
DUALIST THEORIES OF MIND AND BODY

Interactionism and Substance Dualism

Mind–body dualism was a popular view until roughly the 1960s, though it is less and less so these days, at least with professional philosophers. They have for the most part thrown in their lot with those scientists who have adopted a materialistic or naturalistic worldview—nature is all there is.

Dualism is the antinaturalist claim that the mind and the body are two separate and very different things. The two sorts are the *nonphysical* and the *physical*. The nonphysical sort of thing, the mind or soul, is not part of nature. “The mind is a nonphysical thing” was our first proposition, and “The body is a physical thing” the next. The essence of dualism is the claim that both these propositions are true, and that the mind is not part of nature.

Dualism is the anti-naturalist claim that the mind and the body are two separate and very different things. The two sorts are the *non-physical* and the *physical*.

In addition, one important form of dualism tells us that mind and body are distinct *things* that can exist independently of one another. Such independently existing things have been called “substances” in the history of philosophy. A substance is an individual thing that can exist by itself, independently of other substances. Accordingly, *substance dualism* is the view that mind and body are distinct in the sense that they can exist independently of each other, or are substances.

Interactionist substance dualism is the view that these two substances or things exist and can interact causally. So, for example, when the body takes in too much beer, the mind becomes confused, and one’s mood may change. Here the interactionist substance dualist will say that the physical substance or thing called “the body” is interacting, or certainly seeming to, with the nonphysical substance or thing called “the mind.”

The body can exist without the mind, after burial. But what about the other way round? We can imagine the mind existing in the darkness after death, just as it exists in the darkness after bedtime. Just as vividly as we are aware of our mind in the darkness after lights-out, perhaps working on little mathematics problems, or perhaps saying its prayers, or thinking about this or that, so we can imagine activity of these sorts continuing and going on after we die. To some this is a comforting thought, to others unnerving and alien.

There is also a thought experiment that we can perform that is suggestive of dualism. Imagine that I wake up, as usual, and open my eyes, or think I do. To my left I see my cup of coffee in a clean white mug, steaming a bit and smelling good. "Great," I say to myself, "time for coffee." I glance down the bed, and I am surprised, because it looks unrumpled and perfectly made, as it was the night before. Things begin to get even odder when I notice that where my feet should be sticking up under the covers, the cover is completely flat. The next odd thing I see is that my torso also does not make a lump under the covers. When my wife pulls back the covers and asks whether she can hand me the coffee, I become most alarmed: I see no body at all where my body should be. Is this a nightmare? No, I am fully awake, but my body seems to have vanished in the night. It is not merely that it is invisible. It simply isn't there. It has disappeared. Perhaps it no longer exists. Have I turned into a pure consciousness? What philosophers call my mind or consciousness (though these most certainly are not the same thing), including my thoughts and visual and tactile sensations, and all the other sensations, of motion and action, is still there, unchanged. I still have the chronic feeling of pain in my back, where my back ought to be, but seem to be missing the back itself. What am I supposed to think? It seems natural to say that my body has gone, but that my mind is still there. I now see that

my mind and my body are distinct, then, for my mind can exist without my body.

I can imagine all this; and, more importantly, it is all possible, in the sense that a story of waking up without a body does not seem to be a contradictory story. Freedom from contradiction, rather than imaginability, is the proper test of possibility. If there is no contradiction in the description of an event, then the event is possible. Suppose that it is possible that I shall win the lottery. I can imagine that I shall win it, but that is not the important thing. The important thing is that I can describe myself winning tomorrow, going to the office of the lottery, presenting my lottery ticket, picking up my winnings, and so on, and among the descriptions of all these happy events there is no contradiction. Imaginability *may* be an indication of describability, but it does not guarantee it, whereas describability—in the sense of description genuinely free from contradiction—does demonstrate possibility.

