

CHAPTER 7

From Experience to Experienter

Dean Zimmerman

VARIETIES OF DUALISM AND MATERIALISM

Throughout history and pre-history, the majority view of humankind seems always to have been that there is more to a person than the body, and that an “afterlife” is possible because this “something more” — the soul or spirit — does not pass away with the death of the body. Many philosophers have agreed, developing various forms of mind-body dualism. Philosophical dualists such as Plato, Aquinas, and Descartes — and, more recently, Karl Popper, Richard Swinburne, and William Hasker — disagree about many details. But they have this much in common: they believe that, for every person who thinks or has experiences, there is a thing — a soul or spiritual substance — that lacks many or most of the physical properties characteristic of non-thinking material objects like rocks and trees; and that this soul is essential to the person, and in one way or another responsible for the person’s mental life.

Nowadays, this view is often called “substance dualism,” and contrasted with various forms of “property dualism.” Property dualism is the idea that the mental properties of persons are significantly independent of, or in some other way distinct from, the physical properties of persons. The distinction between the two kinds of dualism allows for an intermediate view, the combination of property dualism with substance materialism. On this conception of what it is to be a human person, each of us is a material object — something that, ultimately, is made up entirely of parts that can be found in non-thinking things — but a material object with a special kind of properties — mental properties

or states, varieties of feelings and thoughts — that are at least somewhat independent of our purely physical aspects. This combination of property dualism with substance materialism is sometimes called “the dual-aspect theory.”

The dual-aspect theory, so understood, is a version of substance materialism. As Robin Collins points out in Chapter 9 of this book, there is a version of substance dualism (or “entity dualism,” as he calls it) that also attributes two aspects to persons: a subjective or experiential aspect, and non-subjective aspects that help explain how the soul and brain work together to generate a rich conscious life. Nothing in my chapter turns upon whether the substance dualist should follow Collins in positing a complex, non-subjective aspect to persons. Perhaps *his* sort of dual-aspect theory is correct. What I shall be criticizing, however, are dual aspect versions of substance materialism.

One might worry whether it is even possible for the materialist to formulate a stable kind of dual aspect view. Many substance dualists have claimed that thinking is impossible for mere matter. But this claim surely needs some serious argument; on the face of it, the combination of property dualism and substance materialism seems to be a consistent position. In general, the fact that one class of properties can vary independently of another does not rule out the possibility that some things have both kinds of properties. Substance materialists who are property dualists can point to the example of colors and shapes. Color properties and shape properties seem to be independent of one another. Yet a single object, such as a red ball, can have both; it does not have a part that is red but shapeless and another part that is spherical but colorless. Philosophers who deny substance dualism while advocating a robust form of property dualism are claiming that mental and physical properties are independent in something like the way color and shape are, while affirming that they are attributes of a single object, consisting entirely of ordinary matter.

In this chapter, “dualism” shall usually carry its more traditional meaning: substance dualism. But I will defend property dualism as well, and argue that property dualism makes trouble for the more plausible forms of substance materialism. And this leaves the more plausible versions of substance dualism looking better than one might have thought.

The paper begins with a meditation on the question, “What am I?,” to which dualism and materialism are competing answers. The most plausible versions of materialism and dualism are then described — “garden variety materialism” and “emergent dualism.” Property dualism

is then defended, in its own right. It turns out that, if property dualism is true, objections to dualistic interaction become less pressing. Finally, garden variety materialism is criticized for the vagueness of the material objects it offers as candidates to be me. Their temporal vagueness makes it hard to know what "I" could refer to. And their spatial vagueness leads to more severe problems: given the nature of vague objects, it is hard to see how one of them could be the subject of the fundamental phenomenal states required by property dualism. The upshot: only very weird versions of materialism are left standing, and emergent dualism starts to look like a better alternative. Given the truth of property dualism, it should not be too surprising to discover that substance dualism is true as well.

FINDING ONESELF

Substance dualism and substance materialism are competing answers, at a high level of generality, to the question each of us may ask with the words: "What am I?" (spoken in a metaphysical tone of voice, with emphasis on the word "am"). Answering the question with any specificity turns out to be harder than one might think. These days, many people will take it to be just obvious that we are mere material objects. But there are at once too many candidates, and not enough sufficiently distinguished candidates, for the role of *the* material object with which I am supposed to be identical.

The problem can be approached by way of the unity of consciousness. Each of us knows that whatever is asking the question "What am I?" must be a single thing capable of exemplifying a plurality of psychological properties. Its unitary nature consists in the impossibility of its exhibiting a certain sort of "division of psychological labor." If a single thinker can recognize the difference between sounds and colors, this thinker does not enjoy the ability to compare the two simply by having one part that does its seeing and another that does its hearing, even if these parts are tightly bound together. As Brentano put it, this "would be like saying that, of course, neither a blind man nor a deaf man could compare colors with sounds, but if one sees and the other hears, the two together can recognize the relationship."¹

The unity of consciousness poses a difficult question: what is this single thing that has auditory and olfactory and visual and tactual and gustatory sensations? I know that I am whatever thing it is that has all of the sensations I am now having; how could I fail to be? But can I know

what kind of thing this is; that is, can I know what *other* characteristics I have?

We can't tell, merely by thinking about the meaning of the word "I," what the word refers to in our mouths — nothing about the function of the word "I" will tell us much about the intrinsic nature of persons. If dualism is false, it's what philosophers call an *a posteriori* falsehood (i.e. something we could only discover to be false by learning things about the world). Likewise, if materialism is true, it's an *a posteriori* truth. In other words, I need to learn about what kinds of things the world contains, and which ones are most closely connected to my conscious states, before I should reach any conclusions about the kind of thing I am.

