

# 4

## Materialism and Christian Belief

*Alvin Plantinga*

According to *materialism*, human persons are material objects. They are not immaterial things, or objects, or substances; neither do they contain as parts immaterial selves or souls or entelechies. Their parts are material: flesh and bones and blood, molecules, atoms, electrons and quarks (if in fact there are such things). This view, of course, goes contrary to the vast bulk of the Christian tradition. This is not to say, *pace* Plato (or anyway Socrates), that the body is the prison house of the soul, or that our present attachment to the body is to be deplored, as if it were a temporary, makeshift arrangement (due to sin?) to be jettisoned in the next life. Not at all; on the traditional Christian view, God has designed human beings to have bodies; they function properly only if embodied; and of course Christians look forward to the resurrection of the body. My body is crucial to my well-being and I can flourish only if embodied. As W. H. Auden put it, “I wouldn’t be caught dead without my body.”

Materialism goes contrary to the Christian tradition; even worse (so I’ll argue), it is false. As I see it, therefore, Christian philosophers ought to be dualists. Now most naturalists, of course, are materialists; but so are a surprising number of Christian philosophers.<sup>1</sup> I’ll argue that this is a mistake. In “Against Materialism”<sup>2</sup> I also argue that materialism is false. This paper covers some of the same ground as that one. It differs in that it omits a couple of sections; it

In addition to the people mentioned in the text, I thank Michael Bergmann, E. J. Coffman, Evan Fales, Richard Fumerton, Trenton Merricks, William Ramsey, and the members of the Notre Dame Center for Philosophy of Religion discussion group, in particular Thomas Flint and Peter van Inwagen, as well as the others I have inadvertently overlooked.

<sup>1</sup> See e.g. Peter van Inwagen, *Material Beings* (Ithaca, NY: Cornell University Press, 1990), and “Dualism and Materialism: Athens and Jerusalem?”, *Faith and Philosophy*, 12/4 (Oct. 1995), 475–88; Trenton Merricks, “The Resurrection of the Body and the Life Everlasting”, in Michael Murray (ed.), *Reason for the Hope Within* (Grand Rapids, Mich. Eerdmans, 1999); Nancey Murphy “Human Nature: Historical, Scientific, and Religious Issues”, in *Whatever Happened to the Soul?* (Minneapolis: Fortress Press, 1998), 1–30; Lynne Rudder Baker, “Need a Christian Be a Mind/Body Dualist?”, *Faith and Philosophy*, 12/4 (Oct. 1995), 498–504; and Kevin Corcoran, “Persons and Bodies”, *Faith and Philosophy*, 15/3 (July 1998), 324–40.

<sup>2</sup> In *Faith and Philosophy*, 23/1 (January 2006), 3–32.

also adds sections dealing with (1) the alleged arguments for materialism, and (2) the relevance of Christian theism to the question, and (3) an appendix dealing with the way in which materialists try to explain how it could be that a material structure or event could be a belief. With respect to (2), there are, I believe, at least three points to be made. First, there is Scripture; the New Testament in particular contains much that at any rate strongly suggests that materialism is false. Second, Christian theism is crucially relevant to the epistemology of the situation, and that in at least two ways:

(a) Given Christian theism, we know that it is at any rate *possible* that there be immaterial thinking things. God Himself is an immaterial thinking thing; hence, by the argument form *ab esse ad posse*, the most powerful argument for possibility, it follows that immaterial thinking things are possible. Furthermore, Christian theism strongly suggests that there are *created* immaterial thinking things: angels, for example, as well as Satan and his minions.

(b) Considerations from the Christian faith are powerfully relevant to the alleged objections to dualism and arguments for materialism.<sup>3</sup>

Finally, certain crucial Christian doctrines (for example, Incarnation and the resurrection of the dead) fit better—much better, I'd say—with dualism than with materialism.

I'll restrict myself, for the most part, to the second of these three points. Section 1 of this paper will follow "Against Materialism" in presenting a couple of "strictly philosophical" arguments against materialism; in Section 2 I'll turn to the considerations from Christian theism.

## 1. TWO ARGUMENTS FOR DUALISM

Christian philosophers, so I say, should be dualists; but of course dualism itself is multiple, if not legion. There is the view—embraced by Plato, Augustine, Calvin, Descartes, and a thousand others—according to which a human person is an immaterial substance: a thing, an object, a substance, a suppositum (for my Thomist colleagues), and a thing that isn't material. Second, there is the view the name 'dualism' suggests: the view according to which a human person is somehow a sort of composite substance S composed of a material substance S\* and an immaterial substance S\*\*.<sup>4</sup>

Third, there is also the important but obscure view of Thomas Aquinas and his followers. Is this a form of dualism? The question is vexed. According to Aquinas,

<sup>3</sup> Substance dualism and materialism are not uncontroversial contradictories (perhaps, as some suggest, we aren't substances at all, but events, or maybe momentary collections of mental states, or transtemporal collections of person states or stages). For present purposes, however, I'll take it that substance dualism and materialism are the only relevant positions, and speak indifferently of arguments for materialism and arguments against dualism.

<sup>4</sup> See Richard Swinburne, *The Evolution of the Soul* (Oxford: Clarendon Press, 1986), 145.

a human person is a material substance with an immaterial part, the soul. Aquinas says, of this immaterial part, that it is itself a substance. Furthermore the soul, this immaterial part, has the property of possibly thinking (believing, desiring, hoping, deciding, etc.), and after death, *does* think. But Aquinas, also says that the soul is *the form of* the body.<sup>5</sup> A form, however, at least as far as I can see, is or is like a *property*; and a property, presumably, *can't* think. If the soul is a form, therefore, how can it be capable of thinking?<sup>6</sup> This is a tough question, but perhaps we needn't go into it at the moment. A more pressing question is this: I'll be arguing that it is possible that I exist when my body doesn't; is that a possibility, on Thomas's view? True, on his view my *soul* can exist when my body doesn't; but it also seems, on this view, that I am not identical with my soul. Rather, I am a material object that has an immaterial soul as a part. So (on his view) can I exist when my body does not? If the answer is no, then Aquinas's view is not felicitously counted as a version of dualism; at least it is not among the versions of dualism for which I mean to argue. If, on the other hand, the answer is yes, we can welcome Aquinas (perhaps a bit cautiously) into the dualist camp.

Three more initial comments: (*a*) when I speak of possibility and necessity, I mean possibility and necessity in the broadly logical sense—metaphysical possibility and necessity, as it is also called; (*b*) I won't be arguing that it is possible that I (or others) can exist disembodied, with no body at all, although I believe that this is in fact possible;<sup>7</sup> (*c*) I will make no claims about what is or isn't conceivable or imaginable. That is because imaginability isn't strictly relevant to possibility at all; conceivability, on the other hand, is relevant only if 'it's conceivable that p' is to be understood as implying or offering evidence for 'it's possible that p'. (Similarly for 'it's inconceivable that p'.) It is therefore simpler and much less conducive to confusion to speak just of possibility. I take it we human beings have the following epistemic capacity: we can consider or envisage a proposition or state of affairs and, at least sometimes, determine its modal status—whether it is necessary, contingent, or impossible—just by thinking, just by an exercise of thought.<sup>8</sup>

<sup>5</sup> *Summa Theologiae*, I, Q. 75.

<sup>6</sup> For an interesting suggestion as to the answer, see Brian Leftow's "Souls Dipped in Dust", in Kevin Corcoran (ed.), *Soul, Body and Survival* (Ithaca, NY: Cornell University Press, 2001), 120 ff.

<sup>7</sup> I can't help concurring with David Armstrong, no friend of dualism: "But disembodied existence seems to be a perfectly intelligible supposition . . . Consider the case where I am lying in bed at night thinking. Surely it is logically possible that I might be having just the same experiences and yet not have a body at all. No doubt I am having certain somatic, that is to say, bodily sensations. But if I am lying still these will not be very detailed in nature, and I can see nothing self-contradictory in supposing that they do not correspond to anything in physical reality. Yet I need be in no doubt about my identity" (*A Materialist Theory of Mind* (London: Routledge, 1968), 19).

<sup>8</sup> See my *Warrant and Proper Function* (New York: Oxford University Press, 1993), ch. 6. See also George Bealer, "Intuition and the Autonomy of Philosophy" in Michael DePaul and William Ramsey (eds.), *Rethinking Intuition* (Lanham, Md.: Rowman & Littlefield, 1998), 201 ff.

## The Replacement Argument: An Argument from Possibility

I begin by assuming that there really is such a thing, substance, or suppositum as I, I myself. Of course I'm not unique in that respect; you too are such that there really is such a thing as you, and the same goes for everybody else. We are substances. Now suppose I were a material substance: which material substance would I be? The answer, I should think, is that I would be my body, or some part of my body, such as my brain or part of my brain. Or perhaps I would be something more exotic: an object distinct from my body that is constituted from the same matter as my body and is colocated with it.<sup>9</sup> What I propose to argue is that I am none of those things: I am not my body, or some part of it such as my brain or a hemisphere or other part of the latter, or an object composed of the same matter as my body (or some part of it) and colocated with it. For simplicity (and nothing I say will depend on this simplification) I shall talk for the most part just about my body, which I'll name 'B'. (I was thinking of naming it 'Hercules' or maybe 'Arnold', but people insisted that would be unduly self-congratulatory.)

The general strategy of this first argument is as follows. It seems possible that I continue to exist when B, my body, does not. I therefore have the property *possibly exists when B does not*; B, however, clearly lacks that property. By Leibniz's Law, therefore (more specifically, the Diversity of Discernibles), I am not identical with B. But why think it possible that I exist when my body does not? Strictly speaking, the replacement argument is an argument for this premise. Again, I conduct the argument in the first person, but naturally enough the same goes for you (although of course you will have to speak for yourself).

So first, at a macroscopic level. A familiar fact of modern medicine is the possibility and actuality of limb and organ transplants and prostheses. You can get a new heart, liver, lungs; you can also get knee, hip, and ankle replacements; you can get prostheses for hands and feet, arms and legs, and so on. Now it seems possible—possible in that broadly logical sense—that medical science should advance to the point where I remain fully dressed and in my right mind (perhaps reading the *South Bend Tribune*) throughout a process during which each of the macroscopic parts of my body is replaced by other such parts, the original parts being vaporized in a nuclear explosion—or better, annihilated by God. But if this process occurs rapidly—during a period of one microsecond, let's say—B will

<sup>9</sup> See, e.g. Dean Zimmerman, "Material People", in Michael Loux and Dean Zimmerman (eds.), *The Oxford Handbook of Metaphysics* (Oxford: Oxford University Press, 2003), 504 ff. Zimmerman himself seems attracted to the thought that "the mass of matter" of which one's body is composed is an object distinct from the latter but colocated with it (although of course he is not attracted to the idea that a person just is this mass of matter). He regards the mass of matter as more fundamental (and therefore more ontologically respectable) than the ever-changing body; so he is inclined to regard the latter as a mere "logical construction" or some other sort of entity dependent upon different masses of matter at different times.

no longer exist. I, however, will continue to exist, having been reading the comic page during the entire process.

But what about my brain, you ask—is it possible that my brain be replaced by another, the brain I now have being destroyed, and I continue to exist? It certainly seems so. Think of it like this. It seems possible (in the broadly logical sense) that one hemisphere of my brain be dormant at any given time, the other hemisphere doing all that a brain ordinarily does. At midnight, we can suppose, all the relevant ‘data’ and ‘information’ is ‘transferred’ via the corpus callosum from one hemisphere—call it ‘ $H_1$ ’—to the other hemisphere— $H_2$ —whereupon  $H_2$  takes over operation of the body and  $H_1$  goes dormant. This seems possible; if it were actual, it would also be possible that the dormant half,  $H_2$ , be replaced by a different dormant half (in the same computational or functional state, if you like) just before that midnight transfer; then the transfer occurs, control switches to the new  $H_2$ , and  $H_1$  goes dormant—at which time it is replaced by another hemisphere in the same computational or functional condition. In a period of time as brief as you like, therefore, both hemispheres will have been replaced by others, the original hemispheres and all of their parts annihilated by God. Throughout the whole process I serenely continue to read the comics.

This suffices, I think, to show that it’s possible that I exist when neither my body nor any part of it exists. What about material objects distinct from my body and its parts, but colocated with it (or one of them) and constituted by the same matter as they? I doubt very much that there could be any such things. If objects of this kind *are* possible, however, the above argument also shows or at least suggests that possibly, I exist when none of them does. For example, if there is such a thing as *the matter of which B is composed*—if that phrase denotes a thing or object<sup>10</sup>—it too would be destroyed by God’s annihilating all the parts of my body.

Of course very many different sorts of object of this kind—objects constituted by the matter of my body and colocated with it—have been suggested, and I don’t have the space here to deal with them all. However, we can offer a version of the replacement argument that will be relevant to most of them. Turn from macroscopic replacement to microscopic replacement. This could go on at several levels: the levels of atoms, molecules, or cells, for example. (It could also go on at the level of elementary particles—electrons and quarks, if indeed there really are such things, and if indeed they are elementary particles.) Let’s think about it at the cellular level. It seems entirely possible that the cells of which my body is composed be rapidly—within a microsecond or two—replaced by other cells of the same kind and in the same state, the original cells being instantly destroyed. It also seems entirely possible that this process of replacement take place while I remain conscious, thinking about dualism and marveling at some

<sup>10</sup> *ibid.*

of the appalling arguments against it produced by certain materialists.<sup>11</sup> Then I would exist at a time at which B did not exist.