By this test we should conclude that it *is* possible that the mind and body should exist without one another. It is possible that I should wake up with my mind and consciousness intact and my body gone. That possibility is the central claim of dualism. It is significant that the initial claim is not that they *do* exist without one another, since for the moment anyway they are somehow stuck together, but that they *can*. If they can exist independently, it does not follow that both do, or that they will not exist at the

same time. We can imagine that mind and body can exist independently of one another, but that at death both of them get destroyed at once, though by two different sets of forces, one physical and one nonphysical, assuming we can make sense of the idea of nonphysical forces.

The main difficulty for interactionism is one that stumped Descartes. *How* can the mind and body interact, if one is physical and therefore spatial, and the other is nonphysical and therefore nonspatial? Of course, it is possible to deny that the mind is nonphysical, and I will discuss this important option in the next chapter. But for the moment we are considering the view that the mind is nonphysical and the body is physical, and that the two interact. The question is, *how*? How can mind and body interact if one is physical and the other is not, given that physical and nonphysical things cannot interact? This is the objection to interactionist substance dualism made by Princess Elisabeth and Gassendi, and it is hard for interactionist substance dualism to meet it. Perhaps it is impossible.

There have been some contemporary attempts to make dualism work, but on the whole they have been a bit disappointing. E. J. Lowe, for example, argues for what he regards as a new picture of psychophysical dualist interactionism.¹ He notes that the structure of the causal chains of events is a branching structure, but since the chains get intertwined the structure as a whole has no ends. So mental events cannot interact or indeed fail to interact with

the tips and initiate causal action, since there *are* no tips! When I make to lift my arm, the tree structure as a whole is activated, or some significant part of, but it is as a result not of this but of the desire or wish or intention to move my arm that my arm moves. The activation of the neurophysiological causal tree explains the exact *way* in which the movement of my arm occurs, say, jerkily or smoothly, but it does not explain *that* it occurs in the first place. The tree “mediates” the relationship between mind and action. But mediation is a causal relationship. It remains true that what explains the fact that I raise my hand is the decision to raise it. So there is a direct action of the mental on the physical that still needs explaining. Lowe claims that my mind communicates not with the tips of the tree of causal events in the brain, since there are none, but with the whole tree, and explains the existence of the whole tree-structure of neurophysiological events. But the problem is just the same. How does the mind activate the whole tree? If interacting with the tips of the tree is impossible, so is interacting with the tree as a whole.

Another odd feature of this account is that the intention and the activation of the tree of events take place at the same time, the one responsible for the *how* of the movement, and the other responsible for the *fact* of it. This seems fishy.

Moreover, Lowe’s version of dualist interactionism also does not eliminate the charming “pairing problem”

that arises for dualism.² My mind issues the wish for my arm to rise, and the wish instructs my tree-structure of events to begin. Suppose you are standing next to me. How is it that my mind doesn't go into the wrong tree structure—yours—and not mine? Or how come it doesn't go into both, like a radio broadcast? If mind and body are genuinely distinct, then how is my mind paired with my brain and your mind with yours? Why does my arm rise and not yours?

Property Dualism

There is an answer to the pairing problem, but it means abandoning substance dualism. For dualists who are daunted by the problems facing substance dualism, another kind of dualism may seem to afford some relief. It is called *property dualism*.

The property dualist sees clearly the difficulties of two interacting but distinct substances, and proposes instead a dualism not of substances or *things* but of their *properties*. There is only one substance, says the property dualist, but it has two sorts of properties, physical and nonphysical. In one version of property dualism—the physicalist version—the mind is physical. It is the relevant part of the brain for causal interaction. But it has two sorts of properties.

Consider the fact that a piece of art such as a painting is physical, but it can be said to have nonphysical properties. Though it is made of paint and wood and canvas, the painting can be said to be: accurate; a bit of a caricature; witty; slightly derivative though stylistically effective; and a bit dark. These are aesthetic properties, not physical ones. But there are not two things or substances, a canvas *and* a work of art. If someone attacks a painting with a knife, as people sometimes do, then it might lose some or all of its aesthetic properties. And we cannot say that the higher-level properties of *being witty* or *being a bit dark* are identical with the paint and wood and canvas, though they are dependent on them.