My exploration of the contents of the world, in search of myself, may begin with mere introspection — "looking into my mind" and seeing what it contains — but it cannot end there. Introspection reveals "bodily sensations," but the phenomenon of phantom limbs should convince us that such sensations do not require that one actually *have* the bodily parts one *seems* to feel. Perhaps a version of materialism is true according to which my leg is a part of me; but the mere fact that I have bodily sensations *as of a leg* does not mean that I have a leg as a part. Conversely, if I have taken certain drugs or undergone a "near death" episode, introspection may reveal what feels like an "out-of-body experience"; but again, the experience may be an illusion. Drugs, brain trauma, and psychological illness may cause all sorts of delusional but utterly convincing experiences. The thesis that George Graham calls "weak ontic ignorance" seems very plausible: the intrinsic nature of the self is simply not given in introspection.²

Recent work in the philosophy of language can help explain why the self proves so elusive — how it is that our own intrinsic nature can be opaque to us. Our self-conception "leaves a blank" to be filled in by how things happen actually to be. The function of the word "I," and of the concept of *myself*, is to pick out the thinker of these thoughts, no matter what that thing is like intrinsically (whether material or immaterial or something else), and no matter what I may believe about my own nature. And that is why I cannot figure it out, *a priori*, just by reflecting on the meaning of "I."

A person, like me, who thinks he's an immaterial soul uses "I" in roughly the same way as a person who thinks that he's a brain, or a body, or even (like some "madmen" Descartes mentions) that he is made out of glass. Changing your views about your intrinsic nature doesn't change the general rules in virtue of which "I" picks out its referent.

And such changes may well fail to shift the referent of the word “I” in your mouth.³

This is merely the application, to first-person pronouns, of what has come to be the orthodox view about how natural kind terms and indexicals function. (Indexicals include terms like “I,” “now,” “here,” and “that” or “this” when these latter two words are used by a person pointing to things.) Here is Thomas Nagel’s comparison of “I” to the natural kind term “gold”:

The essence of what a term refers to depends on what the world is actually like, and not just on what we have to know in order to use and understand the term. I may understand and be able to apply the term “gold” without knowing what gold really is — what physical and chemical conditions anything must meet to be gold. My prescientific idea of gold, including my knowledge of the perceptible features by which I identify samples of it, includes a blank space to be filled in by empirical discoveries about its intrinsic nature. Similarly I may understand and be able to apply the term “I” to myself without knowing what I really am. In Kripke’s phrase, what I use to *fix the reference* of the term does not tell me everything about the nature of the referent. . . . Various accounts of my real nature, and therefore various conditions of my identity over time, are compatible with my concept of myself as a self, for that concept leaves open the real nature of what it refers to.⁴

To sum up, then: materialism isn’t shown to be true just by reflection, or by *a priori* reasoning; it’s an hypothesis about the referent of “I” — perhaps the most plausible hypothesis, at least on first blush, but not something revealed to us by armchair reflection. It is not obvious to me that substance dualism is in a better position, given just *a priori* evidence. At least, I shall not try to show that dualism can be conclusively supported by *mere* reflection upon my experience; nor will I try to argue, *a priori*, that no thinking thing could be made out of the kinds of particles that constitute my body. It is time to get out of the armchair and look more closely at the contents of the physical world, in order to see whether there is a place for ourselves.

THE VARIETIES OF “GARDEN VARIETY MATERIALISM”

Some versions of materialism give quite implausible answers to the question, “Which physical thing am I?” Descartes mentions some insane materialists who believe “that their heads are made of earthenware,

or that they are pumpkins, or made of glass.”⁵ A couple of otherwise sane contemporary philosophers have taken seriously the thesis that we are tiny physical particles lodged somewhere in our brains.⁶ But more popular by far (and rightly so) are versions of materialism that pick more familiar physical objects to be me — what I’ll call “garden variety material objects,” the kinds of things we bump into every day, and for which we already have names. Tables and chairs, and many of their detachable parts (for example, chair and table legs, table-tops, seats, cushions, nuts and bolts) are garden variety material objects; as are trees, and many of their parts (such as their seeds, leaves, bark, limbs, and roots). But the best candidates for being a human person, among garden variety material objects, are human organisms and their familiar, naturally demarcated parts — most especially, our brains.

Here is what I mean by calling a part of a living body a “natural part”: its spatial boundaries are reasonably sharply defined, and, assuming that it is made out of parts itself, these further parts work together to perform some function. Examples include: a single atom within a strand of DNA, the heart, the kidneys, the spine, an individual blood cell, the respiratory system, the entire nervous system, the brain, the cerebrum, a single cerebral hemisphere, and the complete organism (the “biggest part”). Basically, if it is worth listing in a book about human physiology or biochemistry, it will count as a natural part, for my purposes. Now that we have had plenty of experience examining the insides of mammalian bodies, including human bodies, all these organs have become “garden variety material objects” to us — just as examination of a flower reveals petals, stamen, pollen, and so on. Among the natural parts of a human body, some are better candidates than others for being the person associated with that body. For one thing, if there is a physical “organ of thought” — a smallest natural part that includes all the parts upon which my ability to think most immediately depends — then I surely ought to have that organ as a part. The parts that fit these criteria are primarily: the complete organism I refer to as “my body,” the entire nervous system within it, the brain, the cerebrum, and perhaps one or the other single hemisphere of that cerebrum. These are garden variety objects, in the sense that they are the physical objects with naturally demarcated boundaries that we find when we stroll through the “garden” of the human body; and I will sometimes call them “the standard candidates” for being me. Garden variety materialism, then, will be the thesis that each human person has the size and shape of one of these naturally demarcated, garden variety parts.

I am only interested in one fact about the standard candidates: all

of them, being garden variety objects, are vague in their spatial and temporal boundaries. I shall argue that this vagueness raises insuperable difficulties for standard materialism, leaving dualism looking better off — for it will then only be competing with versions of materialism that pick unfamiliar physical objects to be me, and so far, the pickings are slim. In these circumstances, materialists must adopt a more speculative form of materialism; and, so far, there has emerged no such version of materialism that would compare favorably with the better versions of substance dualism — to which I now turn.

AN ALTERNATIVE TO STANDARD MATERIALISM: EMERGENT DUALISM

I now describe the version of dualism I find most appealing, the one I believe to have the best chance at competing with substance materialism, namely, the kind of “emergent dualism” defended elsewhere in this volume (Chapter 8) by William Hasker. Although something like it has been widely accepted by ordinary folk all over the world and through the millennia, it may strike many scientifically educated people as ridiculous, an exploded myth. What I hope to show is that, when all is said and done, it is at least no crazier — no more improbable, on first blush, as an answer to the question “What am I?” — than the more speculative versions of materialism to which I shall attempt to drive the dual aspect theorists.

