is not even a headless body, but nothing more than the wall behind you and no body at all. Wondering whether someone has installed a trick mirror or if you've become a vampire, you look down at your torso, arms, and legs but find that you still can't see them, only the floor under you; nor can you feel them, as you realize that your attempt to touch them has failed - there's nothing there to touch! You would now be conceiving of seeing without a body. But seeing is a mental process, as is the frenzied thinking you'd now be engaged in; which means that what you've conceived of is your mind existing apart from a body or

brain. So, again, it's conceivable that the mind exists apart from the brain - in which case they're not identical.

This argument has, as one would imagine, been subject to a lot of criticism. However, some seemingly obvious criticisms simply miss the point of the argument. It is no good, for example, to object that merely conceiving of something can't make it happen in the real world – I can't make myself fly merely by imagining that I can. That's not what the argument is saying. The claim, remember, isn't that being able to conceive of something makes it physically possible, but rather that it shows that it is metaphysically possible. It may not be, given the way the actual world works, but it could have been, had the world been different. Someone might then object that this point is trivial, since anything could have been possible in that sense. But as we've seen, this isn't so: circular squares and $2 + 2$ equalling 5 would not have been possible no matter how different the world might have been; they are absolutely and metaphysically impossible, because they involve contradictions, as running a two-minute mile and existing without a body do not. It might then be insisted that the claim is still trivial, for what needs to be shown is that the mind could exist without the body in the actual world, not merely in some conceivable one. But this too misses the point. For, as with water and H_2O , minds and brains, if identical at all, must be identical in every possible, and thus every conceivable, world. If it is even conceivable that a mind could exist without a brain, then mind and brain can't be the same thing - how could they be, if one could conceivably exist apart from the other? The point is related to the "apples and oranges" argument: you could have apples without oranges, so obviously apples and oranges aren't the same thing. You could also have minds without brains, so obviously they aren't the same thing either. This holds true even if, in the actual world, minds typically are associated with brains - something no dualist denies. Where there's smoke, there's fire, but obviously smoke and fire aren't the same thing. Creatures with hearts are always creatures with kidneys, but obviously hearts and kidneys aren't the same thing. Minds are typically associated with brains, but that doesn't mean they are the same thing.

There are more serious objections, however. The principle that conceivability entails metaphysical possibility, though endorsed in some form or other by philosophers of the stature of Descartes and

David Hume (1711-1776), is often challenged by contemporary philosophers (though usually it should be noted, precisely as a way of avoiding commitment to dualism, rather than for independent philosophical reasons). Take the fact that Neil Armstrong is identical to the first man to walk on the moon. since this is a fact, it is presumably metaphysically impossible for Armstrong to exist apart from the first man to walk on the moon - they're the same person, after all. But isn't it nevertheless conceivable that Armstrong could have failed to be the first? Can't we just obviously imagine a case where the Soviets beat the Americans to the moon and Yuri Gagarin got to leave his boot prints there instead? sometimes even the water/H₂O case is put forward as a counter-example. True, it is said, it is metaphysically impossible to have water without H₂O, since they are the same thing. But isn't it in fact, and contrary to my earlier suggestion, at least conceivable that water could exist apart from H₂O? Can we not coherently imagine a situation in which we have a substance that is clear, liquid, and quenches thirst, freezes and turns to gas at the same temperatures that water does, yet does not have the chemical composition of H₂O but instead turns out to have the composition X_yZ? wouldn't this just be to conceive of water existing apart from H₂O? But if it is conceivable that water could exist apart from H₂O, or that someone other than Armstrong could have been the first to walk on the moon, even though it is metaphysically impossible for water to exist apart from H₂O or for Armstrong to exist apart from the first moonwalker, then the principle that conceivability entails metaphysical possibility must be false. It follows that the fact that we can conceive of the mind existing apart from the body does not show that this is metaphysically possible.