Mental properties, such as having a thought, are *grounded* in the physical brain or mind, says the physicalist property dualist, but they are not themselves *reducible* to physical properties. If my brain is damaged, my capacity for thought can be impaired. But according to the property dualist, it does not follow that mental properties, such as my having a thought, are physical.

Now the property dualist is in a position to respond to the pairing problem that attacks substance dualism. Why do *my* mental activities, if detached from my body, not cause things to happen in *your* body? How do my mental activities reach the correct destination? Why is *this* mind connected with *this* body and not some other?

The property dualist denies that the mind and the body are distinct, since the mind is a physical thing. Mind and body can then interact nicely. Though physical and nonphysical substances cannot interact, the mind is not nonphysical; it is physical. But it *does* have nonphysical properties. These properties, however, do not themselves have effects on the body.

Everything seems to be in order. But there is a large fly in this ointment. Although it may be true that abstract triangles and aesthetic properties do not have actual effects, in the case of the mind, mental properties, such as thoughts and feelings, most certainly do initiate effects. Intent, premeditation, or *mens rea*, the “guilty mind” which is presumably something mental (since “mens” is just the Latin word for mind, and from which the English “mind” and “mental” are derived), is an element that is necessary to demonstrate certain classes of crimes, most notably murder. The evil intentions in the mind are taken to be the properties of the guilty mind that result in the unlawful deed. And the physicalist property dualist has no way to account for these.

Parallelism

There is another dualist possibility, however: parallelism. On this view, mind and body are distinct, but they do not

interact. We can accept dualism, including the proposition that the mind and body can exist independently of one another, along with the proposition that one is physical and the other is nonphysical, and at the same time reject the proposition that they interact. One good reason we can give is that “Physical and nonphysical things *cannot* interact,” and we have seen exactly why this is such an appealing proposition, starting with the arguments offered by Princess Elisabeth and Gassendi. We simply *deduce* “Mind and body cannot interact” from “Physical and nonphysical things cannot interact,” given the further premises that the mind is a nonphysical thing and that the body is a physical thing. We arrive at:

- (1) The mind is a nonphysical thing.
- (2) The body is a physical thing.
- (4) Physical and nonphysical things cannot interact.

As solutions go, this one is as logically appealing and as successful as any other. From (1), (2), and (4), it certainly does follow that mind and body *cannot* interact, and hence that they *do not* interact. But how then are we to account for the appearance that they *do* interact? It strains belief to suppose that they do not, one might think, because the fact of mind–body interaction is so common and familiar as to be undeniable. There is the effect of alcohol on mental state, for example, not to mention drugs of various

sorts. There is the effect of the mind on the body, most obviously of cases of intentional action, but also in cases, say, of a mental state such as blind rage leading to unhappy physical consequences. Psychiatry and psychology are full of examples of interactions, in both directions.

What does parallelism have to say to all this? It is perhaps surprising, but these examples represent no threat whatsoever to the parallelist view. The parallelist can simply assert that though there is no *interaction* between mind and body, there is a *correlation* between what happens in the body and what happens in the mind wherever we thought there was an interaction. The drinking of beer is *followed* by the fogginess of the mind, or correlated with it. And this is a well-established empirical fact that is neutral with respect to interactionism and parallelism. What we must not do, says the parallelist, is to imagine the body emptying beer into the mind, or, what is equally absurd, getting the neurons to fire into the mind, or in some literal sense sending physical messages directly into the mind, so that we have the ridiculous picture of electrical signals going off in the mind as well as in the body. We have no way at all of picturing such an event, as the mind is nonphysical and the signals are physical. We would then be imagining something that does *not* have a position in space (the mind) containing objects that *do* have positions in space (signals from the neurons). As A. J. Ayer has observed,

The physiologist's story is complete in itself. The characters that figure in it are nerve cells, electrical impulses, and so forth. It has no place for an entirely different cast, of sensations, thoughts, feelings, and the other *personae* of the mental play. ... Nor are there such temporal gaps in the procession of nervous impulses as would leave room for mental characters to intervene. In short, the two stories [mental and physical] will not mix. It is like trying to play *Hamlet*, not without the Prince of Denmark, but with Pericles, Prince of Tyre. Each is an interpretation of certain phenomena and they are connected by the fact that, in certain conditions, when one of them is true, the other is true also.³

The impossibility of physical and nonphysical things *interacting*, asserted in proposition (4) of the initial tetrad, does not prevent the *correlation* of the events within the physical body and the nonphysical mind. What the parallelist objects to is the idea that the electrical impulses or neural activity do in any literal sense sidle right up alongside the mind, and, from their close proximity, interact. There can be no literal proximity to the mind, if "literal proximity" is spatial contiguity.