The empirical facts strongly suggest that human minds are dependent, both for their existence and many of their characteristics, upon brains. Some dualists have denied this. Descartes thought that: (i) no mere brain could produce conscious states without interacting with a soul, (ii) brains are not themselves capable of generating souls naturally, and (iii) God does not care to work the miracle necessary to bring a brain into interaction with a soul in the case of non-human animals. And so Descartes was led to deny that any non-human animal has conscious experiences. But surely at least the higher mammals *are* conscious; so at least one of these three theses is false. Do all sentient creatures have souls, then? If (i) is true, and brains alone cannot produce consciousness, then they must. But, according to (ii), brains do not naturally produce souls, and so each animal soul would have to be specially created by God, just like human souls. Some dualists have accepted this result, rejecting (iii): God does intervene in nature, distributing animal souls wherever and whenever they are needed. Assigning God this role

would be a relatively minor departure from fully fledged Cartesianism. But would there really have to be so much divine tinkering, simply in order to insure that each organism with a sufficiently complex nervous system is able to be conscious? Couldn't God have designed creatures in which consciousness arises *naturally*? Many dualists have thought it would be sloppy for God to create a world requiring nearly constant miraculous intervention. And, despite a close association between dualism and theism, many dualists have not believed in a personal God who could intervene, or in any sort of God at all. One needn't be a Christian, like Descartes, in order to be a dualist. And of course atheistic or deistic dualists cannot follow Descartes in supposing that divine intervention is needed to introduce souls into the natural order.

Many dualists, then, will not want to reject (iii). They must, then, reject either (i) or (ii) — and, with it, some other part of the Cartesian picture. One could, I suppose, reject (i), holding instead that, in non-human animals, there are conscious states but no souls. On this view, a mere organism, or a brain, or some other physical part of the animal can have experiences; no soul is needed. But here is an argument that (i) should be retained, and that it is rather (ii) that must go: If the events in the brain of a chimp were causally sufficient to confer conscious states upon its body or brain, then the very similar events in my skull ought to do the same for *my* body or brain.⁷ But dualists want to deny this — it is, in human persons at least, the *soul* that has the conscious states. Consequently, the dualist should reject (ii): brains like ours, and also those that are less complex but still quite similar, are naturally capable of causing souls to exist.

Some contemporary dualists, like William Hasker⁸ and Richard Swinburne,⁹ accept this conclusion, advocating a view sometimes called “emergent dualism”, whereby organisms having nervous systems complex enough to generate conscious states automatically *also* generate nonphysical subjects for those states. Though brains and souls share no parts in common, each soul remains radically dependent upon one brain for its continued existence and for many, if not all, of its powers and dispositions. Since Hasker and Swinburne believe in an afterlife, they affirm that God could (and does) miraculously prevent the dissolution of the soul that would (or at least might) naturally occur when the nervous system upon which it is dependent ceases to function.¹⁰ Hasker also supposes that each nonphysical subject is located somewhere within the nervous system that generates it.

Some will say emergent dualism is not *real* dualism, reserving the name for nothing but Cartesianism. But why should Descartes's version

of dualism be the only game in town? In order to count as a genuine dualism of substances, a theory must say that persons, unlike plants and the bodies of animals, are not made of the same kinds of stuff as ordinary inanimate objects. Where dualists disagree is over just how radically different from ordinary matter these new substances, or souls, must be. A soul posited by a particular version of dualism will seem less physical the fewer the number of properties that are said to characterize both substances capable of thought (and their parts, if any) and substances utterly incapable of thought. As a matter of fact, dualists have disagreed about how much souls have in common with ordinary matter; they have meant different things by calling souls “nonphysical.” The result is a spectrum of dualisms, with Cartesian dualism near one end and emergent dualism closer to the other end.

The maximal difference a dualist might posit between soul and body would be to identify souls with necessarily existing abstract objects, outside of space and time, like numbers or Plato’s Forms. This sort of dualism goes even further than Descartes’; it makes souls out to be even more radically unlike material objects. But almost no one would want to accept a dualism so radical as that.¹¹ Almost all dualists will agree that souls have this much in common with ordinary material things: they are concrete entities, existing in time, and capable of change.

Descartes allowed at least *that* much similarity between souls and ordinary matter, but little more. Cartesian souls have no position in space. Descartes also claimed that souls are “simple,” or without parts. Since he believed that everything in space was infinitely divisible, this was another way in which souls were unlike anything made of ordinary matter. Dualists who deny these aspects of Descartes’s particular form of dualism are merely staking their claim at a different location along the spectrum of possible dualisms about persons.

Emergent dualism is much less extreme than Cartesian dualism. Emergent dualists make souls a part of the natural order, generated by any brain sufficiently complex to support conscious experience. If souls are in space, some of the worst problems of interaction are easily solved. Although I shall not try to defend the claim here, I am convinced that the better arguments for dualism do not require that the soul have all the features Descartes attributed to it, or that it lack all those he withheld from it. Less radical dualisms are, in fact, safer — they posit no more differences between souls and material objects than are required by the reasons for rejecting materialism.

PROPERTY DUALISM

I have claimed that emergent dualism is less implausible than other forms of dualism, and that substance materialism is not conclusively established simply by reflecting upon the nature of my experience — for example, from the fact that I can “feel my limbs” it does not immediately follow that I am identical with the body that has those limbs among its parts. One may, however, grant all of this while nevertheless insisting that, *obviously*, we are mere material objects. Materialism will seem to many to be the default view about our nature. We know that human brains and bodies exist, and that they are entirely and unproblematically physical; we know that our ability to think and feel and act is radically dependent upon their proper function. Why go out on a limb, positing the existence of some extra, nonphysical thing? In the absence of a compelling argument for immaterial souls, the only sensible conclusion is that we are entirely physical in nature. At least, that’s how many materialists will see the matter.