But is it really true that this process of replacement would result in the destruction of B? After all, according to current science, all the matter in our bodies is replaced over a period of years with other matter, without any obvious compromise of bodily integrity or identity. As a matter of fact, so they say, the matter in our brains is completely replaced in a much shorter time.<sup>12</sup> Why should merely accelerating this process make a difference?<sup>13</sup>

Well, speed kills. When a part (a cell, say) is removed from an organism and replaced by another cell, the new cell doesn't become part of the organism instantaneously; it must be integrated into the organism and assimilated by it.<sup>14</sup> This takes time—maybe not much time, but still a certain period of time. At the instant the new part is inserted into the organism,<sup>15</sup> and until the time of assimilation has elapsed, the new part is not yet a part of the organism, but a foreign body occupying space within the spatial boundaries of the organism. (Clearly not everything, nor even everything organic, within the spatial boundaries of your body is *part* of your body: think of the goldfish you just swallowed, or a tapeworm.) Let's use the phrase 'assimilation time' to denote the time required for the assimilation of the new part. To be rigorous, we should index this to the part (or kind of part) and the organism in question; different parts may require different periods of time for their assimilation by different organisms. For simplicity, though, let's assume all parts and organisms have the same assimilation time; this simplification won't make any difference to the argument.

That a given part and organism are such that the time of assimilation for the former with respect to the latter is  $dt$  for some specific period of time  $dt$  is, I take it, a contingent fact. One thinks the velocity of light imposes a lower limit here, but the time of assimilation could be much greater. (For example, it could depend on the rate of blood flow, the rate of intracellular transport, and the

<sup>11</sup> One such argument, for example, apparently has the following form: (a) Many people who advocate p, do so in the service of a hope that science will never be able to explain p; therefore (b) not-p. See Daniel Dennett, *Darwin's Dangerous Idea* (New York: Simon and Schuster, 1995), 27.

Another seems to have the form (a) If you believe p, prestigious people will laugh at you; therefore (b) not-p. (or perhaps (b\*) don't believe p?) See Daniel Dennett, *Explaining Consciousness* (Boston: Little, Brown, 1991), 37.

<sup>12</sup> "But on the kinds of figures that are coming out now, it seems like the whole brain must get recycled about every other month." John McCrone, "How Do You Persist When Your Molecules Don't?" *Science and Consciousness Review* (web-journal, June 2004, No. 1).

<sup>13</sup> Here I am indebted especially to Michael Rea.

<sup>14</sup> See e.g. David Hershenov, "The Metaphysical Problem of Intermittent Existence and the Possibility of Resurrection," *Faith and Philosophy*, 20/1 (Jan. 2003), 33.

<sup>15</sup> Complaint: this new 'part' as you call it, isn't really a part, at first, anyway, because at first it isn't yet integrated into the organism. Reply: think of 'part' here, as like 'part' in 'auto parts store'. Would you complain that the auto parts store is guilty of false advertising, on the grounds that none of those carburetors, spark plugs and piston rings they sell is actually part of an automobile?

rate at which information is transmitted through neuron or nerve.) God could presumably slow down this process or speed it up.

There is also what we might call ‘the replacement time’: the period of time from the beginning of the replacement of the first part by a new part to the end of the time of the replacement of the last part (the last to be replaced) by a different part. The time of replacement is also, of course contingent; a replacement can occur rapidly or slowly. Presumably there is no non-zero lower limit here; no matter how rapidly the parts are replaced, it is possible in the broadly logical sense that they be replaced still more rapidly.

What’s required by the Replacement argument (or at any rate what’s sufficient for it) is

*Replacement* It is possible that: the cells in B are replaced by other cells and then instantly annihilated while I continue to exist; and the replacement time for O and those cells is shorter than the assimilation time.<sup>16</sup>

### Can a Material Thing Think? An Argument from Impossibility

The replacement argument is an argument from possibility; as such, it proceeds from an intuition, the intuition that it is possible that my bodily parts, macroscopic or microscopic, be replaced while I remain conscious. But some people distrust modal intuitions. Of course it’s impossible to do philosophy (or for that matter physics) without invoking modal intuitions of one sort or another or at any rate making modal declarations of one sort or another.<sup>17</sup> Still, it must be conceded that intuition can sometimes be a bit of a frail reed. True, there is no way to conduct philosophy that isn’t a frail reed, but intuition is certainly fallible. Further, some might think modal intuitions particularly fallible—although almost all of the intuitions involved in philosophy have important modal connections. Still further, one might think that intuitions of

<sup>16</sup> “Against Materialism”, 3. Contains some Objections and Replies to the argument just sketched out.

<sup>17</sup> Realists will say that there can’t be similarity without a property had by the similar things, thus resting on an alleged intuition of impossibility; nominalists will deny this claim, thus resting on an alleged intuition of possibility. In his argument for indeterminacy of translation, Quine claims that the native’s behavior is consistent with his meaning ‘rabbit state’ or ‘undetached rabbit part’ or ‘rabbit’ by ‘gavagai’, thus (despite his animadversions) relying on an intuition of possibility. Similarly for his and others’ claims about the underdetermination of theory by evidence. Further, anyone who proposes an analysis (e.g. of knowledge) relies on intuition, as does someone who objects to such an analysis (by proposing a Gettier case, for example). In philosophy of mind we have Jackson’s Mary example, Burge’s arthritis example, twin earth arguments for a posteriori necessities and wide content, refutations of phenomenalism and behaviorism, and much else besides, all of which rely centrally and crucially on intuition. Materialists either take materialism in the basic way, thus relying on intuition, or they accept it on the basis of argument; every argument for materialism I’ve seen relies on intuition (e.g. the intuition that an immaterial thing can’t cause effects in the hard, heavy, massive material world). Indeed, take your favorite philosophical argument or position: it will doubtless rely on intuition.

*possibility* are especially suspect.<sup>18</sup> That is because it seems easy to confuse *seeing the possibility of p* with *failing to see the impossibility of p*. You can't see why numbers couldn't be sets; it doesn't follow that what you see is that they *could* be sets. Maybe I can't see why water couldn't be composed of something other than H<sub>2</sub>O; it doesn't follow that what I see is that water could be something other than H<sub>2</sub>O. And perhaps, so the claim might go, one who finds the replacement argument attractive is really confusing seeing the possibility of the replacements in question with failing to see their impossibility. Granted: I can't see that these replacements are impossible; it doesn't follow that what I see is that they are indeed possible.

To be aware of this possible source of error, however, is to be forewarned and thus forearmed. But for those who aren't mollified and continue to distrust possibility intuitions, I have another argument for dualism—one that depends on an intuition, not, this time, of possibility, but of impossibility. One who distrusts possibility intuitions may think more kindly of intuitions of impossibility—perhaps because she thinks that for the latter there isn't any obvious analogue of the possible confusion between failing to see that something is impossible with seeing that it is possible. Or rather, while there *is* an analogue—it would be confusing failure to see the possibility of p with seeing the impossibility of p—falling into that confusion seems less likely. In any event, the argument I'll now propose is for the conclusion that no material objects can think—that is, reason and believe, entertain propositions, draw inferences, and the like. But of course I can think; therefore I am not a material object.

### *Leibniz's Problem*

I (and the same goes for you) am a certain kind of thing: a thing that can think. I believe many things; I also hope, fear, expect, anticipate many things. I desire certain states of affairs (desire that certain states of affairs be actual). I am capable of making decisions. I am capable of acting, and capable of acting on the basis of my beliefs and desires. I am conscious; and conscious of a rich, kaleidoscopic constellation of feeling, mental images, beliefs, and ways of being appeared to, some of which I enjoy and some of which I dislike. Naturally enough, therefore, I am not identical with any object that lacks any or all of these properties. What I propose to argue next is that some of these properties are such that no material object can have them. Again, others have offered similar arguments. In particular, many have seen a real problem for materialism in *consciousness*: it is extremely difficult to see how a material object could be conscious, could enjoy that vivid and varied constellation of feelings, mental images, and ways of being appeared to. Others have argued that a material object can't make a decision (although of course we properly speak, in the loose and popular sense, of the chess-playing

<sup>18</sup> See below, p. 113 ff.

computer as deciding which move to make next). These arguments seem to me to be cogent.<sup>19</sup> Here, however, I want to develop another argument of the same sort, another problem for materialism, a problem I believe is equally debilitating, and in fact fatal to materialism. Again, this problem is not a recent invention; you can find it or something like it in Plato. Leibniz, however, offers a famous and particularly forceful statement of it:

It must be confessed, moreover, that *perception*, and that which depends on it, *are inexplicable by mechanical causes*, that is by figures and motions. And supposing there were a machine so constructed as to think, feel and have perception, we could conceive of it as enlarged and yet preserving the same proportions, so that we might enter it as into a mill. And this granted, we should only find on visiting it, pieces which push one against another, but never anything by which to explain a perception. This must be sought for, therefore, in the simple substance and not in the composite or in the machine.<sup>20</sup>

Now Leibniz uses the word ‘perception’ here; he’s really thinking of mental life generally. His point, in this passage, is that mental life—perception, thought, decision—cannot arise by way of the mechanical interaction of parts. Consider a bicycle; like Leibniz’s mill, it does what it does by virtue of the mechanical interaction of its parts. Stepping down on the pedals causes the front sprocket to turn, which causes the chain to move, which causes the rear sprocket to turn, which causes the back wheel to rotate. By virtue of these mechanical interactions, the bicycle does what it does, that is, transports someone from one place to another. And of course machines generally—jet aircraft, refrigerators, computers, centrifuges—do their things and accomplish their functions in the same way. So Leibniz’s claim, here, is that thinking can’t arise in this way. A thing can’t think by virtue of the mechanical interaction of its parts.

Leibniz is thinking of *mechanical* interactions—interactions involving pushes and pulls, gears and pulleys, chains and sprockets. But I think he would say the same of other interactions studied in physics, for example those involving gravity, electromagnetism, and the strong and weak nuclear forces. Call these ‘physical interactions’. Leibniz’s claim is that thinking can’t arise by virtue of physical interaction among objects or parts of objects. According to current science, electrons and quarks are simple, without parts.<sup>21</sup> Presumably neither can think—neither can adopt propositional attitudes; neither can believe, doubt, hope, want, or fear. But then a proton composed of quarks won’t be able to think either, at least by way of physical relations between its component quarks, and the same will go for an atom composed of protons and electrons, a molecule composed of atoms, a cell composed of molecules, and an organ (e.g. a brain)

<sup>19</sup> There is also the complex but powerful argument offered by Dean Zimmerman, “Material People”, 517 ff.

<sup>20</sup> *Monadology*, 17. In *Leibniz Selections*, ed. Philip Weiner (New York: Charles Scribner’s Sons, 1951), 536.

<sup>21</sup> Although there are speculative suggestions that quarks may in fact be composed of strings.

composed of cells. If electrons and quarks can't think, we won't find anything composed of them that *can* think by way of the physical interaction of its parts.

Leibniz is talking about thinking generally; suppose we narrow our focus to *belief* (although the same considerations apply to other propositional attitudes). What, first of all, would a belief *be*, from a materialist perspective? Suppose you are a materialist, and also think, as we ordinarily do, that there are such things as beliefs. For example, you hold the belief that Marcel Proust is more subtle than Louis L'Amour. What kind of a thing is this belief? Well, from a materialist perspective, it looks as if it would have to be something like a long-standing event or structure in your brain or nervous system. Presumably this event will involve many neurons related to each other in subtle and complex ways. There are plenty of neurons to go around: a normal human brain contains some 100 billion. These neurons, furthermore, are connected with other neurons at synapses; a single neuron can be involved in several thousand synapses, and there are some  $10^{15}$  synaptic connections. The total number of possible brain states, then, is absolutely enormous, vastly greater than the  $10^{80}$  electrons they say the universe contains. And the total number of possible neuronal events, while no doubt vastly smaller, is still enormous. Under certain conditions, groups of neurons involved in such an event fire, producing electrical impulses that can be transmitted (with appropriate modification and input from other structures) down the cables of neurons that constitute effector nerves to muscles or glands, causing, for example, muscular contraction and thus behavior.

From the materialist's point of view, therefore, a belief will be a neuronal event or structure of this sort. But if this is what beliefs are, they will have two very different sorts of properties. On the one hand they will have *electrochemical* or *neurophysiological* properties ('NP properties', for short). Among these would be such properties as that of involving  $n$  neurons and  $n*$  connections between neurons, properties that specify which neurons are connected with which others, what the rates of fire in the various parts of the event are, how these rates of fire change in response to changes in input, and so on. But if the event in question is really a *belief*, then in addition to those NP properties it will have another property as well: it will have a *content*. It will have to be the belief that  $p$ , for some proposition  $p$ . If this event is the belief that Proust is a more subtle writer than Louis L'Amour, then its content is the proposition *Proust is more subtle than Louis L'Amour*. My belief that naturalism is all the rage these days has as content the proposition *Naturalism is all the rage these days*. (That same proposition is the content of the German speaker's belief that naturalism is all the rage these days, even though she expresses this belief by uttering the German sentence 'Der Naturalismus ist dieser Tage ganz gross in Mode'; beliefs, unlike sentences, do not come in different languages.) It is in virtue of having a content, of course, that a belief is true or false: it is true if the proposition which is its content is true, and false otherwise. My belief that all men are mortal is true because the proposition which constitutes its content is true, but Hitler's belief that the

Third Reich would last a thousand years was false, because the proposition that constituted its content was false.<sup>22</sup>

And now the difficulty for materialism is this: how does it happen, how can it be, that an assemblage of neurons, a group of material objects firing away *has a content*? How can that happen? More poignantly, *what is it* for such an event to have a content? What is it for this structured group of neurons, or the event of which they are a part, to be related, for example, to the proposition *Cleveland is a beautiful city* in such a way that the latter is its content? A single neuron (or quark, electron, atom, or whatever) presumably isn't a belief and doesn't have content; but how can belief, content, arise from, be constituted by, physical interaction among such material entities as neurons? As Leibniz suggests, we can examine this neuronal event as carefully as we please; we can measure the number of neurons it contains, their connections, their rates of fire, the strength of the electrical impulses involved, the potential across the synapses—we can measure all this with as much precision as you could possibly desire; we can consider its electrochemical, neurophysiological properties in the most exquisite detail; but nowhere, here, will we find so much as a hint of content. Indeed, none of this seems even vaguely *relevant* to its having content. None of this so much as slyly suggests that this bunch of neurons firing away is the belief that Proust is more subtle than Louis L'Amour, as opposed, for example, to the belief that Louis L'Amour is the most widely published author from Jamestown, North Dakota. Indeed, nothing we find here will so much as slyly suggest that it has a content of *any* sort. Nothing here will so much as slyly suggest that it is *about* something, in the way a belief about horses is about horses.