In the history of the mind-body problem, parallelism arose partly as a result of a vivid awareness of the reasons for which Descartes's interactionist dualism could not

work. The *dualism* made the *interaction* impossible, and behold: the mind–body problem was born. The parallelists were committed to dualism. So what was left? Mind and body *could* not interact, just *because* dualism was true, they thought. But mind and body *do* seem to operate in tandem—synchronized, as it were. When the desire for coffee enters the mind, it is then that the body, or part of it, reaches out and picks up the cup of coffee. Then the mind says to itself, “Enough. No more,” and the body stops pouring the coffee into its throat, and puts the cup down. But why does it do it *then*, at exactly that moment? How has the mind *made* it come to pass that the body stops pouring coffee down its throat?

Causal interaction, said the parallelists, just *is* synchronization. The most celebrated and extraordinary metaphor for this idea to appear in the post-Cartesian wave of parallelism in seventeenth- and eighteenth-century France was the image of two clocks beating and striking in synchrony, satisfying the very French desire for order and harmony that existed at that time. The parallelists invite us to imagine two synchronized clocks, keeping perfect time together. When one strikes three o’clock, say, so does the other. If we were to imagine a slight time lag between them, it might seem tempting to think that the clock that strikes first *makes* the second clock strike. That would be a false inference, a fallacy that actually has a name: *post hoc ergo propter hoc*, or *after this therefore because of this*.

Leibniz was the most celebrated of the parallelists in the seventeenth and eighteenth century. He may have been struck by a well-publicized phenomenon observed by his teacher Huygens (himself a pupil of Descartes), who was the inventor of the pendulum clock. Huygens had been ill in bed, and while lying there had noticed that the pendulums of two clocks mounted in one case always ended up synchronized, though in opposite directions (“antisynchronized”) irrespective of their starting points, displaying what he called an “odd kind of sympathy.” The clocks were somehow regulating one another, but just how remained a mystery until 2002. In that year a team of scientists from Georgia Institute of Technology were able to explain the phenomenon with a sophisticated mathematical and physical model based on small vibrations in the case that interfere with one another.⁴ After ruling out air motion experimentally, Huygens had himself suspected but not been able to prove that the phenomenon was caused by small motions in the clock case, and the Georgia team proved him right.

Leibniz went considerably further with the thought, however. Mind and body do indeed act *as though* they were synchronized, and although they do not affect one another in a literal way, for him synchronization *is* causation. Nothing could go into or out of a “substance,” in the terminology of early modern philosophy, or a genuinely unified individual thing, which in this respect is rather

like an indivisible atom. The individual thing is “windowless,” in Leibniz’s metaphor. According to him the “mental pendulum” and the “physical pendulum” are synchronized in their behavior—fortunately not antisynchronized—though not by interaction. The synchronization comes with their initial creation by God from the substances’ “complete individual concepts,” which detail everything that will happen to them throughout their futures. This view, of course, has implications for freedom of the will, in which Leibniz was keenly interested. What concerns us, however, is the fact that the whole universe is arranged so that what we observe of it manifests all sorts of remarkable synchronizations. These include but are not limited to the synchronizations of the mind and the body, which, along with all the other synchronizations that constitute the laws of universe, are designed to bring about the best possible universe over time.