Are there positive reasons to posit immaterial substances in addition to the material substances that constitute our bodies? Many philosophers have given arguments for the conclusion that human persons are immaterial souls; some are more impressive than others.¹² The line of thought I shall pursue here is to argue, first, for property dualism. The considerations I raise are familiar ones, and they have convinced many philosophers that, even if we are made entirely of physical particles — just like rocks and trees — we nevertheless must have a “side” to us that is independent of our physical nature. Our conscious experience presents us with an “aspect” or set of properties that is not fully determined, in any very strong sense, by the properties we share with unconscious physical objects. In subsequent sections I will argue that, given property dualism, garden variety materialism cannot be maintained, and more speculative forms of materialism turn out to be even less plausible than emergent dualism.

The case for a dualism of mental and physical properties is most compelling when it focuses on conscious experience, especially the distinctive “way that it feels” to have experiences of different kinds — what philosophers sometimes call the “phenomenal aspects” of conscious experience. Some mental events have no distinctive phenomenal aspects; there is, for example, no single, characteristic way that it feels to believe the Pythagorean theorem. Experiences, however, do include phenomenal aspects (or, in philosophers’ jargon, qualia). Most people who look at a stop sign experience phenomenal redness — the quality

(or quale) that fills an octagonal part of one's visual field when looking at the sign, and which also turns up in red after-images or the hallucination of a bloody dagger. Similar things can be said about the phenomenal aspects of tastes, smells, sounds, pains, tickles, and so on.

Qualia make trouble for materialists. Today's materialists have learned to live with little agreement among themselves about the nature of mental states; and they would not presume to guess what physics will finally say about the nature of matter. In the absence of a positive consensus, they have rallied their forces around the following more general doctrines, under the banner "physicalism": the universe consists entirely of atoms in the void, or particles and fields, or hyper-dimensional superstrings, or whatever physics ultimately settles upon as the terms of the most fundamental causal transactions. These basic physical entities, physicalists suppose, do not include minds or anything else with the tincture of mentality about it. And everything that happens in the universe boils down to nothing but mindless, physical interactions among these basic entities.

There is controversy about what "boiling down" requires, but most who accept the label "physicalist" seem to agree upon two components. (i) "Higher level" phenomena — biological, psychological, sociological, and so on — are determined by what goes on at the basic physical level. A universe that exactly duplicates our distribution of matter throughout space-time must include organisms, thoughts, and political movements exactly like ours. (ii) "Higher level" phenomena must be ultimately explicable in basic physical terms. It may not be very useful, given our usual purposes, to describe a case of cirrhosis of the liver, or a red sensation, or a revolution, in terms of the activity of subatomic particles, and the laws governing such phenomena may not be reducible in any tidy way to laws of physics. But, physicalists will insist, there must in principle be a story that could be told that would show how all the facts about human beings — including facts about diseases, sensations, and wars — are necessitated by the physical facts discussed by physics.

To some philosophers, this version of physicalism seems almost to be part of "enlightened common sense." A vocal minority disagrees, offering several arguments against physicalist orthodoxy. Here, I shall focus exclusively upon two such arguments, based upon famous thought experiments. One can tell stories about creatures who behave like humans, but lack all phenomenal experiences ("zombies"), and one can tell stories in which a subject's phenomenal experiences change in systematic ways, but leave the physical states of the brain unchanged (inverted spectrum arguments), and although these stories may sound

bizarre and improbable, they do not strike us as outright incoherent and impossible. And if these stories really do describe some possible creatures, however “far out” and unlikely they might be, then physicalism is false.

In the philosophy of mind, as in the cinema, zombies are everywhere. But the philosophers’ zombies do not gibber and drool and eat brains. A philosophical zombie is a creature that is outwardly and behaviorally exactly like a normal human being, and is even perhaps identical in its internal physical makeup, but is somehow completely devoid of conscious experience. Philosophers tell stories about these imaginary creatures, hoping to shed light on the relationship between mind and matter. When pricked, a philosophical zombie bleeds, and says “ouch!” But it has no feeling of pain. Its eyes respond to light just as ours do; it says “Bananas are yellow,” and it won’t eat green bananas. But it never experiences yellow qualia, never has a yellowish patch of color in its visual field — in fact, it has no visual field. The zombie experiences none of the qualia we know through taste, smell, hearing, touch, and other forms of sensation (philosophical zombies are more like the angels in the German film *Wings of Desire* than the “living dead”).

Are zombies *really* possible? Consider a body somewhat like ours, but controlled remotely by radio communication with a giant computer. Ordinary human beings may be extremely complex, but it is plausible to suppose that our complexity could, at least in principle, be replicated by a monumentally complex computer. Putting these two thoughts together, there could, in principle, be creatures behaviorally indistinguishable from us but with computers for “brains.” Would they be conscious? Would they feel pain, experience our spectrum of colors, and so on? The natural reaction is: Who knows? How could we tell whether exactly duplicating the human brain’s functional capacities in different “hardware” would generate qualia? It is an open question, one that we human beings might never be able to answer, were we actually confronted with such creatures. A growing number of philosophers of mind believe that this natural reaction is the right reaction to have — it represents a positive insight into the nature of qualia, the presence or absence of qualia within a given computer-brain is a merely *contingent* matter. It is possible — one wants to say, “God could make” — computer-brains with qualia, and ones without. But then it should be a merely contingent matter whether *other* complex objects generate qualia — including objects that look, from the outside, just like our brains. Of course *we* know that *our* brains do — we know this “from the inside.” But if it is a contingent matter whether non-organic beings have conscious experience, it must

be a contingent matter whether organisms more closely resembling us have conscious experience. So zombies are possible.

If creatures completely devoid of experience seem too outrageous, there are less extreme scenarios that work just as well, for the property dualist's purposes. One needn't be a philosopher to wonder whether one's own experiences of color might not be the same as others'. The colorblind miss shades that the rest of us see, and some kinds of animals can see wavelengths of light that make no impression on us. How do the colors they see compare to those present to one's own consciousness? The natural reaction is: there must be an answer, but there is no telling what it is.

