

You are Simple

David Barnett

University of Colorado at Boulder

I argue that, unlike your brain, you are not composed of other things: you are *simple*. My argument centers on what I take to be an uncontroversial datum: for any pair of conscious beings, it is impossible for the pair *itself* to be conscious. Consider, for instance, the pair comprising you and me. You might pinch your arm and feel a pain. I might simultaneously pinch my arm and feel a qualitatively identical pain. But the pair we form would not feel a thing.¹ Pairs of people *themselves* are incapable of experience. Call this *The Datum*. What explains The Datum? I think the following exhaust the reasonable options. (1) Pairs of people lack a sufficient *number* of immediate parts. (2) Pairs of people lack immediate parts capable of standing in the right sorts of *relations* to each other and their environment. (3) Pairs of people lack immediate parts of the right *nature*. (4) Pairs of people are not *structures* (they are unstructured *collections* of their two immediate parts). (5) Some combination of (1) – (4). Finally, (6) pairs of people are not *simple*.

I defend (6). I argue that none of (1) – (4) is individually sufficient to explain The Datum. Then I argue that no combination is sufficient. I conclude that (6) best explains The Datum.

Although my paper has the structure of an argument, its real aim is merely to make credible a position that is rarely taken seriously by contemporary philosophers, namely, that conscious beings must be simple.² The argument itself has escape routes. Indeed, I do not expect the confident materialist to find it persuasive. Still, it has value. At the very least, it serves as a stalking horse for the materialist. The argument involves a long, slow, development of an intuition pump that is designed to give the reader a

¹ As Chisholm says, “You could want the weather to be colder and I could want it to be warmer; but that heap or aggregate which is the pair of us (that thing that weighs 300 pounds if you and I each weigh 150 pounds) does not want anything at all” (1991, p. 172).

² Some recent discussions of the position include Bennett 1967, Hart 1988, Chisholm 1991, Foster 1991, Zimmerman 1991, 2005, Taliaferro 1994, Hasker 1999, Lowe 2001, and Barnett *forthcoming*.

certain perspective on key examples. By the end of the argument, the materialist just might find himself having the intuition that he is wrong.

In the initial sections, as I argue against (1) – (4), I tease out some intuitions related to the intuition behind The Datum, and I introduce some techniques for teasing out further intuitions in more sophisticated settings. Because the main purpose of these early sections is to develop these techniques, the initial examples that I deal with are deliberately simplistic in nature. Only in later sections will I introduce examples that exhibit the requisite degree of complexity for defeating plausible combinations of (1) – (4). Thus, I ask that you withhold your judgment of my ultimate conclusion until you see the whole picture.

1. Number

Do pairs of people lack a sufficient *number* of immediate parts to be conscious?

Perhaps. But this cannot be the full explanation of The Datum. For increasing the number of people in a collection does not have any significant effect on the absurdity of the idea that the collection *itself* might be conscious. For illustration, consider the triplet comprising you, Paul McCartney, and me. Whatever you, McCartney, and I might experience individually, we can be certain that there is no further experience enjoyed by our triplet. Or consider the entire world population. Might this huge collection of people itself be conscious? Might it currently be experiencing, say, the taste of McDonald's French fries? Of course not. For emphasis, try to imagine that each and every human is in excruciating pain, while our collection is itself experiencing pure bliss. This is absurd. No matter how large, a collection of people cannot itself be a subject of experience.

I am not denying that a collection of people might qualify as having an experience in a secondary sense, in virtue of one or more of its members having that experience. Often a whole is said to have a feature in virtue of one of its parts having the feature. For instance, a house is said to be on fire in virtue of its roof being on fire. And a shirt is said to be stained in virtue of its collar being stained. But in these cases it is the part that is the primary bearer of the feature. Perhaps humanity can itself be said to suffer in virtue of one or more of its members suffering. And perhaps a recently divorced couple can itself be said

to experience separation anxiety by virtue of its members experiencing this anxiety. Personally, I doubt that there is any legitimate sense in which a whole can “inherit” an experience in this way from its parts. But in any case this is not what I am denying. What I am denying is that a collection of people might itself be the *primary* bearer of an experience. Hereafter, when I ask whether something might be a bearer of experience, that is, a conscious being, I mean to ask whether it might be a primary bearer of experience.