There is a big difference between Leibniz’s views and those of the so-called occasionalists. The occasionalists took the view that parallelism is true, but that physical events in the body are the *occasion* for God to act in the mind, and vice versa. The occasionalists, such as Geulincx, who before Leibniz had used the simile of the two clocks to illustrate parallelism, were impressed by the absolute power of God, and wanted to make all our actions and every other action in the world into actions whose motive or moving power is God. That this uncommonsensical and

implausible view survived as long as it did is a testament to the religious faith of the time, and to the dedication of the occasionalists in following their reason through to where it led, or seemed to lead. On the other hand, as Leibniz pointed out, the continuous need for action on God's part every time mind and body interact makes for a very hard-working God, and it is itself unacceptable on religious grounds as well as the ground of philosophical economy and the theological drive toward simplicity and piety.

The Role of the Conservation Laws

Another historically important point about interaction needs to be made on behalf of those who, like the parallelists, wish to deny that mind and body interact. It has to do with the laws of conservation in physics. Among these laws, which seem to be about as well established as anything could be in physics, the conservation of mass and energy tells us that in a "closed" system changing over time, the net total of mass or energy in the system stays the same. The system as a whole neither gains nor loses mass or energy. (There are particles with no mass, but they must have some energy, since energy is a function of frequency.)

Suppose that the human body is a closed physical system. In other words, it acts as it does because of the physical energy and mass that it contains, and it is insulated

from the effects of outside energy. This has been called “the causal closure of the physical.” If we want to change anything within the system, we will either have to use the energy that is already within the system, or we will have to introduce energy from the outside. If we use the energy in the system, then the mind, since it is not within the body, can have no effect on the body. If we do not use the energy already in the system, then mass and energy are not conserved, or the system is not closed.

However, if the mind is to effect a change in the body, then it must presumably introduce physical energy into the body. But according to our first proposition, the mind is nonphysical, and so it cannot expend physical energy. Here we can see that the conserved mass and energy are playing the same role as linear dimensions did in our first formulation of the mind–body problem. Lack of linear dimensions and spatial location on the part of the nonphysical is what makes the physical and nonphysical unable to interact. But the same result is obtained if we make mass or energy the defining characteristics of physical things. Physical and nonphysical things cannot interact. The body will not accept nonphysical energy, and the mind will not accept physical energy, in both cases because of the causal closure of the physical.

Versions of the four propositions are often if not always there when the mind–body problem is discussed. A specialized form is present when conservation laws are at

issue. Ernie Lepore and Barry Loewer,⁵ for example, express the mind–body problem as the difficulty of fitting together the following three propositions:

- (5) The mind and the body are distinct.
- (3b) The mental and the physical causally interact.
- (4b) The physical is causally closed.

Roughly speaking, (1) (“The mind is a nonphysical thing”) and (2) (“The body is a physical thing”) give us (5), proposition (3b) works like proposition (3) (“Mind and body interact”), and (4b) implies (4), that the mental and the physical cannot interact. The problem is that the physical world cannot reach out of itself into anything else that is nonphysical, but it must somehow interact with the mental, which is nonphysical. Similarly, the mental world cannot reach out of itself into anything else that is nonmental, but it must somehow interact with the physical, which is nonmental. Describing the inability of the mind to reach into the physical and of the body to reach into the mental is a way of stating the existence of the law of conservation of mass and energy, which has (4b) as a consequence. “Causally closed” means that energy or mass from causes outside the physical world, or outside the closed physical system, cannot get into it, and that it cannot contribute energy and mass to outside and nonphysical systems in such a way as to deplete the net total of its own energy and mass.

Naturally, if the mind is physical, then the body plus the mind can function as a closed system, and there is no difficulty with the laws of conservation. This amounts to denying (5), that the mind and body are physically distinct, which they are not, according to the physicalist.

The only option that does not seem available, given what physics has to say about conservation, is the denial of (4b). So a physicalist will deny (5) and affirm (3b). This involves the interesting claim that the mental is physical, or a denial of the claim that the mind is nonphysical. A parallelist, on the other hand, will affirm (4b) and deny (3b), telling us that the mind and the body are indeed distinct, but that they occupy parallel and noninteracting realms.

The lesson so far is that we should be either parallelists or physicalists, but not interactionists.