The possibility of more radical shifts in visual qualia can turn this reaction into an argument for property dualism. Suppose goggles were constructed that would systematically shift one's experiences of shades of color. Whenever one would normally have an experience as of something red, the goggles cause an experience as of something violet; whenever one would normally experience orange, the goggles cause an experience of indigo; and so on, inverting the entire spectrum of experienced colors (except for one lone shade of green, in the middle, which the goggles leave alone). Someone fitted with the goggles from birth would learn to use color words in the same circumstances as the rest of us, and would discriminate shades just as finely as the rest of us, but her experiences would be radically different, phenomenally, from ours — as she would discover, when the goggles were finally removed. If goggles can invert the spectrum of experienced colors, so could the right sort of interference with the workings of the eyes, or the optic nerves. Presumably, the trick could be pulled off by fiddling with processes still deeper in the brain, where our color experiences are “constructed” (the final character of the visual field is based rather loosely on the information provided by the rods and cones in the eye).

These possibilities show that no amount of attention to the color discriminations a person is able to make, or even the way her eyes respond to light, will rule out the possibility that her experience, when she says something “looks red,” is phenomenally like our experience when something looks violet to us. Still, it would seem safe to conclude that, if two humans have equally good color vision, and similar eyes and optic nerves, then the way their visual cortexes work is probably similar, and that similarity at that level should be reliably correlated with similar experiences — so one could reasonably assume that most humans have the same qualia when they say something “looks red.” But suppose we encounter a species of intelligent alien that registers visual information

using binocular eyes much like ours, and describes experience using words that seem to correspond to “red,” “green,” and so on; however, in these aliens, the brain is made of quite different stuff. We can no longer infer similarity of their qualia with ours on the basis of similarity in the mechanism by which color experience is ultimately generated within the brain. Two genuinely different possibilities present themselves: colored objects may look to them as they do to us now, or objects may look to them as they would look to us were we wearing the spectrum-inverting goggles. (No doubt there are other possibilities as well.)

There will be a physical story to be told about the generation of color experience in alien brains by wavelengths toward the red end of the spectrum; and whatever it is, the question will remain: Is an experience similar to ours, when things appear red, hooked up with this physical process? Or is it an experience of phenomenal violet, instead? (Or does the experience have some quale from a range of phenomenal colors with which we are unfamiliar?) Whatever the answer, it *feels* as though it should be a contingent one. These particular wavelengths of light were associated, in the aliens, with the range of color experiences enjoyed by normal folks, say; but they *could* have been associated with the kinds of experiences enjoyed by the goggle-wearer. And if that is the right thing to say about the aliens, one should also say it about us. As a matter of fact, God (or Nature or whatever) built us in one of these ways; but creatures could have been designed along physically similar lines, but with inverted phenomenal experiences arising from the workings of their otherwise identical brains.

Once you accept the bare possibility of *either* zombies *or* creatures just like us but with inverted spectra, you have rejected physicalism in favor of a dualism of mental and physical properties. Take all the facts about the world that can be stated within a “final physics” that mentions no mentality. If zombies are possible, it would remain an open question whether the world contains phenomenal experience at all. If it is possible for there to be physically similar creatures with systematically shifted color experiences, it would remain an open question *which* individuals had *which* qualia. To settle the matter, one would need to know about some extra laws of nature, linking brain states and qualia.

The conclusion supported by appeal to the possibility of zombies, inverted spectra, and the like is, I take it, a thesis about which kinds of similarity and difference “carve nature at the joints.” Plato’s metaphorical talk of “joints” in nature is a way of expressing the idea that some properties are more “natural” than others, some kinds of similarity are deeper than others. The truly natural properties are the ones responsible

for the most fundamental kinds of objective resemblance among things. Naturalness comes in degrees because resemblance comes in degrees; and property dualism is a claim about where phenomenal similarities and dissimilarities lie on the spectrum from more to less natural. Some phenomenal properties or conditions are less than perfectly natural (for example, highly gerrymandered or disjunctive ones, like seeing-red-or-feeling-an-itch; and very general ones, like hearing some sound or other). But, like other families of properties, the phenomenal ones come in more and less precise forms — some are highly determinate or specific, while others are determinable or more general. Examples include: having some mass or other, being roughly one kilogram in mass, and being precisely one kilogram in mass; and having some shape or other, being quadrilateral, and being square. Similarly, the phenomenal aspects of experience come in more and less determinate forms. If someone describes an experience as being “as of something colored,” “as of something red,” and “as of something scarlet,” the property dualist will suppose there is a corresponding series of properties ascribable to qualia: phenomenal color, phenomenal redness, and phenomenal scarlet. In such families of properties, the more general ones are exemplified because the most precise ones are; and the most precise ones are the most natural ones, the real joints in nature. The arguments for property dualism support the conclusion that similarities and differences among experiences, due ultimately to these precise qualia properties, are independent of the similarities and differences determined by physical properties of things.

How natural are these properties? If they are not determined by the physical properties of things, nor by any other family of properties that does not include them, they must simply be another basic respect in which things can resemble one another. When we ask ourselves whether there could be creatures physically like us, but with inverted spectra, we are not, most of us, imagining, in vivid detail, the true neurophysiological side of color experience — since we don’t know its details. We are simply imagining creatures just like us with respect to *whatever* other properties our brains may have, besides the qualia with which we are familiar in experience. There is little prospect of finding some other family of properties — neither those mentioned in physics, nor those discovered in experience — that could be more basic than our qualia, somehow grounding phenomenal similarity and difference in another realm, beyond the reach of physics or our experience.

The property dualist is, then, committed to the idea that our experiences resemble one another in virtue of some family of most basic

phenomenal properties — the most precise qualia, whatever they are and whatever has them — and that these qualia represent perfectly natural “joints in nature” — as natural as the most natural properties that would be mentioned in an idealized “final physics.”

With qualia fundamental, yet obviously caused by the workings of brains, the property dualist will have to suppose that the catalogue of fundamental properties and fundamental laws includes more than just the kinds one finds in physics as it currently stands. Paul Churchland considers the hypothesis that “mental properties are *fundamental* properties of reality . . . on a par with length, mass, electric charge, and other fundamental properties.” Churchland notes that a property dualist might cite, as historical precedent, other cases in which a property was thought to be reducible but turned out to be fundamental — e.g. “electromagnetic phenomena (such as electric charge and magnetic attraction)” which were once thought to be “just an unusually subtle manifestation of purely *mechanical* phenomena” but ultimately had to be added to “the existing list of fundamental properties.”