Formidable as these examples might seem, they do nothing to undermine the main thrust of the conceivability argument, for reasons made clear by the influential work of the philosopher and logician Saul Kripke. As Kripke has argued, strictly speaking it is identity statements involving what he calls rigid designators that are, if true at all, true of metaphysical necessity, that is, whose falsehood is metaphysically impossible. A rigid designator is an expression that denotes the same thing in every possible world, in every possible way that things might have been. "Water" is an example, as is any term designating a "natural kind" or naturally occurring substance such as gold or iron. 'Water' in essence designates: whatever substance it is in the actual world that has

the properties of liquidity, quenching thirst, freezing and turning to gas at such and such temperatures, etc. Thus "water" also designates whatever it is in any other possible world that fits this precise description, namely, the description of being the substance that has those Properties in the actual world. "H₂O" essentially designates; the substance having specifically such-and-such a chemical composition.'H₂O" thus also designates whatever it is in any other possible world that has that specific chemical composition. We know empirically that the substance in the actual world that is liquid quenches thirst, etc. is exactly the same as the substance having specifically such-and-such a chemical composition. Water, in the actual world, is H₂O. But since 'water" also designates whatever the substance is in any other possible world that in the actual world is the substance that is liquid, quencha thirst, freezes and turns to gas at such-and-such temperatures, etc., and that latter Substance is H₂O (where "H₂O" designates whatever it is in any possible world - including the actual one - that is the substance having specifically such-and-such a chemical composition), it follows that "water" and "H₂O" will refer to the same substance in every possible world. That is, water and H₂O are identical in every possible world.

When we think carefully about the semantics of terms like "water" and 'H₂O," then, we will see that we really can't coherently describe or conceive of a world where water isn't H₂O. When we think we're conceiving of such a world, what we're really conceiving of is a world where there is a substance that is liquid, quenches thirst, freezes and turns to gas at such-and-such temperatures, etc. that turns out to have a chemical composition of X_yZ. But precisely because this substance wouldn't thereby be the substance in the actual world that has these properties, it wouldn't be water, but merely a substance very similar to water to conceive of a substance similar to water that is not H₂O is not , the same thing as to conceive of water existing apart from H₂O. So the water/H₂O case just isn't at all a counter-example to the principle that conceivability entails metaphysical possibility.

What about the Neil Armstrong example? We can indeed coherently conceive of a situation where Armstrong is not identical to the first man to walk on the moon, but this would nevertheless not, on Kripke's analysis, be a counter-example to the principle that conceivability entails metaphical possibility. For it is not

metaphysically impossible for Armstrong to fail to be the first to walk on the moon, even though the identity statement "Armstrong is identical to the first man to walk on the moon" is true. The reason is that at least one of the expressions in this statement is not a rigid designator, namely the expression "the first man to walk on the moon." This expression does not mean "the specific person who, in the actual world, first walked on the moon," but rather merely something like "whichever Person turns out to be the first to walk on the moon." And of course it is metaphysically possible that someone other than Armstrong could have turned out to be that person. So we shouldn't be surprised that it is also conceivable. As long as we note carefully, along Kripkean lines, that it is only identity statements involving rigid designators which, if they are true at all, cannot possibly be false, we will see that there are no genuine counter examples to the principle that conceivability implies metaphysical possibility.

That principle seems highly plausible in any case. Indeed, it is hard to see how even its critics could themselves regard anything as metaphysically possible in the first place, without being implicitly committed to the principle. For why does anyone accept that it is at least metaphysically possible to run a two-minute mile or high jump fifty feet if not on the basis of the fact that one can clearly conceive of this happening, or give a coherent description of it? That is not to say that anything anyone says he or she can conceive is thereby truly conceivable and therefore metaphysically possible; as we've seen, sometimes what someone thinks is conceivable turns out on reflection not to be conceivable after all. This might result not only from a failure to take note of the role of rigid designators in identity statements, but also from the commission of such fallacies as confusing a word for the object named by the word or from a failure to pay careful attention to the precise meaning of a word. (For example, someone might claim that he or she can conceive of a circular square, when in fact all the person is really conceiving of is a circle he or she is calling "a square" or a shape that isn't truly a square at all, but has three straight sides and one round one.) But when we've been careful to avoid such fallacies and find that we still seem capable of conceiving of a certain state of affairs, we surely have strong reason to believe that state of affairs is metaphysically possible.

The Interaction problem

The principle that conceivability entails metaphysical possibility is, then, eminently defensible. But there is another way to challenge the conceivability argument. One could simply deny that it really is conceivable in the first place for the mind to exist apart from the brain. That is, one could argue that, just as someone who thinks it is possible to conceive of water apart from H₂O is mistaken, and just hasn't really thought carefully enough about what he or she claims to be conceiving, so too is someone who thinks it is possible to conceive of the mind existing apart from the body simply mistaken, and will see, on further reflection, that this isn't really what he or she has conceived of at all.