I conclude that the idea that pairs of people lack a sufficient number of immediate parts to be conscious cannot by itself explain The Datum.

2. Relation

Do pairs of people lack immediate parts capable of standing in the right sorts of *relations* to each other and their environment to be conscious?

Perhaps. But this cannot be the full explanation of The Datum either. For the only remotely plausible candidate relations are causal-dispositional relations of the general sort borne by the parts of an ordinary human brain (or some other animal brain, or some entire organism) to one another and their environment; these are the relations that things must stand in if they are to jointly function, on a relevant level, like an ordinary human brain. But there is no metaphysical obstacle to two people standing in any such relation.

For illustration, consider the following scenario. Allowing for some radical changes to the laws of nature, imagine that some clever scientists shrink you and me down to the size of McCartney’s left and right brain hemispheres, respectively. The scientists then train us to function, at a relevant level, just as our respective hemispheres function: in terms of exchanging signals with the peripheral neurons and each other, we are trained to behave just as our respective hemispheres behave. McCartney’s left and right hemispheres are then removed and replaced with you and me, respectively. Someone pinches McCartney’s right arm (or his former right arm, should McCartney not survive the ordeal). When the signal arrives at the top of the spinal cord, I identify it; I notify you; we stimulate certain outbound

neurons; and we move into a new functional state. As a result, McCartney's head turns and faces his right arm; an irritated look appears on his face; and out of his mouth comes the words, "Stop that!" On a relevant functional level, you do just what McCartney's left hemisphere would have done. And I do just what McCartney's right hemisphere would have done. At a relevant level, the causal-dispositional relations borne by you and me are those that McCartney's two brain hemispheres would have borne. Together, you and I function like an ordinary human brain.³

Given our new relations to each other and our environment, is it any less absurd to think that the pair we form might *itself* be conscious? No. To be sure, there is nothing absurd in the idea that *McCartney* might somehow survive the procedure; though unlikely, perhaps he would remain conscious throughout the ordeal. What seems absurd, rather, is that *the pair comprising you and me* might be conscious. Variation in how two people are related to each other and their environment has no significant effect on the absurdity of the idea that the pair itself might be conscious.

I conclude that the idea that pairs of people lack immediate parts capable of standing in the right sorts of relations to each other and their environment to be conscious cannot by itself explain The Datum.

3. Nature

Do pairs of people lack immediate parts of the right *nature* to be conscious?

Perhaps. But this cannot be the full explanation of The Datum either. For varying the nature of the members of the pair does not have any significant effect on the absurdity of the idea that the pair itself might be conscious. It is absurd that a pair of people might itself be conscious. But it is equally absurd that a pair of carrots, rocks, dogs, electrons, or neurons might itself be conscious.

³ This is not quite right; for even if you and I function on a relevant level in the same way that McCartney's two brain hemispheres function, our disposition to function in this way is not law-like in the way that the disposition of McCartney's two brain hemispheres is. To correct for this difference, we can adjust our scenario as follows: it turns out to be a law of nature that when two humans go through the training and shrinking procedure described above, they become disposed—to the same law-like degree—to function, on a relevant level, in the same way that the corresponding brain hemispheres function.

If I tell you only that I have two objects in mind, *a* and *b*, you need more information before you can determine whether *a* is conscious and whether *b* is conscious, but you do not need more information to determine whether the pair comprising *a* and *b* is *itself* conscious. You know by mere reflection that it is not. Pairs of things, we know by mere reflection, cannot themselves be conscious. So there is no need to empirically investigate the individual natures of carrots, rocks, electrons, or neurons to determine whether pairs of them might themselves be experiencing the taste of McDonald's French fries; we know by pure reflection that they are not. In general, the absurdity that any given pair of objects might itself be conscious has nothing to do with the natures of the members of the pair.

I conclude that the idea that pairs of people lack immediate parts of the requisite nature to be conscious cannot by itself explain The Datum.