Perhaps mental properties enjoy a status like that of electromagnetic properties: irreducible, but not emergent. Such a view may be called *elemental-property dualism*. . . . Unfortunately, the parallel with electromagnetic phenomena has one very obvious failure. Unlike electromagnetic properties, which are displayed at all levels of reality from the subatomic level on up, mental properties are displayed only in large physical systems that have evolved a very complex internal organization. . . . They do not appear to be basic or elemental at all.¹³

Churchland’s objection is not a trivial one, and property dualists need to do more than they have so far done to answer it. I will not address it seriously here, beyond a couple of remarks. For one thing, I disagree with Churchland if he is implying that the qualitative properties that characterize my experience — e.g. the particular color qualia exemplified in the parts of my visual field — do not *seem* basic or elemental. They certainly seem so to me (though the complicated patterns in which they occur do not). A second point is relevant to the prospects for substance dualism. Churchland seems to be presupposing that mental properties are exemplified by “large physical systems” that display “complex internal organization”; but, even though property dualists who accept some version of garden variety materialism must accept this, the emergent dualist, at least, need not.

I shall take property dualism to be true, a thesis supported by cogent philosophical arguments having nothing to do with substance dualism

and souls. And I will argue that accepting it makes substance dualism look much better than it otherwise would. Property dualism requires that we posit fundamental laws governing fundamental phenomenal properties. Ultimately, this will raise serious problems for any form of garden variety materialism.

THE STRUCTURE OF PHENOMENAL STATES

Stories about inverted spectra and zombies are supposed to show that, had the laws relating brains and conscious states been different, the objects we see would have appeared differently to us, despite precise similarity in the light waves hitting our retinas and the patterns of neural firing in our brains. Stop signs now appear red, but in the inverted world, they appear violet. Somewhere, qualia have been switched — but what is it that switched properties, what kind of thing has the most fundamental phenomenal properties?

The property dualist has a choice: she can either suppose that qualia are exemplified by some range of things to which the subject is related in experience; or she can regard them as properties had by conscious subjects themselves. Philosophers (and psychologists, when the discipline was younger) have engaged in considerable armchair speculation about the amount and kind of complexity to be found in phenomenal states; and each of the two choices for the subjects of qualia has had its defenders. The most popular versions of the two approaches have been called the “act-object theory” and the “adverbial theory.”

Take the kind of experience I have when I see a stop sign in front of me, or I hallucinate a bright red object before me, or am in some other situation that would lead me to say that something red is in the center of my visual field. To some, it has seemed obvious that *appearing red* is something that can only be done by an object or entity of some kind, distinct from the experiencing subject; to have an experience “as of something red” is to engage in an “act” of sensing which acquires its reddish character from the nature of its “object.” To be an “act-object theorist” about a certain kind of phenomenal experience is to attribute a relational structure to the experience. According to an act-object theory, the distinctive qualia of this type of experience belong to something other than the subject of the experience; and differences among similar types of phenomenal state are construed as differences in the properties had by the entities to which the subject is related. G. E. Moore and other sense data theorists took all phenomenal states to have such an act-object

structure.¹⁴ Sensing is relational; there is no sensing without sensibilia.

Some philosophers reject the analysis of sensory experience in terms of a relation between a thing that is appeared to in a certain way (the subject of the experience) and an appearance (the object). To be sure, in the case of visual experience, it is natural to give an act-object analysis. Even in complete visual hallucination, it seems as though one is related to *some* sort of object — a two-dimensional colored surface, albeit one that turns out not to be part of the surface of any physical object or even to be located in physical space. But it is less obvious that, when experiencing a dull headache, for instance, there is a meaningful distinction to be made between an act of experiencing and a sensible object — a thing that appears to the subject in a dull, painful way. The main alternative to act-object theories of the phenomenal is *adverbialism*: the thesis that all talk about the appearances of things (including visual appearances) should be understood as descriptions of “ways of being appeared to,” so as to avoid commitment to the existence of a special class of phenomenal objects (i.e. the sense-data of Russell and Moore) that can appear even when no physical object is appearing.

When it looks to a person as though there is something red in front of him or her, the person is experiencing “in a reddish way” — “sensing redly,” as Roderick Chisholm put it. The phenomenal quality peculiar to experiences “as of something red” is not borne by a thing to which the experiencing subject is related. “Red,” as a term used to describe types of phenomenal experience, is better construed as an adverb modifying the type of feeling or sensing undergone by an experiencing subject. So such accounts of the structure of experience have been dubbed “adverbial theories of appearing.”¹⁵

Frank Jackson sums up the differences between these two approaches by saying that act-object theorists take an experience to be a “relational state,” involving a person and a sensed particular; while adverbialists take an experience to be a “unitary state,” “a state of that person not essentially involving anything over and above that person.”¹⁶

C. D. Broad considered the relative merits of act-object and adverbial theories under the heading “Are Sensations Analysable into Act of Sensing and Sensum?” Broad discerns a kind of continuum of sensation types:

If we consider the various experiences called “sensations,” we seem to be able to arrange them in an order, starting with those of sight, passing through those of taste and smell, and ending with bodily sensations, like headache. Now, as regards the top members of the series, the analysis into act of sensing and object

sensed seems pretty clear. A sensation of red seems clearly to mean a state of mind with a red object, and not to mean a red state of mind.

If we now pass to the other end of the series the opposite seems true. It is by no means obvious that a sensation of headache involves an act of sensing and a “headachy” object; on the contrary, it seems on the whole more plausible to describe the whole experience as a “headachy” state of mind. In fact the distinction of act and object seems here to have vanished; and, as there is clearly *something* mental in feeling a headache, just as there is in sensing a red patch, it seems plausible to hold that a sensation of headache is an unanalysable mental fact, within which no distinction of act and object can be found.

Now this contrast between the top and the bottom members of the series would not greatly matter, were it not for the fact that the two kinds of sensation seem to melt insensibly into each other at the middle of the series. It is about equally plausible to analyse a sensation of a sweet taste into an act of sensing and a sweet *sumum*, or to treat it as an unanalysable mental fact, having no object, but possessing the property of sweetness.¹⁷

The continuum naturally tempts systematizing philosophers to develop a theory of sensation based on examples from one end or the other, and then to force the whole spectrum of sensory states to fit into a single (possibly Procrustean) bed. Broad resists the unifying impulse; the states we call “sensations” are so called because of their similar causes (each is “the immediate response to the stimulation of a nerve”), but they may be quite different in their intrinsic structure.