The fact is, we can't see how it *could* have a content. It's not just that we don't know or can't see how it's done. When light strikes photoreceptor cells in the retina, there is an enormously complex cascade of electrical activity, resulting in an electrical signal to the brain.<sup>23</sup> I have no idea how all that works; but of course I know it happens all the time. But the case under consideration is different. Here it's not merely that I don't know how physical interaction among neurons brings it about that an assemblage of them has content and is a belief. No, in this case, it seems upon reflection that such an event could *not* have content. It's a little like trying to understand what it would be for the number seven, for example, to weigh 5 pounds, or for an elephant (or the unit set of an elephant) to be a proposition. (Pace the late (and great) David Lewis, according

<sup>22</sup> A materialist might take a leaf from those who accept 'adverbial' accounts of sensation, according to which there aren't any red sensations or red sense data or red appearances: what there are instead are cases of someone's sensing redly or being appeared to redly. Similarly, the materialist might claim that there isn't any such thing as the belief that all men are mortal (or any other beliefs); what there is instead are cases of people who believe in the all-men-are-mortal way. This may or may not make sense; if it does make sense, however, a person will presumably believe in the all-men-are-mortal way only if she harbors a neuronal structure or event that has as content the proposition all men are mortal.

<sup>23</sup> See Michael Behe, *Darwin's Black Box* (New York: Free Press, 1996), 18 ff.

to whom the unit set of an elephant *could* be a proposition; in fact, on his view, there are uncountably many elephants the unit sets of which *are* propositions.) We can't see how that could happen; more exactly, what we can see is that it *couldn't* happen. A number just isn't the sort of thing that can have weight; there is no way in which the number seven or any other number could weigh anything at all. The unit set of an elephant, let alone the elephant itself, can't be a proposition; it's not the right sort of thing. Similarly, we can see, I think, that physical activity among neurons can't constitute content. These neurons are clicking away, sending electrical impulses hither and yon. But what has this to do with content? How is content or aboutness supposed to arise from this neuronal activity? How can such a thing be a belief? You might as well say that thought arises from the activity of the wind or the waves. But then no neuronal event can as such have a content, can be *about* something, in the way in which my belief that the number seven is prime is about the number seven, or my belief that the oak tree in my backyard is without leaves is about that oak tree.

Here we must be very clear about an important distinction. Clearly there is such a thing as *indication* or *indicator meaning*.<sup>24</sup> Deer tracks in my backyard indicate that deer have run through it; smoke indicates fire; the height of the mercury column indicates the ambient temperature; buds on the trees indicate the coming of spring. We could speak here of 'natural signs': smoke is a natural sign of fire and the height of the mercury column is a natural sign of the temperature. When one event indicates or is a natural sign of another, there is ordinarily some sort of causal or nomic connection, or at least regular association, between them by virtue of which the first is reliably correlated with the second. Smoke is caused by fire, which is why it indicates fire; measles causes red spots on your face, which is why red spots on your face indicate measles; there is a causal connection between the height of the mercury column and the temperature, so that the latter indicates the former.

The nervous systems of organisms contain such indicators. A widely discussed example: when a frog sees a fly zooming by, the frog's brain (so it is thought) displays a certain pattern of neural firing; we could call such patterns 'fly detectors'. Another famous example: some anaerobic marine bacteria have magnetosomes, tiny internal magnets. These function like compass needles, indicating magnetic north. The direction to magnetic north is downward; hence these bacteria, which can't flourish in the oxygen-rich surface water, move towards the more oxygen-free water at the bottom of the ocean.<sup>25</sup> Of course there are also indicators in human bodies. There are structures that respond in a regular way to blood temperature; they are part of a complex feedback system that maintains a more

<sup>24</sup> See Fred Dretske's *Explaining Behavior* (Cambridge, Mass: MIT Press, 1988), 54 ff. See also William Ramsey's *Using and Abusing Representation: Reassessing the Cognitive Revolution* (forthcoming). Materialists who try to explain how a material structure like a neuronal event can be a belief ordinarily try to do so by promoting indicators to beliefs; see below, pp. 136–141.

<sup>25</sup> Dretske, *Explaining Behavior*, 63.

or less constant blood temperature by inducing, for example, shivering if the temperature is too low and sweating if it is too high. There are structures that monitor the amount of sugar in the blood and its sodium content. There are structures that respond in a regular way to light of a certain pattern striking the retina, to the amount of food in your stomach, to its progress through your digestive system, and so on. Presumably there are structures in the brain that are correlated with features of the environment; it is widely assumed that when you see a tree, there is a distinctive pattern of neural firing (or some other kind of structure) in your brain that is correlated with and caused by it.

Now we can, if we like, speak of ‘content’ here; it’s a free country. We can say that the mercury column, on a given occasion, has a certain content: the state of affairs correlated with its having the height it has on that occasion. We can say, if we like, that those structures in the body that indicate blood pressure or temperature or saline content have a content on a given occasion: whatever it is that the structure indicates on that occasion. We can say, if we like, that the neural structure that is correlated with my looking at a tree has a content: its content, we could say, is what it indicates on that occasion. We can also, if we like, speak of information in these cases: the structure that registers my blood temperature, we can say, carries the information that my blood temperature is thus and so.

What is crucially important to see, however, is that this sort of content or information has nothing as such to do with *belief*, or belief content. There are those who—no doubt in the pursuit of greater generality—gloss over this distinction. Donald T. Cambell, for example, in arguing for the relevance of natural selection to epistemology, claims that “evolution—even in its biological aspects—is a knowledge process”.<sup>26</sup> Commenting on Cambell’s claim, Franz Wuketits explains that

The claim is based on the idea that any living system is a “knowledge-gaining system.” This means that organisms accumulate information about certain properties of their environment. Hence life generally may be described as an information process, or, to put it more precisely, an information-increasing process.<sup>27</sup>

At any rate Wuketits has the grace to put ‘knowledge’ in scare quotes here. Knowledge requires belief; correlation, causal or otherwise, is not belief; information and content of this sort do not require belief. Neither the thermostat nor any of its components believes that the room temperature is thus and so. When the saline content of my blood is too low, neither I nor the structure correlated with that state of affairs (nor my blood) believes the saline content is less than it should be—or, indeed, anything else about the saline content. Indication,

<sup>26</sup> “Evolutionary Epistemology”, in P. A. Schilpp (ed.), *The Philosophy of Karl Popper* (LaSalle: Open Court, 1974), 413.

<sup>27</sup> “Evolutionary Epistemology”, *Biology and Philosophy*, 1/2 (1986), 193.

carrying information, is not belief; indicator content is not belief content, and these structures don't have belief content just by virtue of having indicator content. And now the point here: I am not, of course, claiming that material structures can't have indicator content; obviously they can. What I am claiming is that they can't have belief content: no material structure can be a belief.

Here someone might object as follows. "You say we can't see how a neural event can have content; but in fact we understand this perfectly well, and something similar happens all the time. For there is, after all, the computer analogy. A computer, of course, is a material object, an assemblage of wires, switches, relays, and the like. Now suppose I am typing in a document. Take any particular sentence in the document: say the sentence 'Naturalism is all the rage these days'. That sentence is represented and stored on the computer's hard disk. We don't have to know in exactly what *way* it's stored (by pluses and minuses, or a magnetic configuration, or something else; it doesn't matter). Now the sentence 'Naturalism is all the rage these days' *expresses* the proposition *Naturalism is all the rage these days*. That sentence, therefore, has the proposition *Naturalism is all the rage these days* as its content. But then consider the analogue of that sentence on the computer disk: doesn't it, too, express the same proposition as the sentence it represents? That bit of the computer disk with its pluses and minuses, therefore, has propositional content. But of course that bit of the computer disk is also (part of) a material object (as is any inscription of the sentence in question). Contrary to your claim, therefore, a material object can perfectly well have propositional content; indeed, it happens all the time. But if a computer disk or an inscription of a sentence can have a proposition as content, why can't an assemblage of neurons? Just as a magnetic pattern has as content the proposition *Naturalism is all the rage these days*, so too a pattern of neuronal firing can have that proposition as content. Your claim to the contrary is completely bogus and you should be ashamed of yourself." Thus far the objector.

If the sentence or the computer disk really *did* have content, then I guess the assemblage of neurons could too. But the fact is neither does—or rather, neither has the right kind of content: neither has *original* content; each has, at most, *derived* content. For how does it happen that the sentence has content? It's simply by virtue of the fact that we human beings *treat* that sentence in a certain way, *use* the sentence in a certain way, a way such that if a sentence is used in that way, then it expresses the proposition in question. Upon hearing that sentence, I think of, grasp, apprehend the proposition *Naturalism is all the rage these days*. You can get me to grasp, entertain, and perhaps believe that proposition by uttering that sentence. How exactly all this works is complicated and not at all well understood; but the point is that the sentence has content only because of something *we*, we who are *already* thinkers, do with it. We could put this by saying that the sentence has *secondary* or *derived* content; it has content only because we, we creatures whose thoughts and beliefs already have content, treat it in a certain way. The same goes for the magnetic pattern on the computer disk;

it represents or expresses that proposition because we assign that proposition to that configuration. But of course that isn't how it goes (given materialism) with that pattern of neural firing. That pattern doesn't get its content by way of being used in a certain way by some other creatures whose thoughts and beliefs already have content. If that pattern has content at all, then, according to materialism, it must have *original* or *primary* content. And what it is hard or impossible to see is how it could be that an assemblage of neurons (or a sentence, or a computer disk) could have original or primary content. To repeat: it isn't just that we can't see how it's done, in the way in which we can't see how the sleight of hand artist gets the pea to wind up under the middle shell. It is rather that we can see, to at least some degree, that it can't be done, just as we can see that an elephant can't be a proposition, and that the number seven can't weigh 7 pounds.

### *Parity?*

Peter van Inwagen agrees that it is hard indeed to see how physical interaction among material entities can produce thought: "it seems to me that the notion of a physical thing that thinks is a mysterious notion, and that Leibniz's thought-experiment brings out this mystery very effectively."<sup>28</sup>

Now I am taking this fact as a reason to reject materialism and hence as an argument for dualism. But of course it is a successful argument only if there is no similar difficulty for substance dualism itself. Van Inwagen believes there *is* a similar difficulty for dualism:

For it is thinking itself that is the source of the mystery of a thinking physical thing. The notion of a non-physical thing that thinks is, I would argue, equally mysterious. How any sort of thing could think is a mystery. It is just that it is a bit easier to see that thinking is a mystery when we suppose that the thing that does the thinking is physical, for we can form mental images of the operations of a physical thing and we can see that the physical interactions represented in these images—the only interactions that can be represented in these images—have no connection with thought or sensation, or none we are able to imagine, conceive or articulate. The only reason we do not readily find the notion of a non-physical thing that thinks equally mysterious is that we have no clear procedure for forming mental images of non-physical things. (*Metaphysics*, 176)

So dualism is no better off than materialism; they both have the same problem. But what precisely *is* this problem, according to van Inwagen? "we can form mental images of the operations of a physical thing and we can see that the physical interactions represented in these images—the only interactions that can be represented in these images—have no connection with thought or sensation or none we are able to imagine, conceive or articulate." As I understand van Inwagen here, he is saying that we can imagine physical interactions or changes in a physical thing; but we can see that the physical interactions represented in

<sup>28</sup> *Metaphysics* (2nd edn; Boulder, Colo.: Westview, 2002), 176.

those images have no connection with thought. We can imagine neurons in the brain firing; we can imagine electrical impulses or perhaps clouds of electrons moving through parts of neurons, or whole chains of neurons; we can imagine neural structures with rates of fire in certain parts of the structure changing in response to rates of fire elsewhere in or out of that structure: but we can see that these interactions have no connection with thought. Now I'm not quite sure whether or not I can imagine electrons, or their movements, or electrical impulses; but it does seem to me that I can see that electrical impulses and the motions of electrons, if indeed there are any such things, have nothing to do with thought.

Another way to put van Inwagen's point: no change we can imagine in a physical thing could be a mental change, that is, could constitute thought or sensation, or a change in thought or sensation. But then we can't imagine a physical thing's thinking: that is, we can't form a mental image of a physical thing thinking. And this suggests that the problem for materialism is that we can't form a mental image of a material thing thinking. But the same goes, says van Inwagen, for an immaterial thing: we also can't imagine or form a mental image of an immaterial thing thinking. Indeed, we can't form a mental image of any kind of thinking thing: "My point", he says, "is that nothing could possibly count as a mental image of a thinking thing" (p. 177). Materialism and dualism, therefore, are so far on a par; there is nothing here to incline us to the latter rather than the former.

Thus far van Inwagen. The thought of a physical thing's thinking, he concedes, is mysterious; that is because we can't form a mental image of a physical thing's thinking. But the thought of an immaterial thing's thinking is equally mysterious; for we can't form a mental image of that either. This, however, seems to me to mislocate the problem for materialism. What inclines us to reject the idea of a physical thing's thinking is not just the fact that we can't form a mental image of a physical thing's thinking. There are plenty of things of which we can't form a mental image, where we're not in the least inclined to reject them as impossible. As Descartes pointed out, I can't form a mental image of a chiliagon, a 1,000-sided rectilinear plane figure (or at least an image that distinguishes it from a 100-sided rectilinear plane figure); that doesn't even suggest that there can't be any such thing. I can't form a mental image of the number 79's being prime: that doesn't incline me to believe that the number 79 could not be prime; as a matter of fact I know how to prove that it *is* prime. The fact is I can't form a mental image of the number 79 at all—or for that matter of any number; this doesn't incline me to think there aren't any numbers.