4. Structure

How can we know just by reflection that the pair comprising *a* and *b* is not itself conscious? Here is one suggestion: we can know by reflection that *pairs* of things are mere *collections*, and we can know by reflection that *conscious beings* are *structures*. Whereas a collection of things exists whenever those things exist, a structure of things exists only if those things stand in a certain relation required to exhibit the structure. This is not to say that a structure of things is essentially a structure of the particular things that currently exhibit it; some structures can survive the destruction or replacement of some of the things that exhibit them. For an example of a collection, consider the atoms that constitute this clay bowl; for an example of a structure, consider the bowl itself. Intuitively, if we were to spread the atoms evenly throughout the universe, their collection would survive, but the bowl they now form would not. On this informal characterization of a structure, the following entities seem initially to qualify as structures, rather than mere collections: systems, such as a galaxy of stars; certain artifacts, such as a table or a clay bowl; certain non-living natural objects, such as a rock or a planet; and certain living objects, such as an organism, one of its organs, or one of its cells.

Might pairs of people be incapable of experience because they are not structures? Perhaps. But this cannot be the full explanation of The Datum either. To see that it cannot, simply shift your attention from the pair of people inside McCartney's skull to the brain-like system they compose. Unlike the pair, the system is a structure: it would cease to exist if you and I were removed from the skull and placed in isolation chambers (the pair we form would not). Yet this system of people is no better a candidate for being a subject of experience than the pair that constitutes it.

I conclude that the idea that pairs of people are incapable of experience because they are not structures cannot by itself explain The Datum.

5. Combination of Number, Relation, Nature, and Structure

Perhaps The Datum is explained, not by any one of *Number*, *Relation*, *Nature*, or *Structure* (the four hypotheses considered above) alone, but by some combination of the four. For instance, perhaps a pair of people cannot itself be conscious because it is a *collection* resulting from the *mere existence* of *two* particular *people*, whereas a conscious being is a *structure* resulting from *many organs*, or *billions of cells*, or *quadrillions of particles*, standing to one another and their environment in certain causal-dispositional *relations*.

Before I evaluate this proposal, I want to discuss an appealing line of reasoning behind it: (i) human bodies—physical structures comprising organs, tissues, cells, molecules, and atoms—are conscious; (ii) the salient differences between human bodies and pairs of people are captured by *Number*, *Relation*, *Nature*, and *Structure*; (iii) none of these four hypotheses alone explains The Datum; hence, (iv) some combination of the four must explain The Datum. In particular, I want to discuss the initial—that is, pre-theoretical—appeal of (i). If asked outside a philosophical context whether human bodies are conscious, most of us are inclined to give a positive answer without hesitation. And yet, if my argument is sound, then it is impossible for a human body to be conscious, for it is not simple. Why, then, are we initially willing to ascribe consciousness to human bodies but not to pairs of people? I have two hypotheses.

First, an ascription of consciousness can be interpreted in one of two ways. On the strong interpretation, an utterance of ‘*x* is conscious’ means that *x* is *identical to* a conscious being; on the weak interpretation, it means that *x* is some conscious being’s body—that is, that *x* *embodies* a conscious being. Pre-theoretically, we interpret ascriptions of consciousness to human bodies in the weak sense. Our initial willingness to accept that human bodies are conscious is simply a willingness to accept that human bodies embody conscious beings. Otherwise we would have difficulty entertaining scenarios involving disembodiment and reincarnation. But we have no such difficulty. For illustration, imagine waking up, looking into a mirror, and discovering that you have swapped bodies with McCartney. No problem. Now imagine waking up and discovering that your favorite chair has swapped bodies with McCartney’s favorite chair. Big problem: the scenario does not make sense. We cannot make sense of chairs “swapping bodies” because our initial conception of them demands that they be *identical* to such bodies. By contrast, we *can* make sense of people swapping bodies, for our initial conception of them does not demand that they be identical to their bodies.