I shall set aside the act-object account of the phenomenal, and assume that adverbialism is correct. There are two reasons it is safe to do so. For one thing, I believe the act-object property dualist can hardly avoid attributing qualia to entities like the sense-data of Russell and Moore — things that have size and shape but are very hard to fit into the three-dimensional space inhabited by physical objects. Sense data, so conceived, are not to be found in unconscious material objects, and they are responsible for the fact that we have conscious lives. So the act-object theory leads, if not exactly to a dualism of thinkers and physical objects, at least to a dualism of parts of our experiences and physical objects.¹⁸ Another reason to ignore act-object theories is that adverbialism about even *one* fundamental phenomenal state that humans experience would be sufficient for the argument based on vagueness that comes next. Headaches, smells, tastes . . . these seem to submit most naturally to an adverbial analysis. Smelling a skunky smell is like waltzing a waltz, not like kicking a tire. When you kick a tire, there is

you, your act of kicking, and the tire you kick. When you waltz a waltz, there is you, and there is the act of dancing, but there is not a third thing, something that you do the dance *to*. “The waltz” is a name for the kind of dancing activity in which you are engaged. Likewise in the case of the skunky smell that you smell; “the smell,” in this context, is a name for the kind of activity in which you are engaged — the olfactory experience you are undergoing.

From now on, I shall assume that the property dualist who would reject substance dualism must also reject the act-object theory in favor of adverbialism: the subject of phenomenal experience is the very thing that bears the qualia.

THE VAGUENESS OF HUMAN BODIES AND BRAINS

Garden variety materialism identifies me with a garden variety object, a thing that already has a place in our commonsense conception of the world. Such an object will have relatively natural boundaries, such as those of an organism, or a brain, or even a single hemisphere of a brain. But animals and their organs belong on a spectrum that includes bushes, branches, clouds, mountains, rivers, tidal waves, and all manner of fuzzy entities. These familiar denizens of the physical world exhibit vagueness or indeterminacy in their spatial (and, for that matter, temporal) boundaries. And the strategies typically implemented to resolve the puzzles posed by vague objects do not seem so satisfactory when applied to *oneself*.

All the sensible material candidates for being the referent of someone’s use of the word “I” appear surprisingly like clouds upon close inspection: it is not clear where they begin and end, in space or time. Many particles are in the process of being assimilated or cast off; they are neither clearly “in,” nor clearly “out.” The boundaries of animals and organs are infected with vagueness. Where does the brain end and the brainstem begin? When I hit my fingernail with a hammer, and it slowly blackens and eventually falls off, when exactly did it cease to be a part of my body? If its cells are dead, and it is barely attached, shouldn’t we say it is no longer a part of me? But at what minute or second, prior to falling off, did it cease to be mine?

One might hope to find answers to such questions from biochemistry. Our bodies and brains are made of cells — perhaps biochemistry can fill in the missing pieces, locating a precise parcel of matter that stands out from the rest of the physical world as uniquely my own. If the boundaries

of an organism or organ are to be perfectly precise, the boundaries of the cells that make it up must be perfectly precise. So, what are those boundaries like? The membranes of cells allow liquids and gases to pass through them. Are these substances always, or only sometimes, parts of the cells through which they pass? Does an H_2O molecule, for example, count as part of a cell, no matter how briefly it falls within its borders? What about larger molecules that are selected for transport into the cell by proteins in the cell's membrane? When *precisely* does such a molecule become part of a cell, as a transport protein binds with it, and causes it to pass, as a whole or in parts, through the membrane and into the cell? And when *precisely* does a useless molecule cease to be part of a cell? As soon as it no longer serves a function? Or only after it passes through the cell's outer membrane?

What makes such questions seem hopelessly hard to answer, or simply misguided, is the fact that the binding of proteins and the breaking of molecular bonds include the movements of parts of the molecules in question, and motion seems always to be a continuous process — that is to say, one can take the period during which a motion happens, and divide it up as finely as you like, and for each part of the period there will be a stage of the motion, slightly different from those that came before and after. The chemical reactions involved in one of these changes will be measured in nanoseconds; nevertheless, as a continuous process, the passage of a particular molecule into or out of a cell can be broken down, at least in principle, into an indefinite number of smaller movements and changes in the relationships among the parts of the molecule and other parts of the cell. The withdrawal of the glaciers was a slow, continuous motion; the metabolic processes within a cell are astonishingly fast, but they may be continuous for all that, in which case they pose the same problem, when one considers sufficiently brief periods of time. Choosing the picosecond, say, at which a molecule truly becomes part of a cell is like dating the end of the last ice age to within 1,000 years.¹⁹

If an object is perfectly precise — if it in no way resembles a cloud — there must be an answer, at each instant of the object's existence, to the question: Which other objects (including cells, molecules, and atoms) are to be counted among its parts at that time? When the question is asked about garden variety objects like organs, and organisms, it simply does not admit of a precise answer. To be a part of one of these things is to be caught up in its metabolic processes, and whether a thing is caught up in the metabolism of a cell, say, is often a vague matter.

When dealing with a vague object, it is tempting, when issues about

boundaries become important, to simply *stipulate* answers to questions about precisely which things are or are not among its parts. And why not? When there is no sharp boundary in nature, but we need to draw one for legal or other purposes, we can choose, more or less arbitrarily, among the many equally eligible ways to do it. Where does Mount Everest end, and its foothills begin? We are free to lay down conventions in various ways to settle such questions, when they need to be settled. Similarly, one could lay it down that a certain kind of molecule does not truly become part of a cell until it is entirely within the cell membrane, but one could just as well have chosen the beginning of the process, when it is first bound to a transport protein.