Along these lines, one might assume that the Kripkean framework we've appealed to in defense of one premise of the conceivability argument (the premise that conceivability entails possibility) might be applied here too, in opposition to another premise of the argument (the premise that we can conceive of the mind existing apart from the body). But Kripke himself would disagree. Expressions referring to mental states and brain states are in his view both rigid designators. "The firing of C-fibers," designates: whatever it is that in the actual world is a brain process of such-and-such a type, and "pain", designates: that mental state that has such-and-such a feel. So if pain is identical to the firing of C-fibers (and, by extension, if the mind in general is identical to the brain), then they must be identical in every possible world, as a matter of metaphysical necessity. As what we have already said implies, it appears we can conceive of a possible world where pain exists in a disembodied mind, apart from the firing of any C-fibers or any other brain state; and thus it would follow they can't be identical. It might seem at first as if one could get around this argument the way we saw the dualist can get around the purported water/H₂O counter example. In fact, there is a crucial difference between that case and this one. In the water/H₂O case, we saw that something could be liquid, thirst-quenching, liable to freeze and turn to gas at such-and-such temperatures, etc. (that is, it could have many of the Properties that water does) without being water. So to conceive of such a substance apart from H₂O is not to conceive of water apart from H₂O. But nothing could have the feel that pain does without being pain, for pain is nothing more than that feel itself. So to conceive of something that feels like

pain existing apart from any brain state just is to conceive of pain itself existing apart from any brain state. In the case of the conceivability argument, unlike the case of water and H₂O, what we think we're conceiving of really is what we're conceiving of. Thus, there seems no way for the critic of dualism to appeal to Kripkean semantics to respond to the conceivability argument.

There is another way for the opponent of dualism to press this sort of objection. In the previous chapter we examined the view that the mind is, in perception, only indirectly aware of the external, physical world, with this indirect awareness mediated by a causal connection between the mind and the things it perceives. But let's consider this causal element in perception more carefully. It seems clearly to be a necessary part of your perception of the book you're now reading that you have some causal connection to it, that the book itself is what is causing you to experience it. Obviously, if there were in fact no book there - if you were merely hallucinating, because someone had slipped drugs in your coffee - then you wouldn't really be seeing it at all, but only seeming to see it. But even if there were a book there, that wouldn't by itself be enough. For suppose you were right now having such a hallucination, your brain malfunctioning and your mind totally cut off from the outside world, and suppose also that, just by chance, someone has put a copy of this book down on the table in front of you. Would you really be seeing the book? Surely not, because even though you're having an experience of seeing a book, and there really is a book there, the book itself is not what's causing the experience - the drugs are causing it. So truly to see the book, not only do you have to have the experience of seeing it, and not only does the book have to be there, but the book must be what's causing the experience.

With this in mind, says the critic of the conceivability argument, examples like the "seeing without a body" scenario take on a new complexion. For if we're really to conceive that we are seeing without a body, it follows that we must also conceive that there is a *causal connection* between our mind and the things we are seeing. But it is hard to see how we can conceive of this. In normal cases of perception, we know that what occurs is something like this: light bounced off an object is reflected off the mirror, and travels in the form of photons to your eyeballs, where, the retina being stimulated, a series of complicated neural signals is initiated which results in the experience

of seeing the object. But what happens in the "seeing without a body" example? Light bounced off the object is reflected off the mirror, and travels in the form of photons to ... where, exactly? There are no eyeballs there for it to enter, indeed nobody at all for it to travel to. So where does it go? It's no good to say that it goes to the mind, for on Descartes's view, remember, the mind is outside space and has no physical properties whatsoever - no shape, mass, length, width, or height at all. So how can the light, which is physical, possibly get in "contact" with it? It seems just impossible that it could. But if it can't get in contact with it, then there can be no causal connection between a non-physical mind and the physical objects outside it; which entails that the mind couldn't truly see or perceive such objects without a body. But then it turns out that we really can't conceive of seeing without a body after all. If it seems that we can, that's only because we haven't thought carefully enough about what's involved in seeing something.

Strictly speaking, this objection doesn't quite undermine the conceivability argument, for that argument requires not that we can conceive of "seeing without a body" specifically, but only that we can conceive of the mind existing apart from the body in some fashion or other. Even if we accept the criticism that the causal conditions necessary for true seeing to occur entail that one cannot genuinely conceive of "seeing without a body," we can, nevertheless, insist that it is still possible to conceive of being a disembodied mind which seems to see - that it is possible to conceive, as in Descartes's evil spirit scenario or in solipsism, of being a disembodied mind which has a stream of hallucinatory visual experiences. Obviously those experiences wouldn't truly count as literal seeing *per se*, for there would be no causal contact with the external physical world. But even hallucinatory experiences are still experiences, and to imagine having them while disembodied is still to imagine the mind existing apart from the body. So the gist of the conceivability argument still stands. The dualist could accept that the mind cannot literally see or in general perceive the world of physical objects unless it is joined to a body; cut off from a body, it becomes, as it were, trapped within itself. But that just means the mind needs the body in order to do anything other than merely hallucinate; it doesn't mean it is identical to the body or any part of it.