For emphasis, suppose that chairs *are* conscious, as they are depicted in Disney movies. Then we have no trouble making sense of their swapping bodies. But if we have no trouble imagining their swapping bodies on this supposition, then we must be interpreting the supposition in the *weak* sense, as the supposition that chairs *embody* conscious beings (what we imagine, then, is that the embodied beings swap bodies). For, interpreted in the strong sense, the supposition is obviously incompatible with body swapping.

By contrast, we initially interpret the proposal that a pair of people might itself be conscious in the strong sense—as I intend it. For, because we do not consider a pair of people itself to be a body, we never even entertain the idea that it might embody a conscious being.

So one hypothesis as to why we are initially willing to accept that a human body is conscious, but not that a pair of people is conscious, is that we initially interpret the proposal that a human body is conscious in the weak sense, whereas we initially interpret the proposal that a pair of people is conscious

in the strong sense. And it seems possible for a human body to embody a conscious being, yet impossible for a pair of people to be identical to a conscious being.

A second hypothesis centers on a difference between the way that a human body is typically presented to our minds and the way that a pair of people is typically presented to our minds. As the human body is typically presented, we are able to ignore its composite aspect. On a daily basis, we see human bodies as single, solid, human-shaped objects. The fact that these objects have left halves, right halves, fingers, hands, arms, and legs is obvious. However, because these parts appear to be spatially continuous with one another, we do not typically see human bodies for what they truly are: structures of organs, tissues, and cells—more fundamentally, structures of quadrillions of tiny particles separated by relatively vast amounts of space. As a result, in certain ways we are able to conceive of our bodies as simples.

Hume makes a related point:

An object, whose different co-existent parts are bound together by a close relation, operates upon the imagination after much the same manner as one perfectly simple and indivisible, and requires not a much greater stretch of thought in order to its conception. From this similarity of operation we attribute a simplicity to it, and feign a principle of union as the support of this simplicity, and the center of all the different parts and qualities of the object. (*Treatise* I.iv.6)

By treating our bodies in many respects as simples, we can take seriously the idea that our bodies are *identical* to subjects of experience.

However, when our bodies are presented to us in a way that makes it difficult to ignore their composite aspect, we resist ascribing consciousness to them. Leibniz makes the point as follows:

If we imagine that there is a machine whose structure makes it think, sense, and have perceptions, we could conceive it enlarged, keeping the same proportions, so that we could enter into it, as one enters into a mill. Assuming that, when inspecting its interior, we will only find parts that push one another, and we will never find anything to explain a perception. And so, we should seek perception in the simple substance and not in the composite or in the machine. (*Monadology*, paragraph 17)

Of course, it is nearly impossible to ignore the composite aspect of a pair of people. And so a second hypothesis as to why we are initially willing to accept that a human body is conscious, but not that a pair of people is conscious, is that human bodies are typically presented to our minds in a way that allows us to ignore their composite aspect, whereas pairs of people are not.

Given the initial plausibility of these two hypotheses, it is important, as we investigate the proposal that some combination of *Number*, *Relation*, *Nature*, and *Structure* explains The Datum, first that we focus only on the strict interpretation of the question of whether a given thing is conscious, and second that we not ignore the composite aspect of any composite object that is under consideration. Now, one way to show that no combination of *Number*, *Relation*, *Nature*, and *Structure* can explain The Datum is to consider the human body, not as we ordinarily do, as a solid, human-shaped, object, but rather as a structure of many organs, or of billions of cells, or of quadrillions of particles. We need to make salient the composite aspect of the body. The more salient we make this aspect, the less comfortable we will be ascribing consciousness to the body itself, until, at the limit, the whole idea will seem absurd. My strategy is to close the gap between a pair of people and a human body in stages. I will eliminate the difference first in the *number* of parts, then in the *relation* of parts, then in the sort of whole—*structure* versus collection—that comprises the parts, and finally in the *nature* of parts. My motivation for proceeding in stages is to help keep salient the composite aspect of the relevant candidates for consciousness. To ensure that the gap is completely closed, in the final stages we will consider the human body itself, not as we

typically do, but rather in ways that make it impossible for us to ignore its composite aspect. By the end of the exercise, we will see that no combination of *Number*, *Relation*, *Nature*, and *Structure* can explain The Datum; and we will see that the human body is no better a candidate for being a subject of experience than a pair of people.