Or is all that a mistake? Is it really true that I can't form a mental image of the number seven, for example? Maybe I *can* form an image of the number seven; when I think of the number seven, sometimes there is a mental image present; it's as if one catches a quick glimpse of a sort of partial and fragmented numeral 7; we could say that I'm appeared to numeral-7ly. When I think of the

actual world, I am sometimes presented with an image of the Greek letter alpha; when I think of the proposition *All men are mortal* I am sometimes presented with a sort of fleeting, fragmentary, partial image of the corresponding English sentence. Sets are nonphysical, but maybe I can imagine the pair set of Mic and Martha; when I try, it's like I catch a fleeting glimpse of curly brackets, enclosing indistinct images that don't look a whole lot like Mic and Martha. But is that really imagining the number seven, or the actual world, or the pair set of Mic and Martha? Here I'm of two minds. On the one hand, I'm inclined to think that this isn't imagining the number seven at all, but instead imagining something connected with it, namely the numeral 7 (and the same for the actual world and the set of Mic and Martha). On the other hand I'm a bit favorably disposed to the idea that that's just how you imagine something like the number seven; you do it by imagining the numeral 7. (Just as you state a proposition by uttering a sentence or uttering certain sounds.) So I don't really know what to say. Can I or can't I imagine non-physical things like numbers, propositions, possible worlds, angels, God? I'm not sure.

What is clear, here, is this: if imagining the numeral 7 is sufficient for imagining the number seven, then imagining, forming mental images of, has nothing to do with possibility. For in this same way I can easily imagine impossibilities. I can imagine the proposition *all men are mortal* being red: first I just imagine the proposition, for example, by forming a mental image of the sentence 'All men are mortal', and then I imagine this sentence as red. I think I can even imagine that elephant's being a proposition (I imagine the relevant sentence and then imagine it in the shape of an elephant). David Kaplan once claimed he could imagine his refuting Gödel's Incompleteness Theorem: he imagined the *Los Angeles Times* carrying huge headlines: 'UCLA PROF REFUTES GÖDEL; ALL REPUTABLE EXPERTS AGREE'. In this loose sense, most anything can be imagined; but then the loose sense has little to do with what is or isn't possible. So really neither the loose nor the strong sense of 'imagining' (neither the weak nor the strong version of imagination) has much to do with possibility. There are many clearly possible things one can't imagine in the strong sense; in the weak sense, one can imagine many things that are clearly impossible.

What is it, then, that inclines me to think a proposition can't be red, or a horse, or an even number? The answer, I think, is that one can just see upon reflection that these things are impossible. I can't form a mental image of a proposition's having members; but that's not why I think no proposition has members; I also can't form a mental image of a set's having members. It's rather that one sees that a set is the sort of thing that (null set aside) has members, and a proposition is the sort of thing that cannot have members. It is the same with a physical thing's thinking. True, one can't imagine it (in the strong sense). The reason for rejecting the idea, thinking it impossible, however, is not that one can't imagine it. It's rather that on reflection one can see that a physical object just can't do that sort of thing. I grant that this isn't as clear and obvious, perhaps, as

that a proposition can't be red; some impossibilities (necessities) are more clearly impossible (necessary) than others. But one can see it to at least a significant degree. Indeed, van Inwagen might be inclined to endorse this thought; elsewhere he says: "Leibniz's thought experiment shows that when we carefully examine the idea of a material thing having sensuous properties, it seems to be an impossible idea."<sup>29</sup> But (and here is the important point) the same clearly doesn't go for an immaterial thing's thinking; we certainly can't see that no immaterial thing can think. (If we could, we'd have a quick and easy argument against the existence of God: no immaterial thing can think; if there were such a person as God, he would be both immaterial and a thinker; therefore . . .).

Van Inwagen has a second suggestion:

In general, to attempt to explain how an underlying reality generates some phenomenon is to construct a representation of the working of that underlying reality, a representation that in some sense "shows how" the underlying reality generates the phenomenon. Essentially the same considerations as those that show that we are unable to form a mental image that displays the generation of thought and sensation by the workings of some underlying reality (whether the underlying reality involves one thing or many, and whether the things it involves are physical or non-physical) show that we are unable to form any sort of representation that displays the generation of thought and sensation by the workings of an underlying reality. (*Metaphysics*, 177–8)

The suggestion is that we can't form an image or any other representation displaying the generation of thought by way of the workings of an underlying reality; hence we can't see how it can be generated by physical interaction among material objects such as neurons. This much seems right—at any rate we certainly can't see how thought could be generated in that way. Van Inwagen goes on to say, however, that this doesn't favor dualism over materialism, because we also can't see how thought can be generated by the workings of an underlying *non*-physical reality. And perhaps this last is also right. But here there is an important dissimilarity between dualism and materialism. The materialist thinks of thought as generated by the workings of an underlying reality—that is, by the physical interaction of such physical things as neurons; the dualist, however, typically thinks of an immaterial self, a soul, a thing that thinks, as *simple*. An immaterial self doesn't have any parts; hence, of course, thought isn't generated by the interaction of its parts. Say that a property P is *basic* to a thing x if x has P, but x's having P is not generated by the interaction of its parts. Thought is then a basic property of selves, or better, a basic activity of selves. It's not that (for example) there are various underlying immaterial parts of a self whose interaction produces thought. Of course a self stands in causal relation to its body: retinal stimulation

<sup>29</sup> "Dualism and Materialism: Athens and Jerusalem?", *Faith and Philosophy*, 12/4 (Oct. 1995), 478. That is (I take it), it seems to be necessary that material things don't have such properties. Van Inwagen's examples are such properties as *being in pain* and *sensing redly*; the same goes, I say, for properties like *being the belief that p* for a proposition p.

causes a certain sort of brain activity which (so we think) in turn somehow causes a certain kind of experience in the self. But there isn't any *way* in which the self produces a thought; it does so immediately. To ask, "How does a self produce thought?" is to ask an improper question. There isn't any how about it.

By way of analogy: consider the lowly electron. According to current science, electrons are simple, not composed of other things. Now an electron has basic properties, such as having a negative charge. But the question, "How does an electron manage to have a charge?" is an improper question. There's no how to it; it doesn't do something else that results in its having such a charge, and it doesn't have parts by virtue of whose interaction it has such a charge. Its having a negative charge is rather a basic and immediate property of the thing (if thing it is). The same is true of a self and thinking: it's not done by underlying activity or workings; it's a basic and immediate activity of the self. But then the important difference, here, between materialism and immaterialism is that if a material thing managed to think, it would have to be by way of the activity of its parts: and it seems upon reflection that this can't happen.<sup>30</sup> Not so for an immaterial self. Its activity of thinking is basic and immediate. And it's not the case that we are inclined upon reflection to think this can't happen—there's nothing at all against it, just as there is nothing against an electron's having a negative charge, not by virtue of the interaction of parts, but in that basic and immediate way. The fact of the matter then is that we can't see how a material object can think—that is, upon reflection it seems that a material object can't think. Again, not so for an immaterial self.

True, as van Inwagen says, thought can sometimes seem mysterious and wonderful, something at which to marvel. (Although from another point of view it is more familiar than hands and feet.) But there is nothing here to suggest that it can't be done. I find myself perceiving my computer; there is nothing at all, here, to suggest impossibility or paradox. Part of the mystery of thought is that it is wholly unlike what material objects can do: but of course that's not to suggest that it can't be done at all. Propositions are also mysterious and have wonderful properties: they manage to be about things; they are true or false; they can be believed; they stand in logical relations to each other. How do they

<sup>30</sup> But couldn't a material thing also just directly think, without depending on the interaction of its parts? According to Pierre Cabanis, "The brain secretes thought as the liver secretes bile"; couldn't we think of this as the brain (or, if you like, the whole organism) directly thinking, not by way of the interaction of its parts? Well, if that's how a brain thinks, it isn't like the way a liver secretes bile; the latter certainly involves the liver's having parts, and those parts working together in the appropriate way. Further, the idea of a physical thing's thinking without the involvement of its parts is even more clearly impossible than that of a physical thing's thinking by virtue of the interaction of its parts. Aren't those neurons in the brain supposed to be what enables it to think? You might as well say that a tree or my left foot thinks. Consider any non-elementary physical object—a tree, an automobile, perhaps a horse: such a thing does what it does by virtue of the nature and interaction of its parts. Are we to suppose that some physical object—a brain, let's say—does something like thinking apart from involvement of its parts? Talk about appealing to magic!

manage to do those things? Well, certainly not by way of interaction among material parts. Sets manage, somehow, to have members—how do they do a thing like that? And why is it that a given set has just the members it has? How does the unit set of Neil Armstrong manage to have exactly *him* as a member? What mysterious force, or fence, keeps Leopold out of that set? Well, it's just the nature of sets to be like this. These properties can't be explained by way of physical interactions among material parts, but that's nothing at all against sets. Indeed, these properties can't be explained at all. Of course if you began with the idea that everything has to be a material object, then thought (and propositions and sets) would indeed be mysterious and paradoxical. But why begin with that idea? Thought is seriously mysterious, I think, only when we assume that it would have to be generated in some physical way, by physical interaction among physical objects. That is certainly mysterious; indeed it goes far beyond mystery, all the way to apparent impossibility. But that's not a problem for thought; it's a problem for materialism.

## 2. THE BEARING OF CHRISTIAN BELIEF

As I said above (p. 100) there are three ways in which Christian belief is relevant to the issue of dualism vs. materialism. First, there is Scripture and perhaps also creedal and conciliar declaration. Second, Christian belief is relevant to the epistemology of the situation, and that in two ways: (a) given Christian theism, we know that it is at any rate possible that there be immaterial thinking things, since God Himself is such a thing, and (b) these considerations from the Christian faith are powerfully relevant to the objections to dualism and arguments for materialism. Finally, certain crucial Christian doctrines (for example, Incarnation and the resurrection of the dead) fit better—much better, I'd say—with dualism than with materialism. Here I'll confine myself to the second,<sup>31</sup> beginning with just a brief remark on the first.

The Scripture obviously contains a great deal that is relevant to our question; and in my opinion these scriptural declarations heavily favor dualism. I am no Scripture scholar, however, and hence am not well qualified to develop this case. Fortunately enough, then, there is clear and authoritative work by someone who does have credentials in this area: John Cooper's philosophically sensitive examination of the bearing on biblical teaching on our question.<sup>32</sup> I have little to add to Cooper's balanced and nuanced discussion;<sup>33</sup> I would

<sup>31</sup> For a discussion of the bearing of Christian belief on the doctrine of Incarnation, see my "On Heresy, Mind, and Truth", *Faith and Philosophy*, 16/2 (April 1999).

<sup>32</sup> *Body, Soul, and Life Everlasting: Biblical Anthropology and the Monism–Dualism Debate*, 2nd edn. with a new preface (Grand Rapids, Mich.: Eerdmans Publishing Co., 2000).

<sup>33</sup> In the passages with which I am concerned, Cooper is arguing that Paul asserts or presupposes, not merely that a person is not identical with his body, but that in addition there is an 'intermediate

simply like to call your attention to three Pauline passages, together with Cooper's comments on them. These passages (among many others) are, I believe, vastly more smoothly and plausibly understood in terms of dualism than in terms of materialism. People have indeed come up with interpretations in accord with materialism; these interpretations, in my opinion, are strained and implausible.

The first passage is 2 Corinthians 5: 6–9 (Cooper's comments: pp. 141–9):

Therefore we are always confident and know that as long as we are at home in the body, we are away from the Lord. We live by faith, not by sight. We are confident, I say, and would prefer to be away from the body and at home with the Lord. So we make it our goal to please him, whether we are at home in the body or away from it.

Second, a parallel passage: Philippians 1: 21–4 (Cooper's comments pp. 151–6):

For to me, to live is Christ and to die is gain. If I am to go on living in the body, this will mean fruitful labor for me. Yet what shall I choose? I do not know. I am torn between the two: I desire to depart and be with Christ, which is better by far; but it is more necessary for you that I remain in the body.

Third, 2 Corinthians 12: 1–4 (Cooper's comments pp. 88, 28, 149–51):

I must go on boasting. Although there is nothing to be gained, I will go on to visions and revelations from the Lord. I know a man in Christ who fourteen years ago was caught up to the third heaven. Whether it was in the body or out of the body I do not know—God knows. And I know that this man—whether in the body or apart from the body I do not know, but God knows—was caught up to Paradise.<sup>34</sup>

### Parity Again

Turning now to the epistemological considerations, return first to the discussion of parity (pp. 113–8 above). Peter van Inwagen concedes that the idea of a thinking material thing seems to be an impossible idea; but he thinks or is inclined to think that the same goes for the idea of an immaterial thing's thinking. Here I believe he is mistaken: as far as I can see, there is no apparent impossibility in the idea of an immaterial thing's thinking. It is not the case that when we consider the state of affairs consisting of an immaterial thing's thinking, that state of affairs

state' between death and resurrection during which a person exists disembodied. I'm not concerned to argue for or against the claim that human persons exist disembodied at some points in their careers; I want only to call attention to the point that in these passages Paul certainly appears to endorse dualism.

<sup>34</sup> Cooper doesn't comment on 2 Peter 1: 13–14, a non-Pauline passage expressing the same sentiment: "I think it is right to refresh your memory as long as I live in the tent of this body, because I know that I will soon put it aside, as our Lord Jesus Christ has made clear to me." Here Peter pretty clearly distinguishes himself from "the tent of this body" and thinks of death as putting aside, separation from, the body.

seems impossible. Nor is the dualist committed to the existence of an underlying immaterial reality whose workings somehow generate thought; that may be an impossible idea, but the dualist isn't committed to it. (Of course I agree that in the strong sense of 'imagine' (above, pp. 113–115) it isn't possible to imagine an immaterial thinking thing.) But suppose van Inwagen were right; suppose the state of affairs of an immaterial thing's thinking seemed, on reflection, quite as clearly impossible as that of a material thing's thinking. What would follow? Would it follow that these two states of affairs are on an epistemic par?