Even if the dualist can in this way defend the conceivability argument for dualism against the objection under consideration, that objection still raises questions about dualism itself. Perhaps we needn't claim that we can conceive of the mind and the physical world interacting in order to get the conceivability argument off the ground; it is enough to conceive of the mind existing all by itself totally cut off from the physical world. But the dualist also wants to hold that the mind, though distinct from the brain and body, nevertheless does interact with them. And just as it is hard to see how photons could get into causal contact with a disembodied Cartesian mind, so too it is hard to see how the brain and body could either. The brain is an extended object like any other, with a mass, shape, and particular location in space, while the mind is, in Descartes's view, none of these things. So how can the mind and brain possibly interact? Of course, it seems obvious that they do; the problem is that dualism appears to have no way of explaining how this is possible.

The "interaction problem" has been the main difficulty facing dualism since the time of Descartes and various solutions have been suggested. One of them, known as occasionalism, holds that God serves as the link between mind and brain: observing that light reflected from the cheeseburger has impacted your retinas and set up a series of neural firing patterns in your brain, God causes your mind to have an experience of seeing the burger; observing that experience has led you to decide to eat the burger, he then causes a set of neural firing patterns to occur in your brain that result in you picking up the burger, putting it in your mouth and eating it. Parallelism holds, alternatively, that the mind and the brain are not linked even in this indirect fashion. Rather, they are simply so constructed that the events occurring in the one are always exactly appropriate to the events occurring in the other, yet without having any mutual influence: the brain and body are so ordered that light reflected from a cheeseburger results in certain neural firing patterns, which results in the body's limbs moving it toward the burger, while the mind is so ordered that, at precisely the same time that sequence of events is occurring in the body, it undergoes a parallel series namely, it has the experience of seeing a cheeseburger; which results in a desire for the cheeseburger, which results in an intention to go pick it up. Mind and body are like two clocks operating entirely independently, but keeping up with each other so perfectly that it seems

that there is interaction between them. There is a "pre-established harmony" between them - pre-established by God, who is responsible for having wound up the clock of mind and body in the first place.

It is easy to scoff at such theories if one simply takes for granted the general materialist world picture. But if one believes, as proponents of these theories do, that there is already independent evidence for the existence of God as well as for the distinction between mind and body, it is hardly unreasonable to suggest that God might have something to do with the connection (or apparent connection) between mental and material substance. As in so many other cases in philosophy, what one regards as a plausible theory is largely determined by the background assumptions entailed by one's general metaphysical commitments. Still, it is always preferable, if possible, to avoid having to defend one controversial position by appealing to another which is at least as controversial and to avoid contradicting common sense – something these theories clearly do, denying as they do that there really is a direct causal connection between mind and body.

Another, and more widely accepted theory, which only partially denies this is epiphenomenalism, which holds that events in the brain and body produce events in the mind, but that those mental events in turn have no causal influence on what happens in the brain and body. They are mere "epiphenomena," in effectual by-products of the operation of the physical processes of the brain. The light striking your retinas causes you to have the experience of seeing the cheeseburger' and further brain events cause you to form the desire to eat it but that desire itself is not what causes you to proceed to eat it. The experience, the desire, and everything else that goes on in your mind have no effects at all; what causes your action are just further, purely material, unconscious brain Processes. The appeal of this theory is partly that it does not, as occasionalism and parallelism do, appeal to anything as controversial as the existence of God, and partly that it is consistent with the notion that bodily behavior can be entirely explained in terms of processes occurring in the brain and nervous system - a notion that has gained widespread acceptance following the rise of modern neuroscience. Epiphenomenalists hold, as opponents of-dualism do, that we can completely explain human behavior by appealing to such physical bodily processes; there is thus no need to try to explain how immaterial mental processes interact with the body, for they don't. They

also hold, however, as dualists do, that mental processes are non-physical. Epiphenomenalism thus constitutes a kind of compromise between dualists and their opponents.