First we eliminate the difference in the *number* of parts. Instead of considering a pair of people, we consider a collection of several billion people. We have already seen that a mere increase in the number of people has no effect on the absurdity of the idea that their collection might itself be conscious.

Second we eliminate the difference in the *relation* of parts. Here we can borrow any of a trio of examples from Ned Block 1978.

In Block's "Nation of China" example, we are to imagine that the head of an otherwise normal human contains only miniature two-way radios hooked up to inbound sensory neurons and outbound motor neurons. The radios send and receive signals to and from citizens of China, who are themselves equipped with two-way radios. Block chooses citizens of China because their number is on the order of the number of neurons in a typical human brain. A satellite system displays symbols that can be seen from anywhere in China. Each citizen is given a simple set of instructions: if a given symbol is displayed, then if certain radio signals are received from the sensory neurons, send a given signal to the motor neurons. Together, the billion or so citizens function, on a relevant level, just as a normal human brain functions.⁴ Yet the idea that this collection of people might itself be conscious is absurd.

In Block's "Miniature Men in the Head" example, we are to imagine that the head of an otherwise normal human is filled with a group of little men. Block never says how many men; let us assume that it is several billion. Also inside the head is a bank of lights connected to inbound sensory neurons, a bank of buttons connected to outbound motor neurons, and a bulletin board on which a symbol (designating the

⁴ As with our earlier scenario concerning a pair of people in McCartney's cranium, Block's scenario needs to be adjusted to address the fact that the disposition of the population of China to function like a human brain is not law-like to the same degree as the disposition of a typical brain. A further stipulation will suffice: it is a law of nature that, when a group of people aggregate in the way envisaged in Block's example, their collection becomes disposed—to the same law-like degree—to function, on a relevant level, in the same way that a typical brain is disposed to function. This sort of adjustment is required of all three of Block's examples.

current state of the system) is posted. Each man is given a simple set of instructions: if a given symbol is posted, then if certain lights are illuminated, press a given button. Together, the billions of men function, on a relevant level, just as a normal human brain functions. Yet the idea that this collection of tiny men might itself be conscious is absurd.

In Block's "Elementary-Particle People" example, we are to imagine that the sub-atomic particles in our bodies are gradually replaced with functionally equivalent spaceships piloted by tiny aliens. Our brains (or the objects that replace them, should they not survive the procedure) would continue to function, down to the sub-atomic level, just as they ordinarily would. It seems likely that we would continue to have experience just as we ordinarily would. The question arises: might we, or any other subject of experience, be *identical* to the envisaged collection of alien-piloted spaceships? Might this collection *itself* be conscious? If we could see this collection on a greatly magnified level, our visual experience would be much as it would be if we were to witness an invasion of earth by a giant armada of spaceships. The idea that, in addition to the experiences had by the pilots of the ships, there might be a further experience had by the collection comprising the pilots and ships is absurd.

(Block nevertheless concludes that the system constituted by this collection *would* be conscious. This is because Block assumes from the start that we are composite objects. He says, "Since we know that *we are brain-headed systems*, and that *we have qualia*, we know that brain-headed systems can have qualia" (p. 281). Of course, *supposing* that we are *identical* to brain-headed systems, and supposing that these systems would survive the envisaged alien invasion, *then* the elementary-particle-people system would be conscious. But this begs the question of whether we are, or can be, identical to composite objects. The relevant question is not: supposing that we are identical to composite objects, might such-and-such composite object be conscious? But rather: suppositions and philosophical theories aside, might the collection of elementary-particle-people, or the system that comprises it, itself be a subject of experience? To which the intuitive answer is: *no*.)