Epiphenomenalism, Emergentism, and Supervenience

There is another form of dualism that was especially popular at the end of the nineteenth century. It has seen a modest resurgence recently, in much more sophisticated forms, though more as an object of study, perhaps, than as a view actually to be believed. Known by its formidable Greek-derived name, *epiphenomenalism*, it is the claim that mental events and the mind are “epiphenomena.” “Epi-” is

a Greek prefix that means “on the occasion of” or “in addition” to. “Phenomena” are the things that appear, or happen, so epiphenomena are things that appear in *addition* to what might be called the basic phenomena. For most epiphenomenalists, if not all, the basic phenomena are those of the physical world, and mental phenomena and events are attendant on physical phenomena. Epiphenomenalism is the view that physical events cause mental events but mental events do not cause physical events.

There is an obvious comparison to be made with shadows. My hands curled up in the right way can be made to cast a shadow that looks like an eagle’s head onto a wall or screen. The shadow is dependent on my hands, but what my hands do is not dependent on what the shadow does. It would be amusing but physically difficult for the eagle on the screen to open its beak, say, and force my fingers to move. The image of the eagle projected onto the screen is *just* a shadow.

Almost nobody holds or has held the reversed epiphenomenalist view that mental events cause physical ones but not the other way round, and it is, I think, fairly obvious why. For one thing, there are the obvious phenomena to think about, such as brain damage. But at a deeper level the epiphenomenalists are those physicalists who want to be strict physicalists but who cannot quite see their way to deny the existence of fully mental events, though they also find it hard to see how mental events can exist at all. In

Epiphenomenalism is the view that physical events cause mental events but mental events do not cause physical events.

conformity with their physicalism, they then downgrade the importance and causal power of mental events as far as possible in the physical scheme of things.

Still, if epiphenomenalists are really physicalists under the skin, they are inconsistent ones, since epiphenomenalism admits the existence of genuinely mental events. There are mental events, it claims, but they have no causal power, unlike physical events.

According to the Victorian biologist Thomas Henry Huxley, our consciousness is a “collateral product” of the “mechanism of the body” and “as completely without any power of modifying the working of the body as the steam-whistle which accompanies the work of a locomotive engine.”⁶ Volition too is an emotion that “indicates” physical changes but does not “cause” them.

There is something clearly wrong with Huxley’s simile of the steam-whistle, since nothing prevents us from rigging up a steam-whistle so that every time it blows, the steam activates a fan rigged to an electrical circuit that brakes the train, and that when the whistle is not blowing, the train resumes its normal speed. The whistle then has definite and specifiable physical effects, and there is nothing in the nature of the physical world to prevent this sort of causal loop.

In the case of the mind or consciousness or soul, Huxley would rule out the causal loop. Yet why is it impossible? *Why* is the mind causally inert? Huxley does not address

the question, but it certainly seems to push epiphenomenalism hard in the direction of property dualism. For it is an odd thing indeed, an odd substance, that can have no effects whatsoever. Properties seem more suited than substances to causal inactivity. Even so, one might think that the property of *being hot* can cause me to have the property of *wanting a drink*. Why do I have the property of *wanting a drink*? Because I have the property of *being hot*.

We should keep clearly in mind the fact that epiphenomenalism is a form of dualism. It allows interaction between mind and body in one direction, from mental to physical, but not the other. But there is still a contradiction here. Epiphenomenalism has cut the mind–body problem down to half its original size, so to speak, but what remains is every bit as intractable as the original full-scale version. We don't have to deal with the mind acting on the body, but how can the body act on the mind, if the mind is nonphysical, and physical things cannot act on nonphysical things? In fact I think epiphenomenalism counts as a rough-and-ready philosophy of mind, but not as a genuine solution to the mind–body problem. This may explain why philosophers have on the whole been less than interested in it, and why it has been referred to as “the curse of epiphenomenalism” by one writer (Stephen Law) in the philosophy of mind.⁷

Emergentism is a view of the relation between mind and body of roughly the same kind as epiphenomenalism,

in the sense that the physical is dominant and the mental is a sort of by-product, but it is important to see what the difference is between the two. Epiphenomenalism is a kind of dualism, in which two separate kinds of events exist and are causally related. With emergentism, the relation between the mental and the physical is much closer. It should perhaps be discussed later, in the next chapter, as a form of physicalism, but it seems to me that the comparisons and contrasts between epiphenomenalism and emergentism are interesting ones.