Not at all. For suppose we take Christian theism seriously. Then we are already committed to the existence of a thinking immaterial being: God himself. (We'll probably also be inclined to suppose that there are other immaterial thinkers: angels, perhaps, and Satan and his minions.) The appearance of impossibility in an immaterial object's thinking, if there were such an appearance, would therefore be an illusion, a sort of inexplicable tendency on our part to form a suite of false beliefs, all related to the false intuition that it is not possible that an immaterial thing think. Here, then, is a way in which Christian theism is related to the question of materialism vs. dualism: even if (contrary to fact, as I see it) it did seem on reflection impossible that an immaterial thing think, so that dualism and materialism would be on a par in this regard, Christian theism would lead us to see that there isn't epistemic parity here after all. What it would lead us to think instead is that the apparent impossibility of an immaterial thing's thinking is an illusion.

### Objections to Dualism (Arguments for Materialism)

The above arguments for dualism and others like them are, I believe, powerful arguments. Like philosophical arguments generally, however, they are not of that wholly apodictic and irrefragable character Kant liked to claim for his arguments; they are defeasible. It is possible to disregard or downgrade the intuitions of possibility and impossibility to which they appeal, just as it is possible to produce convoluted interpretations of the relevant scriptural evidence. Further, if there were really powerful arguments against dualism or for materialism, then perhaps the appropriate course would be to embrace materialism, or to take refuge in agnosticism. But *are* there any such powerful arguments? You might think so. As Paul Churchland, Jaegwon Kim, and many others say, dualism is the natural, baseline belief of humankind, not an invention of Plato or Descartes; but according to Daniel Dennett, "The prevailing wisdom, variously expressed and argued for, is *materialism*: there is only one sort of stuff, namely *matter*—the physical stuff of physics, chemistry, and physiology—and the mind is somehow nothing but a physical phenomenon. In short, the mind is the brain."<sup>35</sup> Presumably there must be some pretty powerful arguments to move

<sup>35</sup> Dennett, *Explaining Consciousness*, 34.

so many from the baseline position of dualism to materialism. Paul Churchland concurs, “Arguments like these have moved most (but not all) of the professional community to embrace some form of materialism.”<sup>36</sup> Where are these powerful arguments? The fact is there aren’t any. Most of them seem to have very little force; even the best doesn’t survive a closer look. Here there is a clear bearing of Christian theism: the fact is, I think, none of the usual objections to dualism has any purchase at all on someone committed to Christian theism. In particular, the most widely cited and influential argument against dualism—the claim that an immaterial object can’t cause changes in the hard, heavy, massive physical and material world—should carry no weight whatever with someone so committed.

Of course many arguments have been proposed for materialism; I’ll restrict myself to seven that seem to be among the most important and significant.

### *Soul Stuff?*

The first argument needn’t detain us long. According to Michael Levin and others (i.e. Churchland and Dennett), substance dualism fails because the stuff a self is supposed to be made of is mysterious, or obscure, or even inconceivable:

The trouble, I suggest, is this: we can say what sort of stuff a material thing is an individual piece of, while no one has any idea of the sort of stuff a self is an individual piece of. . . . It is in this sense that it is impossible to form an idea of what the substance dualist’s self is. While there are descriptions that can identify a self, we cannot refer to it as a P of S, for we do not know and evidently cannot imagine the *stuff* it is a piece of, or the sort of piece it could be.<sup>37</sup>

But this objection is massively unimpressive. First, note that it would equally be an objection to propositions, properties, states of affairs, sets, numbers, and other abstract objects. Consider, for example, the proposition *All men are mortal*: we don’t know and can’t imagine the sort of stuff that proposition is made out of or is a piece of. More poignantly, from Levin’s perspective, the same would go for many of the entities postulated by contemporary physics: what is the stuff an electron is a piece of? According to the most widely accepted theories, an electron is a perturbation in a field—so is the stuff in question a *field*? But is a field ‘stuff’? Or a piece of stuff?

More important, though: the objection rests on a misunderstanding. Selves, according to the dualist, aren’t made of any stuff at all, not even very fine, filmy, gossamer, ghostly soul stuff. Levin apparently assumes that everything there is has to be made of stuff of some kind or other: but why think a thing like that? Propositions, properties, sets, possible worlds—these things are not made of

<sup>36</sup> P. Churchland, *Matter and Consciousness* (Cambridge, Mass.: MIT Press, 1984), 21.

<sup>37</sup> M. Levin, *Metaphysics and the Mind–Body Problem* (Oxford: Oxford University Press, 1979), 79.

stuff and are not pieces of stuff. So why think selves, if immaterial, would have to be made of stuff? Perhaps Levin and others will reply that it's perfectly fine and good for abstract objects like sets and propositions not to be made of stuff, but concrete objects can't enjoy that luxury; they can only be pieces of stuff. But again, why think a thing like that? And once more there is contemporary physics: electrons and fields do not appear to be pieces of stuff; but are they not concrete? Well, perhaps that is part of the problem posed by the mysterious character of the entities postulated by contemporary physics. These things are mysterious in a variety of ways; perhaps we shouldn't be surprised that they are an enigma in this way as well.

In any event, there is a much more decisive answer from the perspective of Christian (or other) theism: God, clearly enough, is not an abstract object; equally clearly, God is not made of any stuff and is not himself a piece of stuff. From a Christian perspective, therefore, this objection to dualism has no bite at all; the Christian is already committed to the existence of concrete beings that are not pieces of stuff.<sup>38</sup> But even apart from such commitment: would anyone seriously want to hold that we have here a significant new argument for atheism? Could anyone argue with a straight face that God, if he existed, would be a concrete object that wasn't a piece of stuff; but every concrete object must be a piece of stuff; therefore there is no such person as God?

### *Dualism Unscientific?*

Dennett, Churchland, and others complain that dualism should be rejected because it is *unscientific*:

There is the lurking suspicion that the most attractive feature of mind stuff is its promise of being so mysterious that it keeps science at bay forever.

This fundamentally antiscientific stance of dualism is, to my mind, its most disqualifying feature, and is the reason why in this book I adopt the apparently dogmatic rule that dualism is to be avoided *at all costs*.<sup>39</sup>

But our question here is whether dualism is *true*, not whether it, or more likely its proponents, are properly reverential towards science. What I claim for dualism is only that it is true, not that those who embrace it are of good character, or are appropriately deferential towards science, or in other ways estimable. Perhaps those who promulgate dualism adopt wholly unacceptable stances, even going so far as *lese majesty* towards modern science itself; perhaps they are in still

<sup>38</sup> There is also a sort of general and widespread impression that the very idea of an immaterial concrete substance (an immaterial self or thinker) is weird or crazy or implausible. This shows, once more, the importance of fashion and zeitgeist in philosophy; prior to (e.g.) 100 years ago that idea wasn't considered weird; and we haven't learned anything in the last 100 years to show that it really is weird. In any event, however, from a Christian or theistic perspective the idea is anything but weird; the first being of the entire universe is an immaterial thinking substance.

<sup>39</sup> Dennett, *Explaining Consciousness*, 37.

other ways wholly objectionable: what has that to do with the truth or falsehood of dualism? Materialists like Daniel Dennett sometimes adopt an unpleasantly triumphalist tone; Dennett also suggests that Baptists should be confined to zoos, lest they contaminate the rest of us with their noxious views on evolution, etc.:<sup>40</sup> should we conclude that materialism must be rejected?

But the fact is there is no reason at all to think dualists do or must display anything as heinous as an unscientific attitude—at least they need not do so just by virtue of being dualists. We have discovered many fascinating things about the brain and its organization, about the structure and behavior of neurons, about the ways in which damage to various parts of the brain is correlated with mental and physical disorders, about the correlation between certain kinds of mental activity (memory, vision) and increased blood flow and electrical activity displayed in certain areas of the brain, (leading us to say that those activities are ‘located’ in those areas), and much else. Need a dualist reject these discoveries? Need she decry, downgrade, denigrate, or disapprove of the scientific activity that leads to these discoveries? But is this a serious question? *Of course* she needn’t do those things. Indeed, there is no reason at all why dualists can’t enthusiastically join the scientific enterprise here. The fact is some dualists have done exactly that, and have been leaders in the field, with no conflict whatever with their dualistic views and no compromise whatever to their intellectual integrity.<sup>41</sup> According to dualism, I am an immaterial object intimately linked to a body; nothing follows with respect to whether and in what way appropriate brain condition is a necessary condition of proper mental function (see below, pp. 133–5). Therefore nothing prevents a dualist from being wholly enthusiastic about brain science. This whole issue is nothing but a red herring.

### *Explanatory Impotence?*

Paul Churchland objects that dualism is *explanatorily impotent*:

Compare now what the neuroscientist can tell us about the brain and what he can do with that knowledge, with what the dualist can tell us about spiritual substance, and what he can do with those assumptions. Can the dualist tell us anything about the internal constitution of mind-stuff? Of the nonmaterial elements that make it up? Of the laws that govern their behavior? Of the mind’s structural connections with the body? Of the manner of its operations? Can he explain human capacities and pathologies in terms of its structures and its defects? The fact is, the dualist can know none of these things, because no detailed theory of mind-stuff has even been formulated. Compared to the

<sup>40</sup> *Darwin’s Dangerous Idea* (New York; Simon and Schuster, 1995), 515–16.

<sup>41</sup> An example would be Wilder Penfield (who made impressive contributions to the ‘localization’ of memory); see his *The Mystery of the Mind* (Princeton: Princeton University Press, 1975). Another would be Nobel Prize winner John Eccles; see his *Facing Reality: Philosophical Adventures by a Brain Scientist* (New York and Berlin: Springer, 1970) and *The Wonder of Being Human* (New York: Springer, 1984).

rich resources and explanatory successes of current materialism, dualism is less a theory of mind than it is an empty space waiting for a genuine theory of mind to be put in it.<sup>42</sup>

Here we have once more the mistaken idea that the dualist is committed to some kind of soul stuff. But there are two further and fundamental problems with Churchland's objection. First, this might be a good objection to a *scientific hypothesis* to which there was a much more fruitful and explanatorily powerful alternative. But why think dualism is a scientific hypothesis? What Churchland offers is an objection to dualism only if the latter is proposed as hypothesis, something designed to explain the phenomena, something that gets whatever warrant it enjoys by virtue of the excellence of the explanation it provides. But why think of dualism like this? Perhaps the dualist accepts dualism because she believes, first, that there is such a thing as she herself, and secondly, that she couldn't be a material object; she knows she is conscious, for example, and believes that no material object can be conscious. The question how much dualism does or doesn't explain is irrelevant; maybe it doesn't explain much of anything, but why should that be anything against it?<sup>43</sup> I believe that propositions, unlike sets, don't have members; maybe that doesn't explain much, but so what? It's not being proposed as a scientific hypothesis. Similarly an atheologian might complain that many characteristic Christian doctrines—Trinity and Incarnation, for example—aren't good explanations of the phenomena. But that would be an objection only if those doctrines were proposed as hypotheses, explanations of some range of phenomena; and they aren't.<sup>44</sup>

Secondly, the objection seems to suggest that the materialist can or does have an explanation of all these things, but the dualist doesn't or can't. That is of course mistaken; as I argued above, brain science is just as open to the dualist as to the materialist. Well, perhaps the idea is that the materialist can explain these things *as a materialist*, but the dualist can't do so as a dualist. But this looks like an illicit attempt to credit materialist metaphysics with the warrant enjoyed by the relevant science. It isn't as a materialist metaphysician that the materialist has these explanations; it is rather as someone who knows something about the brain and its connections with human behavior and pathologies. And of course there is nothing to prevent the dualist from knowing the very same things. Happily, you don't have to be a materialist to engage in brain science. Indeed, perhaps the shoe is on the other foot. Brain science investigates, among other things, the relation between brain activity and mental activity. Clearly it is arguable (proved by the fact that I've been arguing it) that if materialism were true, there wouldn't be any

<sup>42</sup> Churchland, *Matter and Consciousness* (Mass.: MIT Press, 1984), 19.

<sup>43</sup> Of course the term 'explain' is something of a weasel word, and explanations are multifarious. We can imagine a dualist suggestion that, given that material objects can't be conscious, think, believe, make decisions, take actions, and the like, dualism 'explains' the fact that human persons can do those things. That would be a slightly different but analogically connected sense of 'explain'; and in that sense, says the dualist, dualism can explain these things and materialism cannot.

<sup>44</sup> See my "Is Theism Really a Miracle?", *Faith and Philosophy*, 3/2 (1986).

such thing as mental activity; hence, from that perspective, it's the materialist who can't sensibly engage in brain science, at least of the sort that investigates those connections.

### *Conservation of Energy?*

Still another scientific or quasi-scientific objection: according to Daniel Dennett and others, dualism violates the scientifically approved principle of conservation of energy:

concentrate on the return signals, the directives from mind to brain. These, *ex hypothesi*, are not physical; they are not light waves or sound waves or cosmic rays or streams of subatomic particles. No physical energy or mass is associated with them. How, then, do they get to make a difference to what happens in the brain cells they must affect, if the mind is to have any influence over the body? A fundamental principle of physics is that any change in the trajectory of any physical entity is an acceleration requiring the expenditure of energy, and where is this energy to come from? It is this principle of the conservation of energy that accounts for the physical impossibility of "perpetual motion machines", and the same principle is apparently violated by dualism. This confrontation between quite standard physics and dualism has been endlessly discussed since Descartes's own day, and is widely regarded as the inescapable and fatal flaw of dualism.<sup>45</sup>

Here Dennett conflates two separate objections to dualism: first, the claim that an immaterial substance can't have causal consequences in the hard, ponderous, massive physical world, so that if dualism were true, human beings would be unable to act in the physical world; and second that the principle of conservation of energy prohibits an immaterial object from acting in the physical world. I'll turn to the first below; here I am concerned with the second. Note first that, again, the theist is already committed to the thought that an immaterial substance—God—can (indeed does) act in the physical world. God has created the world, and also sustains it. Further, according to Christian doctrine, God does much more; for example, he raised Jesus from the dead. And of course many Christians believe God has acted in the world on many occasions, enabling the Israelites to cross the Red Sea, appearing to the apostle Paul, multiplying the loaves and fishes, and much else. Indeed, many Christians believe that God is at present constantly active in the world and constantly active in our lives, strengthening us in time of trouble, offering grace, answering prayers. Clearly this objection, if it has any merit, is as much an objection to Christian belief as to dualism.