Next we shift our attention from collections to the *structures* they sometimes exhibit. Consider again Block's example of the miniature men in the head. Imagine that the miniature men got inside the

head as follows. Very gradually, every neuron of a healthy human brain was replaced with a miniature, functionally equivalent man. At the end of the process, billions of miniature men came to constitute a brain-like structure inside the head. Now, there is no problem imagining that the person whose brain undergoes this process survives; however unlikely, the person might remain conscious throughout the ordeal. What seems hard to imagine, rather, is that the structure constituted by the billions of little men might *itself* be conscious. To my mind, the idea that this structure of little men might itself experience, say, the taste of blueberries, seems no less absurd than the idea that the collection of little men which constitutes the structure might itself experience the taste of blueberries. Shifting our attention from the collection of little men to the structure they exhibit does not seem to make any difference. Combining *Number*, *Relation*, and *Structure* will not, then, suffice to explain The Datum.

Last we eliminate the difference in the *nature* of parts. Here I adapt an example from Peter Unger (1990), who adapts his example from Arnold Zuboff (1981). Imagine that the neurons of your brain are gradually separated from one another without interrupting the flow of communication within your nervous system. The separation proceeds in stages. First, your brain is removed from your body and separated into halves: the hemispheres are placed in nutrient-rich vats miles apart from each other and the de-brained body; radio transceiver devices are implanted at the interfaces of both hemispheres and the peripheral nervous system. Because radio signals travel at the speed of light, and because ordinary cross-synaptic signals travel at far lower speeds, normal communication within your nervous system can be preserved. In the next stage, the halves are themselves halved: each brain quarter is fitted with transceivers and placed several miles from the others. The process is repeated until each of your neurons sits in its own container, miles from the others, hooked up to a complex radio transceiver. Throughout the procedure the system as a whole maintains its functional integrity. Now, to add a further twist to Unger's Zuboffian story, imagine that each neuron is paired up with an understudy—a person who learns to function, at the level of radio inputs and outputs, just as the neuron does. Mondays are then declared “Give a Neuron a Break Day”; on this day the neurons are allowed to rest, while their respective understudies operate their radio transceivers for them.

One question is whether you would survive such a procedure. Perhaps you would. Perhaps you would not. We need not take a stance on this question. For the question that concerns us is whether you (or any other conscious being) might be *identical* to the scattered system that controls your body. On Mondays, this system comprises billions of people operating billions of radio transceivers. On other days, it comprises billions of neurons operating billions of radio transceivers. In any case, it comprises billions of objects scattered about the surface of the earth, and it interacts with your body just as your brain would have, had it remained confined to your cranium. Now, the idea that you might be identical to this system on Mondays is absurd. But so is the idea that you might be identical to this system on any other day. Whether the system controlling your body comprises billions of people or billions of neurons seems irrelevant to whether the system might itself be a subject of experience.

One might worry that the scattered state of the system disqualifies it from being a genuine *structure*. To address this worry, and to ensure that we have completely closed the gap between a pair of people and the human body, we now consider the human body itself, with all its parts intact.

We can make salient the composite aspect of the body without envisaging any changes to the body itself. Instead of manipulating the body, we can manipulate our images of it. Imagine for instance that we are fitted with a series of magical goggles. Each pair provides a higher resolution image of McCartney's body than the preceding pair. Without any goggles, McCartney's body looks like a solid, human-shaped, blob. The first pair enables us to see the billions of individual cells that make up McCartney's outer layer of skin. The cells are packed so tightly together that body still looks like a solid blob, though one with an intricate pattern on its surface. The second pair is truly magical: it enables us to see the atoms that make up McCartney's body. Because the atoms are separated by relatively large regions of space, McCartney's body now looks like a scaled-down galaxy of stars. This effect is exaggerated when we don the final pair: it provides us with ultra-fine-grained vision that allows us to see the sub-atomic particles that make up McCartney's body. Our visual experience is now very much like it would be if we were to gaze into outer space on a clear night.

With our most powerful goggles on, we ask: might the system of widely scattered particles before

us itself be a subject of experience? Here it is easy to heed Hume’s warning not to “attribute a simplicity” to this system, and not to “feign a principle of union as the support of this simplicity.” It is easy to take the system for what it is: a *structure* of quadrillions of particles. The structure is not some simple object that pops into existence once the particles are so related; at any moment, it *consists* in the particles’ being so related. To be sure, a simple object may pop into existence whenever particles are so related, and such an object may itself be a subject of experience. But the idea that this system of particles—considered *as* a system of widely separated objects—might *itself* experience something, say, the taste of McDonald’s French fries, seems no less absurd than the idea that a galaxy of stars might itself experience something.