It is a notoriously unsatisfying compromise, however. Occasionalism and parallelism may deny common sense in holding that mind and body have no direct effects on one another, but at least this denial serves the purpose of solving the interaction problem, and at least they provide some explanation of why mind and body seem to interact. Epiphenomenalism, in denying at least that the mind has any effect on the body, also defies common sense, but it fails to compensate by providing any explanation in return of how the body can (as the theory claims) have an effect on the mind. Worse still, epiphenomenalism makes mysterious how we can even so much as talk about the mind. Presumably, for our written and spoken words to refer to the mind, they have in some sense to be the effects of what is going on in it. But the mind has no effects at all, in the epiphenomenalist view. So how are we able to talk about it? How are epiphenomenalists able to tell you anything about the mind when, in their own view, the mind cannot have any effect whatsoever on what they say?

There is more to be said about the interaction problem, and it will be said in later chapters. Suffice it for now to make two points. First, the interaction problem by itself does nothing to undermine the arguments for dualism we have considered so far. Merely to note that the Cartesian concept of the mind leads to a mystery about how mind and body interact is not to uncover any fallacy in the conceivability, indivisibility, or apples and oranges arguments. Dualists can, therefore, reasonably hold that so long as the arguments for their position have not been proved fallacious, they are in their rights in continuing to maintain that position - while also, of course, continuing to look for ways to solve the interaction problem. Dualism is in this respect really no worse off than those most fundamental theories of modern physics: relativity and quantum mechanics. Notoriously there are respects in which these theories seem to be in conflict, and yet the evidence for each is very powerful. There are various ways of trying to reconcile them, but as yet no consensus as to which, if any, is the right one. It would be silly to insist that physicists must reject these theories, or at least one of them, until some generally accepted solution to the problem of reconciling them has been worked out. Physicists must continue the

search for a scheme that will unify quantum mechanics and relativity, but there is no reason for them simply to ignore the strong considerations that favor these theories until such time as that unifying scheme has been arrived at. Similarly, it is unreasonable to expect the dualist to give up dualism simply because the interaction problem exists, when there are arguments in favor of dualism that are at least as powerful and worthy of consideration as any others in philosophy.

Second, contemporary philosophers have nevertheless taken the interaction problem to be, at least, a strong motivation for seeking an alternative to dualism, and they have not necessarily been unreasonable in doing so. The mere fact that interaction between a nonphysical substance and a physical one is difficult to explain does not refute dualism. But the philosophers in question take the problem to go deeper. The difficulty in their view, is not merely that it is hard to see how a cause and effect relationship between such substances might work; it is that modern science seems to present us with a picture of the nexus of causes and effects in the physical world that leaves nothing for a non-physical substance to do. We are not in the position of failing to understand how such a substance can play the role it plays; we are rather in the position of failing to understand how it could even have a role to play in the first place. For the law of the conservation of energy entails that the amount of energy in the physical universe is constant. A Cartesian immaterial substance, being outside space, is outside this universe. For it to affect the physical world, and in particular the brain, it would have to introduce energy into the physical universe; and for the brain in turn to affect an immaterial substance, it would seemingly have to transfer energy out of the physical universe. Either way, the amount of energy in the physical universe would fail to be constant. So the very idea of causal interaction between Cartesian material and immaterial substances seems to violate the laws of physics.

Most contemporary philosophers have accordingly sought to develop a materialist conception of the mind in which it is, contrary to appearances, just another part of the physical world. More modest versions aim to show that such an alternative account will be at least as plausible as dualism, and equally capable of explaining the various aspects of our mental lives. The idea would then be that, though both dualism and materialism have strong arguments in their favor, materialism, being (allegedly) more in harmony with modern physics,

ought to be preferred. More ambitious materialists would go further than this and claim that a materialist conception of the mind will, when fully worked out, show dualist arguments to be not only inconclusive, but positively fallacious or incoherent.

The case for dualism, then, cannot fully be evaluated until it is compared with the case for materialism. If the materialist can indeed show that the various features of the mind can be accounted for in purely physical terms, then dualism will, at the very least, have much of the wind taken out of its sails. But if the materialist fails to do so, that failure will itself provide some further support for dualism - indeed, many of the most influential dualist arguments in recent philosophy are precisely attempts to undermine various arguments for materialism. And if there remains a mystery of how mind and matter can possibly interact, we will see that some dualists have argued that this reflects, not a problem with the dualist's conception of the mind, but rather a problem with the materialist's conception of the physical world.