*Does* it have any merit? In a word, No. It is perfectly possible for God to create *ex nihilo* a full-grown horse in the center of the Notre Dame campus without in any way violating the conservation principles. God says: "Let there be a horse in the middle of the North Quad!" The horse suddenly appears in the

<sup>45</sup> *Consciousness Explained* (Boston: Little, Brown), 35.

middle of the quad; there need be no violation of conservation of energy. Clearly this needn't violate *global* conservation; for he could deduct an equal amount of energy elsewhere in the universe; the total energy of the system, that is, the universe, would then remain constant. But of course there is local conservation as well as global; and it is harder to see how there could be local conservation of energy if God created that horse *ex nihilo*. That is because it's not easy to find, for each relevant system, an analogue of creating a horse in one part of the universe and deducting the appropriate amount of energy elsewhere. So perhaps creating a horse *ex nihilo* is incompatible with local conservation: if God were to create that horse, energy would fail to be conserved in at least one system.

It doesn't follow, however, that God's creating that horse is precluded by any of the conservation laws of physics or that his doing so violates those laws. That is because the conservation laws are deduced from Newton's Laws; those laws are conditionals whose antecedents include the condition that the system in question is closed; the conservation laws—of momentum, charge, mass, energy, mass/energy, etc.—are therefore said to hold for *closed* or *isolated* systems. Thus Sears and Zemansky,

This is the principle of conservation of linear momentum: When no resultant external force acts on a system, the total momentum of the system remains constant in magnitude and direction.

More generally,

*The internal energy of an isolated system remains constant.* This is the most general statement of the *principle of conservation of energy*. The internal energy of an *isolated* system cannot be changed by any process (mechanical, electrical, chemical, nuclear, or biological) taking place *within* the system. The energy of a system can be changed only by a flow of heat across its boundary, or by the performance of work. (If either takes place, the system is no longer isolated.)<sup>46</sup>

But of course a system—the physical universe, say—in which God creates *ex nihilo* a full-grown horse is not, obviously, a closed or isolated system. It is clearly not one that is subject to no resultant external force. Therefore the conservation laws do not imply that the quantity in question remains constant in it. More specifically, from the system's being closed it follows that the relevant reference frame is inertial, and hence that the Lagrangian (roughly, a function giving the difference between the kinetic and potential energy of the system) of the system is independent of time (its partial derivative with respect to time is zero). It also follows that the Lagrangian is unaffected by a translation of the entire system in space.<sup>47</sup> But these conditions can't both hold for any system in which this horse

<sup>46</sup> *University Physics* (Reading, Mass: Addison-Wesley, 1964), 186 (bold and italics removed from original), and 415 (italics in original).

<sup>47</sup> See Marion and Thornton, *Classical Dynamics*, 4th edn (New York: Harcourt Brace College Publishers, 1995), 217, 219. (I am indebted to Brian Pitts for this reference, and for help throughout this section.)

suddenly appears. For example, if the space of the system is just the space into which the horse is suddenly introduced, the Lagrangian of the system will depend on time; it will assume different values before and after the horse is created.

The same considerations clearly apply to Dennett's claim that dualism (taken as involving the claim that an immaterial self can cause effects in the physical world) is incompatible with the law of the conservation of energy. He neglects the fact that the law in question applies only to closed systems, ones not subject to any outside force. This condition clearly won't hold for any physical system—my body, or brain, or part of my brain—in which an immaterial self causes a change. This objection, therefore, is wholly without force. It's not that it gives one some reason, perhaps only a weak reason, for rejecting dualism; it provides no reason at all.<sup>48</sup>

### *Can an Immaterial Substance have Causal Consequences in the Material World?*

Dennett appears to be confusing conservation of energy with the (alleged) causal closure of the physical—the idea, as he puts it, that “anything that can move a physical thing is a physical thing” (*Consciousness Explained*, 35.) Strictly speaking, this is not an objection to dualism as I defined it: the thought that human beings are not material objects but are immaterial substances. That view, just as it stands, doesn't entail that (human) immaterial substances *can* cause effects in the physical world. Dualism as thus defined is compatible, first, with *occasionalism*, the doctrine embraced by Malebranche and others. According to occasionalism, it is only God who causes changes in the physical world, but, for example, God takes my willing to raise my arm as an occasion to cause my arm to rise. Dualism is compatible, secondly, with Leibniz's *pre-established harmony*, according to which mental events don't cause physical events, but from time immemorial God has instituted a correlation between mental events, such as my willing or trying to raise my arm, and physical events, such as my arm's rising. Dennett's objection is really to *dualistic interactionism*, according to which human beings are immaterial substances that can act, can cause changes, in the physical world. I don't mean to argue against either pre-established harmony or occasionalism, and in fact the latter has its attractions.

So consider the current objection as directed against dualistic interactionism; even so, it still has no force. First, the doctrine or dogma of the causal closure

<sup>48</sup> Another definition of closure for a system: a system is closed if and only if there is no flow of energy across its boundaries. As Sears and Zemansky put it above, “The energy of a system can be changed only by a flow of heat across its boundary, or by the performance of work.” But this is clearly not a correct definition of closure; if an external agent causes something to occur within a system S (creates a horse within S, say) but without causing a flow of energy across the boundaries of S, S is still clearly not a closed system. This definition would be accurate up to logical equivalence only if it were impossible, in the broadly logical sense, that God create a horse within a system without causing an energy flow across the boundaries of the system.

of the physical is not a deliverance of current science: it is more like an article of faith or perhaps a pious hope on the part of materialists. Science says nothing at all to imply that there aren't any immaterial substances, and nothing at all to imply that if there are some, they can't cause changes in the physical world. Is there then any reason to believe this dogma? Although not strictly relevant to my case, it is of interest to note that causal closure depends heavily on the correct analysis or account of causation. On one of Hume's accounts, causation is fundamentally a matter of constant conjunction (with the 'cause' preceding the effect): "we may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.*"<sup>49</sup> But of course there is no reason in the world why a mental event (i.e., an event involving only an immaterial substance) shouldn't be related, in this way, to a physical event (one involving only a physical substance). There is no reason in the world, therefore, why my willing to raise my arm shouldn't cause my arm to rise, even if I am an immaterial substance.

Immediately after the above passage from the *Inquiry*, Hume proposes a different account of causation: "Or in other words, *where, if the first object had not been, the second never had existed.*" David Lewis presents a fuller version of this second account. Say that an event *d* depends causally on an event *e* iff the counterfactual *If e had not occurred, d would not have occurred* is true. Then

Let *c, d, e*, be a finite sequence of actual particular events such that *d* depends causally on *c, e* on *d*, and so on throughout. Then this sequence is a *causal chain*. Finally, one event is a *cause* of another iff there exists a causal chain leading from the first to the second.<sup>50</sup>

Again, it is obvious that there can be this kind of counterfactual relation between mental events and physical events. Suppose, for example, that I am an immaterial substance and that something like Leibniz's pre-established harmony is the truth of the matter: from before the foundation of the world, God has decreed and established a correlation between my mental states—my tryings and willings, my efforts and endeavors—and what happens in the physical world. I will to raise my arm; my arm rises; if I had not willed to do so, it would not have risen. On the Lewisian account, therefore, my willing to raise my arm causes it to rise, and this despite Leibniz's explicit aim to propose a theory according to which mental and physical events are correlated but not causally related. The moral is this: given a Humean/Lewisian account of causality, causal closure of the physical isn't plausible—unless, of course, there *aren't* any immaterial substances, in which case it is trivial. Indeed, given a Humean/Lewisian account of causality, the difference between dualistic interactionism and pre-established harmony can hardly be so much as stated. The same goes for another venerable

<sup>49</sup> *An Inquiry Concerning Human Understanding*, (LaSalle: Open Court Publishing Co., 1956), section VII, 83.

<sup>50</sup> "Causation", *Journal of Philosophy*, 70/17 (Oct. 1973), 563. Lewis later added some bells and whistles to fend off certain counterexamples.

contrast: that between dualistic interactionism and occasionalism. For if, as on occasionalism, my willing to raise my arm is the occasion for God's causing my arm to rise, then presumably God would not have caused it to rise if I had not willed to raise it; hence the counterfactual *If I had not willed to raise my arm, my arm would not have risen* is true, so that on the Lewis account my willing to raise my arm causes it to rise. In order to state these distinctions, and this objection to dualism, we must suppose that causation involves more than constant conjunction and more than counterfactual dependence: it must involve something further, something in the neighborhood of *production, making*, a sort of causal oomph or force, a necessary connection of some kind.<sup>51</sup> So to consider this objection, let's assume that causality is more than counterfactual dependence.

Now the objection that an immaterial substance can't have causal effects in the material world is usually stated as a rhetorical question: "How is this utterly insubstantial 'thinking substance' to have any influence on ponderous matter? How can two such different things be in any sort of causal contact?"<sup>52</sup> The answers, of course, are supposed to be "It can't" and "They can't". This objection is perhaps the most widely urged of all the objections against dualism; according to Churchland and Dennett it is widely thought conclusive.<sup>53</sup> But what is there to be said for it? From a Christian or theistic perspective, obviously, nothing at all. The claim is that no immaterial substance can cause effects in the hard, heavy, massy physical world. But this is a claim a theist can't take seriously: for of course *God* is an immaterial substance who causes effects in the hard, heavy, massy physical world. Therefore it can't be a true general principle that immaterial substances can't have causal effects in the physical world. This objection, even if the most widely accepted and respected of them all, should carry no weight with Christian theists.

I suppose someone might say that God is an immaterial substance that can have effects in the physical world, but he is the only immaterial substance that can do a thing like that; no finite immaterial substance can do such a thing. But why believe that? What is the or even a reason to think it true? True: we have little or no insight into how it is that an immaterial substance can cause changes in the physical world; but we have equally little insight into how it is that a *material* substance can cause changes in the physical world. Causation as a non-Humean relation among finite substances is something of a mystery; but it is no more mysterious where one of the relata is material and the other immaterial than where both are material or both immaterial.

<sup>51</sup> It is the difficulty of making clear this kind of connection that is part of the charm of occasionalism; in the case of divine causation, the connection isn't obscure at all; it's just broadly logical necessity. Every world in which God says 'Let there be light!' is a world in which there is light.

<sup>52</sup> Churchland, *Matter and Consciousness*, 11.

<sup>53</sup> See also Anthony Kenny, *Descartes* (New York: Random House, 1968), 222–3.

### *The Pairing Problem*

Objection 5 is usually formulated (if that is not too strong a word) by way of rhetorical questions; thus, for example, Churchland: “How is this utterly insubstantial ‘thinking substance’ to have any influence on ponderous matter? How can two such different things be in any sort of causal contact” (above, p. 129)? Jaegwon Kim provides a notable exception; he actually develops a serious and responsible statement of the alleged problem. (In his case ‘formulate’ is certainly not too strong a word.) Kim’s efforts here go so far beyond the usual that they deserve to be treated as a separate objection to dualism.

Kim begins by pointing out that the usual rhetorical-question formulations of the objection have nothing to be said for them. By way of a more serious effort, he asks us to suppose that “Smith and Jones are ‘psychophysically synchronized’: each time Smith’s mind wills to raise his hand, so does Jones’s, and vice versa, and every time they will to raise their hands, their hands rise.”<sup>54</sup> What is it that makes it the case that it is *Smith’s* willing, not Jones’s, that causes Smith’s hand to rise? After all, both willings are spatiotemporally related to the event of Smith’s hand rising in the same way: they occur at the same time, and neither is spatially related to that event. So in virtue of *what* is it that Smith’s willing, rather than Jones’s willing, causes Smith’s hand to rise? We can’t answer by pointing out that Smith wills that *Smith’s* hand rise, while Jones wills that *Jones’s* hand rise for, says Kim, what makes a given body B the body of a given person S is that S is able to cause changes in B *directly*. (I can raise my arm directly; I can raise yours only by taking hold of it with my hand and then raising it.) But then bodily ownership, for the dualist, is explained by way of psychophysical causation; therefore we can’t use bodily ownership to explain psychophysical causation.

Kim presumably won’t be satisfied with the answer, “Well, we don’t so far have a problem except in cases of people who are psychophysically synchronized, and people are very seldom psychophysically synchronized.” His idea would be that the dualistic interactionist (hereafter dualist) is committed to the *possibility* that there be cases of psychophysically synchronized people where nonetheless it is Smith’s willing, not Jones’s, that causes Smith’s arm to rise. And if there were such cases, there would have to be something, some further factor X, that determined, grounded, made it the case, that Smith’s willing, not Jones’s, causes Smith’s arm to rise. In the case of causation on the part of material beings, that further factor would involve spatio-temporal relations; but those aren’t available to the dualist.

Still, there may be an easy answer: there is an asymmetry about these willings. What Smith wills is that *this* (pointing to the hand) hand rise. Of course this hand is Smith’s; but willing that Smith’s hand rise is not the same thing as willing that *this* hand rise, even if this hand is Smith’s. So Smith but not Jones wills that

<sup>54</sup> “Lonely Souls: Causality and Substance Dualism”, in *Soul, Body, and Survival* (Ithaca, NY: Cornell University Press, 2001), 30–43 (33).