I conclude that no combination of *Number*, *Relation*, *Nature*, and *Structure* can explain The Datum.

6. Simplicity

Is a pair of people itself disqualified from being conscious because it is not *simple*?

I think this is the best explanation of The Datum. In all of the hypothetical scenarios we have considered, a composite entity is presented to our minds *as a composite*, and we are asked whether the entity might itself be a subject of consciousness. It does not matter whether the entity has two, two hundred, or two trillion parts; it does not matter whether its parts are people, dogs, neurons, stars, or sub-atomic particles; it does not matter whether its parts bear the relations typically borne by stars of a galaxy, neurons of a brain, or sub-atomic particles of an entire human body; and it does not matter whether it is a mere collection or a structure. What matters is whether the entity is presented to our minds *as a composite*. If so, we find absurdity in the idea that it might be *identical* to a subject of experience. This suggests that what explains The Datum is *Simplicity*: pairs of people are disqualified from being conscious because they are not simple. The only reasonable rival explanations are the various combinations of *Number*, *Relation*, *Nature*, and *Structure*, but we have seen that no such combination can explain The Datum. I conclude that *Simplicity* best explains The Datum.

7. Conclusion

I have provided an argument that conscious beings must be simple. Granted that we are not simple physical particles, this argument goes against materialism. In addition to my argument, there are a variety of other anti-materialist arguments. Then there are arguments on the other side. While it is admittedly difficult to see which of the arguments is stronger, I want to conclude by giving an initial consideration for thinking that my argument is stronger than what is perhaps the most salient argument in favor of materialism, namely, that the best explanation of the systematic correlations between our mental states and our brain states is that we are identical to our brains.

My argument rests on the following premises:

- P1** **The Datum** (for any pair of people, it is impossible for the pair itself to be conscious).
- P2** **Something must explain The Datum** (there is at least one non-trivial feature that no pair of people could itself have, but which every conscious being must have).
- P3** **If no combination of *Number, Relation, Nature, and Structure* explains The Datum, then *Simplicity* explains The Datum.**
- P4** **No combination of *Number, Nature, Relation, and Structure* explains The Datum.**

P1 and **P2** seem obviously true. A reason to accept **P3** is that the various combinations of *Number, Nature, Relation, and Structure*, on the one hand, together with their rival, *Simplicity*, on the other, appear to exhaust the reasonable options for explaining The Datum. Reasons to accept **P4** were given in my discussion above.

On the other hand, the correlations between our mental states and the states of our brains give us some reason to identify ourselves with our brains, for this identification would begin to explain the correlations. Of course, if we are identical to our brains, we are not simple. So these correlations give us some reason to think that conscious beings are not simple and thus need not be simple.

I doubt that this reason is strong enough to countervail our reasons to accept **P1 – P4**. For

imagine a scenario in which it is common knowledge that our heads are filled with little men who, in concert, control our bodies, and that our mental states are correlated with the states of the little men in a way that mirrors the actual correlation between our mental states and the neural states of our brains. Certainly this hypothetical correlation would give us *some* reason to identify ourselves with the systems of little men in our heads. But would it be strong enough to countervail our reasons for accepting **P1** – **P4**? Could we really take seriously the idea that we were *identical* to collections, or systems, of people? I doubt that we could. For I doubt that we could take seriously the idea that a collection, or a system, of people might itself be a subject of experience. But if this hypothetical correlation would not give us strong enough reason to abandon my argument, then it is doubtful that the actual correlation between our states of mind and the states of our brains gives us such reason. For the only difference between the two cases is a difference in the *nature* of the parts of the candidate systems. And above I gave some considerations for doubting that a difference in the nature of the parts could be relevant.