There is a mystery, very much at the center of the modern mind-body problem, of why it should be *pain* that emerges from the brain areas that are activated by A δ or C fiber stimulation. (The A δ fibers are associated with acute and sharp pain, the C fibers with dull or burning pain.) But, according to the emergentist, there simply is *no answer* to the question why it is *pain* that ultimately emerges from the brain areas that are activated by A δ or C fiber stimulation, and not something else entirely. Pain does not emerge from the stimulation of the fibers in the way that sixteen ounces just turns into one pound; but then one wonders how on earth it *is* related to the stimulation of the fibers. Well, it *emerges* from them, comes the answer.

The emergentists both accept and deny (1), that the mind is nonphysical. On the one hand, the mind is physical, because it is really driven by the structures from which it emerges. On the other hand, the mind is nonphysical,

because it has “emerged” from the physical. But how can truly novel properties, such as pain, emerge from the stimulation of the fibers? If they “emerge” *from* the physical, then they are nonphysical. But if they are genuinely nonphysical, how can they “emerge,” and why do they need to?

It seems to me that the emergentists must make up their minds. If with the mind we have a genuinely new phenomenon, a nonphysical and nonspatial one that has emerged like a butterfly out of the chrysalis of matter, then it cannot affect the body, since the body has exclusively the wrong kind of properties to interact with the mind, that is, physical and spatial ones. From this point of view it is hard to see how mind could emerge in the first place, since in emerging it makes itself spatial. If, on the other hand, the new phenomenon has a complete dependence on the physical and spatial phenomena, and can engage with them, it is hard to see how it is anything *but* them, and therefore not a new and emergent property at all.

Emergentists accept the fact that mind can turn around and *do* things to matter, but they do not explain *how* this can happen if the mind has “emerged” and is not physical. If the mind has emerged as nonphysical, we need to understand the way in which it can then interact with the physical. And that is the mind–body problem.

There is a concept that may seem to help with understanding how something can both be nothing but its base properties, and at the same time something emergent,

something in addition to the base properties. Many, even most, emergentists have used the concept of *supervenience* and been grateful for the light it casts on the relation between mind and body. According to these emergentists the mind *supervenes* on the body.

The concept of *supervenience* is a difficult one, but the main idea is something like this. Suppose a property *A* supervenes on a property *B*. For example, some geometrical or aesthetic property *A* supervenes on the properties, collectively *B*, for “base,” of a spatial figure or of painting, such as being thus-and-such a closed figure, or having thus-and-such colors, lines, and forms. We say that *A* supervenes on *B* when there cannot be a change in *A* without there also being a change in *B*. One cannot suppose the aesthetic properties of the painting changing without the physical properties having changed as well. If the aesthetic properties are to differ, so too must the physical properties. In this sense it can be said that the *A*-properties are generated by the base properties.

The emergentists who make use of the concept of *supervenience* believe that (2) the proposition “The body is physical,” but will reject (3), the proposition “Mind and body interact.” Instead they will say that mind *supervenes* on body, or more particularly on the part of the body with the right kind of tissue, namely, the brain.

And yet there does seem to be a kind of causal power possessed by the human mind and consciousness,

recognized in (3). Mind interacts with body, which is to say it has effects in the physical world. Emergentism, even with the more refined concept of supervenience on board, cannot do justice to this causal power. The reason is that emergentism is actually a form of physicalism, and it attempts to deny the existence of the nonphysical except in a very diluted form that cannot accomplish what philosophers call “mental causation.”

The reason emergentism is not a very popular view is that it is not a very clear one. On the one hand, the mind “emerges” and engages in mental causation. On the other hand, it is the creature of the forces from below on which it supervenes, and cannot attain to any sort of causal power.