*this* hand rise, and Jones but not Smith wills that *that* hand rise. And the further factor X that makes it the case that Smith's willing causes Smith's hand to go up is that Smith wills that this hand go up; similarly, of course, for Jones and *that* hand. But Kim can easily amend his example so as to sidestep this reply: suppose that both Smith and Jones will that Smith's hand rise and do so at the same time: by virtue of what is it that *Smith's* willing, as opposed to Jones's, causes Smith's hand to rise? As Kim puts it,

There are two souls, A and B, and they perform a certain mental action, as a result of which a change occurs in material substance M. We may suppose that mental actions of the kind involved generally cause physical changes of the sort that happened in M, and, moreover, that in the present case it is soul A's action, not soul B's, that caused the change in M. Surely such a possibility must exist. But ask: What relation might perform the job of pairing soul A's action with the change in M, a relation that is absent in the case of soul B's action and the change in M? Evidently, no spatial relations can be involved to answer this question, for souls are not in space and are not able to bear spatial relations to material things ('Lonely Souls', 36).

Kim's thought, then, is that in any case where an event A causes an event B, there must be some factor, some X, in virtue of which it is A that causes B, in virtue of which A is paired with B. In the case of material events, this factor X, he suggests, will be a matter of spatio-temporal relations, although he doesn't say what, in general, these spatio-temporal relations would be. Spatio-temporal relations aren't available to do the job for the dualist, however, because temporal relations by themselves obviously can't do the job, and the soul isn't in space. But there don't seem to be any other candidates for the pairing relation; so there is a deep difficulty for the dualist here, one in virtue of which dualism should be rejected.

What can the dualist say for herself? First, is it really clear that in any case of causation, there must be this factor X that pairs up event A with event B, that makes it the case that A is the cause of B? I have two worries here. First, it isn't clear that spatio-temporal relations suffice for the pairing job; at any rate if we take quantum mechanics seriously. On some interpretations of quantum mechanics objects don't have a determinate position, or indeed any position at all, between collapses of the wave function; presumably the same goes, therefore, for events involving those objects. Of course there are other interpretations of quantum mechanics that lack this feature; so perhaps this isn't a serious worry for Kim.

Second and more important: why must we suppose that *there is* such a factor X? The question is: when event A causes event B, what is it that pairs A with B, rather than with C or D? What is it that makes it the case that A causes B? But maybe this is a confused question, or at any rate a question that conceals a contentious philosophical position. Consider the similar and oft-asked question about identity over time. What is it that makes it the case that object A at time t is identical with object B at some earlier time t\*? Similarity? Causal connections of

certain kinds? Many answers have been proposed, but none seems to work. And perhaps the right answer to the question is: there isn't anything (anything *else*, so to speak) that makes it the case that A is identical with B. Identity doesn't have to supervene on other properties. Of course there are necessary conditions of A's being identical with B. For example, both A and B must exist, and (perhaps) must have existed at each time between  $t$  and  $t^*$ ; and if A and B are physical objects, then (perhaps) there must be a continuously occupied spatio-temporal path between the location of A at  $t$  and that of B at  $t^*$ ; but there isn't anything that makes it the case that A at  $t$  is identical with B at  $t^*$ . Couldn't it be the same in the case of causation? Why does there have to be something, a state of affairs or something else, that makes it the case that event A causes event B? This is not an easy question. It is intimately connected with this question: which is prior: causal laws, or individual examples, cases, of causation? That is also a difficult question, and it may have different answers for divine causation on the one hand and creaturely causation on the other.

But we don't have to have answers to these difficult questions in order to see that the pairing problem, if there really is a pairing problem, is not a problem for dualists who are also theists. Medieval and Renaissance theists held, of course, that God creates the universe and sustains it and its parts in existence. But they also held that God *concurs* with every causal transaction that takes place; this concurrence is both necessary and sufficient for a given event (or substance) A to cause a given event B. Now one might suspect that this concurrence doctrine is metaphysical overkill—little more, really, than an attempt to pay God unnecessary (and unwanted) metaphysical compliments. If there really is a pairing problem, however, divine concurrence offers an easy solution: the relevant factor distinguishing Smith's willing from Jones's willing is that God concurs with the state of affairs *Smith's willing causing Smith's arm to rise*, but does not concur with *Jones's willing causing Smith's arm to raise*. That's the further factor X that makes it the case that it is Smith's willing that does the causing.

Perhaps Kim would want to reply as follows: divine concurrence is a solution to the pairing problem only if theism is viable, and theism is viable only if it is possible that God cause events in the world. Now Kim apparently thinks this pairing problem would hold for *any* alleged cases of causation on the part of an immaterial substance: "the difficulty we have seen with Loeb's interpretation of Descartes as a Humean in matters of causation, I believe, points to a more fundamental difficulty in the idea that mental substances, outside physical space, can enter into causal relations with objects in physical space" (p. 35). He might therefore suppose that the pairing problem affects alleged divine causation just as much as creaturely causation. For according to theism, God is not in space (and, some say, not in time either). Therefore the factor X that answers to the pairing problem in the case of material objects, that is, some relation to space and time (or space–time) isn't present in cases of divine causation. No doubt the theist is committed to the possibility that both God and someone else, an angel,

perhaps, will that something happen; what is it that makes it the case that it is *God's* willing that causes the event, rather than the angel's? Not spatio-temporal relations, clearly; but then what? What is that factor X in the case of alleged divine causation?

But here there appears to be an easy answer. According to classical theism, it's a necessary truth that whatever God wills, takes place. It's a necessary truth that if God says, "Let there be light," then there is light. Necessarily, if God says, "Let Adam come into existence," Adam comes into existence. So what is it that makes it the case that God's intentions cause what they cause? To ask that question is like asking, "What is it that makes an equiangular triangle equilateral?" The answer is (broadly) logical necessity; it's necessary that whatever God wills comes to be, just as it's necessary that every equiangular triangle be equilateral.<sup>55</sup> Accordingly there isn't a problem about that factor X in the divine case; but then divine concurrence solves the pairing problem, if there really is such a problem, for the case of immaterial created substances.<sup>56</sup> Here is another objection to dualism, or argument for materialism, that ought to have no purchase at all upon a Christian (or other) theist.

### *Localization and Dependence*

According to Nancey Murphy:

In particular, nearly all of the human capacities or faculties once attributed to the *soul* are now seen to be functions of the brain. Localization studies—that is, finding regional structure or distributed system in the brain responsible for such things as language, emotion and decision making—provide especially strong motivation for saying that it is the brain that is responsible for these capacities, not some immaterial entity associated with the body. In Owen Flanagan's terms, it is the brain that is the *res cogitans*—the thinking thing.<sup>57</sup>

Localization studies show that when a given sort of mental activity occurs, certain parts of the brain display increased blood flow and increased electrical activity. Paul Churchland adds that mental activity is also in a certain important way *dependent* on brain activity and brain condition:

<sup>55</sup> This is part of the attraction of occasionalism. It is hard to see what causality amounts to in the case of secondary or created causes, just as it is hard to see what necessity amounts to in the case of "natural" necessity, the sort of necessity that natural laws are supposed by, for example, D. M. Armstrong in *What is a Law of Nature?* (Cambridge: Cambridge University Press, 1983) to have. (Armstrong has since revised the views.) But it is easy to see what causality amounts to in the case of God's causing something: it's just a matter of logical necessity.

<sup>56</sup> Kim, obviously, is certainly among the most thoughtful materialists; and he finds both reductive and non-reductive materialism deeply problematic. This should incline him towards dualism; but of course he also thinks there is this pairing problem for dualism. If he thinks the pairing problem is the only serious problem for dualism, and if he agrees that divine concurrence offers an easy (theistic) solution to that problem, then wouldn't he have here a powerful theistic argument?

<sup>57</sup> Warren Brown, Nancey Murphy, and H. Newton Malony (eds.), *Whatever Happened to the Soul?* (Minneapolis: Fortress Press, 1998), 1.

Alcohol, narcotics, or senile degeneration of nerve tissue will impair, cripple, or even destroy one's capacity for rational thought. Psychiatry knows of hundreds of emotion-controlling chemicals (lithium, chlorpromazine, amphetamine, cocaine, and so on) that do their work when vectored into the brain. And the vulnerability of consciousness to the anesthetics, to caffeine, and to something as simple as a sharp blow to the head, shows its very close dependence on neural activity in the brain. All of this makes perfect sense if reason, emotion and consciousness are activities of the brain itself. But it makes very little sense if they are activities of something else. We may call this the argument from the *neural dependence* of all known mental phenomena.<sup>58</sup>

It isn't true at all that it makes very little sense to say that activities of the immaterial self or soul are dependent in this way on the proper function of the brain. Still, this argument from localization and neural dependence is perhaps the strongest of the arguments against dualism. That may not be much of a distinction, given the other arguments are substantially without any force, at least for someone committed to Christian theism. But this argument does seem to carry a certain minimal force; at any rate dependence and localization phenomena do suggest the possibility that the brain is all there is here. Taken as a serious argument, however, and looked at in the cold light of morning, it has little to be said for it. What we know is that for at least many mental functions or actions *M*, there are parts of the brain *P* such that (1) when *M* occurs, there is increased blood flow and electrical activity in *P*, and (2) when *P* is damaged or destroyed, *M* is inhibited or altogether absent. Consider, therefore, the mental activity of adding a column of figures, and let's assume that there is a particular area of the brain related to this activity in the way suggested by (1) and (2). Does this show or tend to show that this mental activity is really an activity of the brain, rather than of something distinct from the brain?

Not obviously. There are many activities that stand in that same relation to the brain. There is walking, or running, or speaking, or waving my arms and moving

<sup>58</sup> *Matter and Consciousness*, 20. See also Thomas Nagel's "Concealment and Exposure and Other Essays" (New York: Oxford University Press, 2002); in the course of a long, detailed, and subtle discussion, Nagel argues that there is a logically necessary connection between mental states and physical states of the following sort: for any mental state *M* there is a physical state *P* such that there is some underlying reality *R*, neither mental nor physical but capable of having both mental and physical states, which has essentially the property of being such that necessarily, it is in *P* just if it is in *M*. (And perhaps it would be sensible to go on from that claim to the conclusion that it is not possible that I exist when my body *B* does not.) Nagel concedes that it seems impossible that there be such a reality; his argument that nonetheless there really is or must be such a thing is, essentially, just an appeal to localization/dependency phenomena: "the evident massive and detailed dependence of what happens in the mind on what happens in the brain provides, in my view, strong evidence that the relation is not contingent but necessary" (p. 202), and "The causal facts are strong evidence that mental events have physical properties, if only we could make sense of the idea" (p. 204). The particular route of his argument here is via an argument to the best explanation: he suggests that the only really satisfactory explanation of those localization/dependency phenomena is the existence of such an underlying reality. (Of course if that is what it takes for a really satisfying explanation, we may wonder whether there *is* a really satisfying explanation here; are we guaranteed that all phenomena have what we take to be really satisfying explanations?)

my fingers: for each of these activities there is a part of my brain related to it in such a way that when I engage in that activity, there is increased blood flow in that part, and when that part is damaged or destroyed, paralysis results so that I can no longer engage in the activity. Who would conclude that these activities are really activities of the brain rather than of legs and trunk, or mouth and vocal cords, or arms? Who would conclude that my fingers' moving is really an activity of my brain and not of my fingers? Ric's rock climbing is dependent on appropriate brain activity; it hardly follows that rock climbing just *is* an activity of his brain. Digestion will occur only if my brain is in the right condition; how does it follow that digestion is really an activity of the brain, and not an activity of the digestive system? My brain's functioning properly depends on blood flow and on the proper performance of my lungs; shall we conclude that brain function is really circulatory or pulmonary activity? All of my activities depend upon my ingesting enough and the right kind of food; shall we see here vindication of the old saw 'you are what you eat'?

The point, obviously, is that dependence is one thing, identity quite another. Appropriate brain activity is a necessary condition for mental activity; it simply doesn't follow that the latter just is the former; nor, as far as I can see, is it even rendered probable. We know of all sorts of cases of activities A that depend upon activities B but are not identical with them. Why should we think differently in this case?<sup>59</sup>

Well, perhaps someone will say that in the cases I've been citing, we know on independent grounds that there are two kinds of activities; we know that digestion is an activity of stomach, intestines and the like, and not just of the brain, even if brain activity is a necessary condition of digestion. But (so the objector continues) just that knowledge is what is lacking in the case of mental activity; we don't know of something distinct from the brain that is involved in mental activity. Suppose that were so: we would still have at best a massively weak argument for materialism, for (obviously) the fact that we don't know of such a thing hardly shows that there isn't any such thing. Should we pay much attention to an atheologian who argued that since we don't know of an all-powerful, wholly good, all-knowing being who has created the world and sustains it in existence, there isn't any such being? But in any event it isn't true that we don't know of something distinct from the brain that is involved in mental activity. The above arguments for dualism, I claim, gives us, at the very least, good reason to hold that thinking is not, or not merely, an activity of the brain. But then it is not the case that thinking is an activity of the brain and nothing else.

In conclusion, then: there are powerful arguments against materialism. When we consider the bearing of Christian belief upon materialism, we find still more

<sup>59</sup> A related argument for materialism has it that the great theoretical benefits of *identifying*, for example, pain with C-fiber firing, warrant accepting materialism. For discussion of this claim, see the last section of "Against Materialism".

reasons to reject it in favor of dualism. Were it not for my respect for my materialist colleagues, I would certainly say “Never has so implausible a doctrine been so widely accepted!”<sup>60</sup>

## APPENDIX: INDICATION AND CONTENT

I argued above that a material structure or event isn't the right sort of thing to have belief content; this problem has not been lost on materialists, canny lot that they are. In trying to deal with it, they typically ignore Leibniz's problem and instead offer suggestions as to how it might be that a neural object or event could have (original) content after all. Most attempts to do so begin with *indicators*, or *indication*, or *indicator meaning* as outlined above.

The first step is to call these structures, the ones correlated with external or internal conditions of one kind or another, 'representations'. Indeed, the idea that such structures are representations has become so common that it is part of the current background assumptions in cognitive neuroscience.<sup>61</sup> Those patterns of neural firing in the frog's brain are said to be representations of flies, or bugs, or small flying objects, or small black objects (there is usually considerable latitude of choice as to what gets represented); those magnetosomes in anaerobic bacteria are said to represent north, or the direction towards oxygen-free water, or the lines of the earth's magnetic field; the structures in your body that respond to the temperature of your blood are said to represent that temperature.