Evidently, materialists are by and large receptive to these sorts of considerations; for similar considerations are behind the materialist trend to abandon identity-theory in favor of functionalism—a theory on which the nature of the parts of a system is irrelevant to whether the system is itself conscious. One of the main appeals of functionalism is that it respects the idea that a mere difference in the nature of the parts of a system cannot matter to whether the system is conscious. Hilary Putnam defended this idea on empirical grounds, when he first proposed functionalism in 1967. Ironically, Putnam was simultaneously unwilling to swallow the idea that a system of conscious beings might itself be conscious. In the midst of arguing that the nature of the parts of a system is *not* relevant to whether something is in a given mental state, Putnam added to his functionalist analysis a condition that forbids a system of conscious beings from itself being conscious: “No organism capable of feeling pain possesses a decomposition into parts which separately [are capable of feeling pain]” (1967, p. 227). Why did Putnam do this? The purpose of the condition, according to Putnam, was “to rule out such ‘organisms’ (if they can count as such) as swarms of bees as single pain-feelers” (op. cit.). Apparently, Putnam did not see the tension in his own views. He held (a) that the nature of the parts of a system is irrelevant to whether the

system is conscious (because it is irrelevant to whether the system has a given functional organization) and (b) that it is impossible for a system of conscious beings to be conscious. But he did not hold what follows: that it is impossible for *any system* of things to be conscious. (To see how this follows, consider a system of people with an arbitrary functional organization. By (b), the system is not conscious. Now replace the people with anything you like, say, neurons. By (a), this change should not matter to whether the system is conscious. So the new system is not conscious. Thus, it is impossible for any system of things to be conscious.) Instead, Putnam added a condition to his functionalist analysis that goes against the very spirit of functionalism. Better, I think, for Putnam to have endorsed what follows from his own convictions—(a) and (b)—namely, that conscious beings must be simple.

I doubt, then, that the systematic correlations between our mental states and our brain states gives us a reason to identify ourselves with our brains that is strong enough to countervail the reasons that I have given to accept that conscious beings must be simple. I leave it as a challenge for the materialist to provide some other countervailing reason.⁵

REFERENCES

- Barnett, D. *Forthcoming*. “The Simplicity Intuition and Its Hidden Influence on Philosophy of Mind,” *Noûs*.
- Bennett, J. 1967. “The Simplicity of the Soul,” *Journal of Philosophy* 64: 648-60.
- Block, Ned. 1978. “Troubles with Functionalism,” In *Readings in Philosophy of Psychology* 1980, edited by N. Block, Cambridge: Harvard University Press, pp. 268-305.
- Chisholm, Roderick. 1991. “On the Simplicity of the Soul,” *Philosophical Perspectives*, 5, pp. 167-81.
- Foster, John. 1991. *The Immaterial Self*, New York: Routledge.
- Hart, W.D. 1988. *The Engines of the Soul*. Cambridge, U.K.: Cambridge University Press.
- Hasker, William. 1999. *The Emergent Self*. Ithaca: Cornell University Press.

⁵ For helpful comments and discussion, I am grateful to George Bealer and Adam Pautz.

- Lowe, E.J. 2001. "Identity, Composition, and the Simplicity of the Self," In *Soul, Body, and Survival*, edited by K. Corcoran. Ithaca: Cornell University Press.
- Putnam, Hilary. 1967. "The Nature of Mental States," In *Readings in Philosophy of Psychology 1980*, edited by N. Block, Cambridge: Harvard University Press, pp. 223-31.
- Taliaferro, Charles. 1994. *Consciousness and the Mind of God*. Cambridge, U.K.: Cambridge University Press.
- Unger, Peter. 1990. *Identity, Consciousness, and Value*, New York: Oxford University Press.
- Zimmerman, D. 1991. "Two Cartesian Arguments for the Simplicity of the Soul," *American Philosophical Quarterly* 28: 217-26.
- . 2005. "Dualism in the Philosophy of Mind," in *Encyclopedia of Philosophy* 2nd Edition, edited by D. M. Borchert, MacMillan, 2006.
- Zuboff, Arnold. 1981. "The Story of a Brain," In *The Mind's I*, edited by D. Dennett and D. Hofstadter, New York: Basic Books, pp. 202-11.