Now the terms 'represent', 'representation', and 'representative' are multiply ambiguous. Webster's *Third International* gives a whole host of analogically connected meanings: you can send your representative to a meeting; your state or national representative represents your interests (we hope); an artist can produce a representation of a battle; a musical passage can represent a storm; x's and o's can represent football players, and a dotted line can represent where the tight end is supposed to go, a scale model of Mt Rainier can represent Mt Rainier. This term is therefore something of a weasel word, a property that often gets exploited in philosophy of mind or cognitive science contexts. Since the term is ordinarily used without explicit definition, it is often hard to know just what is meant by calling those indicators 'representations'; shall we say that wherever you have causal or nomological correlation, you have representation? Shall we say that smoke represents fire (and fire represents smoke), that the rate at which the wheels of my car turn represent the speedometer reading, and that trees budding represent spring or warmer weather (and vice versa)? I guess we can say these things if we like; it's a free country, and the term 'representation' is flexible enough to allow it.

But here is the crucial next step: efforts to understand belief materialistically typically try, somehow, to promote these representations to *beliefs*. In so doing, they ordinarily simply ignore Leibniz's problem—the fact that it looks as if a material thing can't think, or be, a belief. But this procedure is also unpromising in its own right: representation of this sort is nowhere near sufficient for belief. The gas gauge on my car may represent

<sup>60</sup> Well, *almost* never. Verificationism, which was as widely accepted in the 1940s and 1950s as materialism is now, is at least equally implausible.

<sup>61</sup> See Ramsey, *Using and Abusing Representation: Reassessing the Cognitive Revolution*.

the amount of gasoline in the tank, and the weight on the bolts holding the tank to the frame, and the volume of air in the tank, and other things as well; nothing in the relevant neighborhood has beliefs on these scores. Those magnetosomes perhaps represent the direction to oxygen-free water; neither they nor the bacteria that contain them believe that's the way to oxygen-free water. Those internal structures that indicate and thus represent your blood pressure do not believe that your blood pressure is thus and so, and neither (most of the time) do you. The thing to see is that no amount of this indication and representation, no matter how gussied up, is sufficient for *belief*. Clearly a material object *can* be a representation in *some* sense: Michelangelo's *David* for example, is a representation of David, and a few weird lines in a cartoon can represent George Bush. But it doesn't follow that a material structure can be a belief, or that it can have propositional content.

There are many ways in which materialist thinkers try to promote indication or representation to belief. I can't of course comment on them all; I'll content myself with brief comments on a couple of the most prominent.

First, there is Jerry Fodor's suggestion. It is plausible to suppose that there is a certain neuronal structure or event that is involved in the perception of cows, and that is caused by cows, and that indicates cows. These structures, says Fodor, have the content *cow*. But note further that these structures can also, under certain conditions, be caused by other things—a moose in the twilight, or under certain conditions maybe a large cat, or a perhaps a scale model of a moose. What confers content on such a structure—the content *cow*—is that there being structures of that sort that are *not* caused by cows, is asymmetrically dependent upon there being structures of that sort that *are* caused by cows: “But ‘cow’ means *cow* and not *cat* or *cow* or *cat* because *there being cat-caused ‘cow’ tokens depends on there being cow-caused ‘cow’ token, but not the other way around.*”<sup>62</sup> This also seems monumentally unpromising, at least if taken as presenting a necessary and sufficient condition.<sup>63</sup> Perhaps we can rewrite Fodor's suggestion more explicitly as follows:

- (F) A token cow indicator C\* of type T has the content *cow* just if there being non-cow-caused tokens of T depends on there being cow-caused tokens of T, but not conversely.

Thus there is, we may suppose, a certain neural structure ordinarily caused, in human beings, by the perception of a cow; in certain circumstances, however (twilight, for example, or great distance) that token will be caused, not by a cow, but by a moose or horse, or cat, or cow picture, or too much whiskey, or whatever. Tokens of this type T have the content *cow*, however, because if there weren't any cow-caused tokens of T, there wouldn't be any non-cow-caused tokens of T; but there could perfectly well be cow-caused tokens of T even if there weren't any non-cow-caused tokens of T.

Taken as a specification of a necessary and sufficient condition, (F) has two problems: the proposed condition is not necessary, and it is not sufficient. First, there are objects about which we have beliefs, of which we have concepts, and denoted by our terms, such that there probably aren't any indicators of them at all. These would comprise the

<sup>62</sup> *A Theory of Content and Other Essays* (Cambridge, Mass.: MIT Press, 1990), 91.

<sup>63</sup> And if it isn't intended as a sufficient condition, it won't really be relevant to our current concerns, i.e. it won't suffice to show how representations can be promoted to beliefs.

whole realm of abstracta: properties, propositions, numbers, sets, and the like. These things do not enter into relevant causal relations with us; hence there aren't indicators of them in our brains—or, if there are, they aren't caused by these abstract objects. Hence the condition (F) proposes isn't necessary: we have beliefs about, concepts of, and terms denoting objects that don't cause the relevant indicators. But secondly, (F) is also insufficient: it is much too generous with content. Consider cow pies, for example; they apparently fill the bill specified by (F). Cow pies are cow indicators; furthermore, there wouldn't be non-cow-caused cow pies if there weren't cow-caused cow pies, although there could certainly be cow-caused cow pies even if there weren't non-cow-caused cow pies. But then cow pies, according to (F), have the content COW—which, not to put too fine a point on it, is no more than a load of organic lawn food.

Second, there is Fred Dretske's work, perhaps the most sophisticated and accomplished attempt to explain belief from a materialist perspective.<sup>64</sup> Like the other two, Dretske ignores Leibniz's problem; like them he simply assumes that it is possible for a material thing to think and for a material assemblage of neurons to be a belief. And like nearly everyone else, Dretske begins with the notion of indication, correlation (perhaps nomic, perhaps causal) between events of one kind and events of another. His attempt to explain belief in terms of indication involves two additional ideas. First, the notion of *function*. All beliefs are representations, and representations essentially involve functions: "The fundamental idea [of representation] is that a system, S, represents a property F, if and only if S has the function of indicating (providing information about) the F of a certain domain of objects."<sup>65</sup> So not all cases of indication are cases of representation: the fuel gauge in my automobile indicates the amount of gasoline in the tank, the weight on the bolts holding the tank to the frame, the amount of air in the tank, the air pressure, the altitude, the temperature, the potential across a certain circuit, and many other things; its *function*, however, is to indicate the amount of gasoline in the tank. Hence it represents the amount of fuel in the tank and does not represent those other properties and quantities, fascinating as they may be. This appeal to function enables Dretske to see representational contexts as like belief contexts in being intentional: it may be that it is the function of something or other to indicate a property *p*, while it isn't its function to indicate a nomically or logically equivalent property *q*.

But just as not every case of indication involves representation, so, according to Dretske, not every case of representation is a case of belief (or proto-belief, as he tends to put it). He cites the case of the noctuid moth, which, upon detecting the bursts of high frequency sound emitted by the bat's sonar, executes evasive maneuvers. Here we have representation; it is the function of those neural structures registering that sound to indicate the presence of bats, to carry the information that bats are present. But these structures, says Dretske, are not beliefs and do not have belief content. Where *C* is a structure representing something or other (and now we come to the second additional idea), belief content is present *only if C causes some motor output or movement M, and the explanation of C's causing M is C's carrying the information that it does*. That is not so in the case of those structures in the noctuid moth:

<sup>64</sup> See esp. *Explaining Behavior and Naturalizing the Mind* (Cambridge, Mass.: MIT Press, 1995).

<sup>65</sup> *Naturalizing the Mind*, 2.

the explanation of why *this* C is causing *this* M, why the moth is now executing evasive maneuvers—has nothing to do with what *this* C indicates about this moth's surroundings. The explanation lies in the moth's genes. (*Explaining Behavior*, 92)

Take a given moth and the neural circuit C whose firing causes those maneuvers M: the explanation of C's causing M is not that C indicates the presence of bats, but the way the neural circuitry of this moth is deployed. The fact that in these moths C represents the presence of bats may explain or help explain why moths of this type have survived and flourished; but the fact that in a given moth C represents bats does not explain why C causes M.

If we don't get belief here, where do we get it? Where there is *learning*, says Dretske (here learning, on pain of circularity, does not entail or presuppose belief). Consider a bird that learns to peck at a red spot because it is rewarded when it does. At first the bird pecks aimlessly, now at the red spot, now at the black spot, now at a shadow on the walls of its cage. But then we reward it when it pecks at the red spot. Soon it will peck only or mainly at the red spot; it has learned something. What has happened here? Well, the bird had a red spot detector to start with; by virtue of learning, that structure came to cause the bird to peck at the red spot. And the structure in question causes the motor output in question *because* that structure indicates a red spot, carries the information that the figure in front of the bird is a red spot. Here, therefore, we do have a case of belief content, says Dretske, and the bird can be said to believe (or proto-believe) that there is a red spot in front of it.

As far as I can see, Dretske's complete account of belief can be put as follows:

- (D) x is a belief if and only if (1) x is a state of an indicating element E in a representational system (e.g. the event consisting in the system's being 'on'), (2) x's function is to indicate something F, (3) x is in the mode or state it is in when it indicates something F, (4) x causes some movement M, and (5) the explanation of x's causing M is that it indicates F.

A comment on (3): it's not necessary that, on the occasion in question, x is actually indicating something F; perhaps on this occasion x is misrepresenting. We fix red-colored spectacles on the bird: now its red spot indicator causes it to peck at any spot, red or not. But the red spot indicator is still on, as we might say, even when in fact the spot in front of the bird is black.

This is a complex and sophisticated account. Still, sophisticated as it is, Dretske's account, I think, won't anywhere nearly do the job. First, a couple of semi-technical objections. I believe that  $7 + 5 = 12$ ; nothing, however, carries the information that  $7 + 5 = 12$ , and indeed  $7 + 5$ 's *being equal to 12* isn't information. That is because, according to Dretske's (Shannon) conception of information, information is always a matter of reduction of possibilities; but  $7 + 5$ 's *equaling 12* doesn't reduce the possibilities with respect to anything. The account is therefore too strong; it rules out beliefs that are logically necessary in either the broad or the narrow sense. And just what kind of possibilities are we thinking of here? If causal or nomic possibility is relevant (if carrying information requires the reduction of causal or nomic possibilities), then the account also fails to work for nomologically necessary beliefs, such as that (as current physics has it, anyway) nothing travels faster than light (more exactly, nothing accelerates from a velocity less than that of light to a velocity greater than that of light). This doesn't

reduce the nomic possibilities. And what about beliefs about the past? Given that past propositions are 'accidentally necessary', does anything *now* carry the information that Brutus stabbed Caesar?

Further, I believe that Proust is more subtle than L'Amour; is it even remotely plausible to suppose that I must therefore have a Proust-is-more-subtle-than-L'Amour-indicator, a neural structure correlated with *Proust's being more subtle than L'Amour* whose function it is to indicate that Proust is more subtle than L'Amour? Or a structure that fires when in the (virtual?) presence of a pair of writers, one of whom is more subtle than the other? And even if there were such structures, would they have to cause *motion* of one sort or another, for me to believe that Proust is more subtle than L'Amour? Maybe I've always believed this, but never said so, or in any other way displayed this belief in my behavior.

Still further, return to that noctuid moth. Perhaps it was designed by God; and perhaps God designed it in such a way that C, the structure causing that evasive motion, causes that motion because C indicates the presence of bats. Then it would be true that C causes M because of what it indicates (God chooses C to cause M, because C indicates the presence of bats) and, on Dretske's account, the moth would on the appropriate occasions believe that there are bats present. So if the moth came to be by undirected evolution it doesn't have beliefs (or at least doesn't have the belief that bats are present when its bat indicator is activated); if God has designed it, however, then it does have that belief on those occasions. Can that be right? In the same way there are all those internal indicators I mentioned a bit ago: structures whose function it is to indicate blood pressure, temperature, sodium level, sugar level, and the like. These indicators are in fact so constituted that they cause certain kinds of movements. If human beings have been designed by God, then presumably they cause those movements because of what they indicate; that's why God designed the system in such a way that they *do* cause those movements. So on Dretske's account, these structures, or we who contain them, would hold the associated beliefs about our blood temperature, pressure, sodium level, sugar level, and the like. But we don't; if Dretske's account were right, therefore, this would constitute an argument against the existence of God. Clearly it doesn't.

Insofar as they can't accommodate necessary beliefs and beliefs about the past, Dretske's conditions are too strong: they aren't necessary for belief. But they are also too weak: they aren't sufficient either. If his account were correct, then if we have been designed by God, we hold all those beliefs about blood pressure, temperature, sodium content, and the like; but we don't. You may or may not think we have in fact been designed by God; but even if we haven't it is certainly *possible* that we have; hence it's possible that Dretske's conditions hold when no beliefs are present. And really, why should the fulfillment of Dretske's conditions have anything at all to do with belief? So there is this structure that has the function of indicating something and causes what it does because of what it indicates; does that really so much as slyly suggest that something in the neighborhood of this structure holds the appropriate belief, or any belief at all? Consider the thermostat. The bimetallic strip indicates the temperature, and has the function of indicating it. Further, when it bends enough to close the circuit, thereby causing furnace ignition, it causes what it causes because of what it indicates. We designed the thermostat in such a way that when that strip indicates 67°F, it causes the furnace to ignite; so the explanation of its causing that movement is that it is indicating that the temperature is

67 °F. But neither the bimetallic strip nor the thermostat, nor the furnace nor anything else need believe that the temperature is 67 °F. Even if we set aside Leibniz's problem, we must conclude, I think, that Dretske's account, subtle and powerful though it is, won't anywhere nearly serve as an explanation of how there could be beliefs if materialism about human beings is true.