The existence of many, equally good, ways of stipulating a boundary for a vague object provides a key that unlocks the nature of the phenomenon of vagueness. The indeterminacy of the borders of garden variety objects — including mountains, clouds, plants, and organisms — is due to an embarrassment of riches. Wherever we say there is “one” mountain, cloud, tree, or animal, there is really a plethora of what are, in some sense, equally good candidates for being the object in question. I confine my discussion to what I take to be the most popular, and most plausible, approach to the vagueness of mountains, clouds, living bodies, and organs: that it is an essentially linguistic phenomenon, due to “semantic indecision.” We have, by means of the conventions governing our languages, decided to use “mountain” to refer to mountains (rather than to lakes), “cloud” to refer to clouds (and not, say, trees), and so on. But, in doing so, we have failed to specify very precisely what the boundaries are of these things we call “mountains” and “clouds”; there are many different chunks of the earth’s crust, each of which is a good candidate to be “the” mountain to which we refer on some occasion, and many different masses of water vapor that could be selected as “the” cloud to which someone points. Instead of trying (hopelessly) to select just one of them, we gesture indeterminately at them all. We could, if we like, specify somewhat more precise meanings, thereby cutting down on the number of these potential referents of “mountain” and “cloud”; but, given our limitations, there will always be some vagueness about exactly where we want to place the boundaries for objects belonging to these garden variety kinds. Given the vagueness of human bodies and brains, they deserve similar treatment. As with clouds and mountains, the vagueness of bodies and brains is to be accounted for by pointing out that there are many equally eligible candidates for being “the body” and “the brain,” and we have failed to do enough to determine which one we are talking about. We speak of a human body or brain as though

there were just one physical object in the vicinity, when in fact there are many largely overlapping, perfectly precise things.²⁰

If our terms and names for garden variety physical objects refer only indeterminately— if our uses of these words are really vague “gestures” in the direction of a host of objects — how do we ever manage to say anything *true* about these things? The vagueness of our ordinary terms is, generally speaking, harmless, because much of what we say about a garden variety object will be true of *every one* of the precise candidates for being that object. They will all have nearly the same mass, size, shape, color, and so on. When we use a sentence with vague terms, and all the things we could mean by it are true, the sentence itself should count as true. We manage to say true things about vague objects because we are usually not interested in the tiny differences in the properties they have. But suppose some feature — say, *being entirely in Switzerland* — is *not* had by all the good candidates for being “that mountain”; there are places that could reasonably be judged to be parts of the mountain — for example, some piece of ground where the foothills meet the slopes — that are in Italy; though the bulk of the mountain’s mass falls squarely within Switzerland. Then it would be problematic to say: “the mountain is entirely in Switzerland.” Suppose only a few of the good candidates for being “that mountain” stick out into Italy. If I said these words — “The mountain is entirely in Switzerland” — I would *almost* speak the truth. If, however, almost all the good candidates for being “that mountain” overlap Italy, the sentence would be more problematic; it would be *almost* flat-out false.

Applying the theory of vagueness to bodies and brains, one reaches the following conclusion. If I am a garden variety body or brain, there are many human-shaped or brain-shaped physical objects, each an equally good candidate to be what I refer to when I use the word “I.” And, when I say “I have such-and-such feature,” I speak truly only if all the candidates have the feature. Otherwise, what I say is false, or at best *kind of* true. In the next section, I argue that it is very difficult to accept this account of the vagueness of our boundaries, while affirming property dualism and adverbialism about qualia.

ADVERBIAL QUALIA AND VAGUE SPATIAL BOUNDARIES

Given what we know about the close connections between brain activity and phenomenal experience in our own case, laws of qualia generation dictate that, *very* roughly, whenever some neurons are organized and

behaving like so — e.g. like the ones in my brain right now — something-or-other will be caused to have such-and-such fundamental phenomenal property. (The fundamental laws might not be about neurons, per se; they might relate qualia to some more general feature of the brain's activity — e.g. to changes in some kind of “pattern,”²¹ or in information-state.²²) Given adverbialism, whatever has this phenomenal property will be a conscious subject — one that feels a very precise pain, senses a very precise smell, etc. But what is the something-or-other that is caused to have the property in question? According to garden variety materialism, it is a familiar object such as a brain or a complete human organism.

If “brain” or “human organism” are terms for garden variety, vague material objects, and I am such a thing, then there must be many equally eligible candidates for being this brain or this organism. There is no problem, in principle, with vague macroscopic objects exemplifying fundamental, perfectly precise properties. All that is necessary is that each of the eligible candidates has the fundamental property. But, since the candidates differ from one another in tiny ways, and these tiny differences are fully determined by differences at more fundamental levels, it should be very surprising if it ever happens. It is easy for a vague object such as a table to weigh *about* 20 kilograms, because every eligible candidate for being the table has a mass very close to 20 kilograms — some a little more, some a little less. It is much harder for a table to weigh precisely 20 kilograms; some table candidates will, but very many will be ever so slightly heavier or lighter, rendering it less than completely accurate to say that the table has *exactly* that mass.

Adverbialism about some fundamental phenomenal properties requires that there be a family of perfectly natural properties which can be had only by conscious beings. If I am conscious in one of these precise ways, and I am an ordinary vague object, the laws governing the generation of qualia must insure that every eligible candidate for being me has this perfectly precise property. How likely is it that the fundamental laws select all and only the eligible candidates?

I suppose a property dualist should grant that it is *possible* that the natural process of qualia generation is *prodigal* in the production and distribution of fundamental phenomenal properties; that the brain generates very many instances of each phenomenal type, one for each of very many distinct but overlapping physical objects. But the defender of garden variety materialism must hope for more than that. The firing of neurons that causes something to have adverbial qualia must somehow target *all and only* the precise objects that are eligible candidates for being what we mean by “organism” or “brain.” The fundamental

physical laws governing qualia generation, even if they are prodigal in the number of instances produced, should not be expected to choose precise objects in exactly the same way that our everyday terms for brains and bodies choose many objects – that would be to attribute to nature itself a touching deference to our linguistic practices and to our rough-and-ready concepts.

If fundamental laws of adverbial-qualia-generation select fewer than every single one of the eligible candidates for being this organism or brain, the organism or brain will be at best *sort of* conscious. Whatever else I know about myself right now, I know that I am *definitely* conscious; so if a smaller thing or things *definitely* have the adverbial qualia, I am not the thing that is only indefinitely conscious, I am that smaller thing, or I am one of those things, or perhaps I am indefinitely identical with each of them – “I” might be a vague term, indeterminate in reference among many of the things that truly have the qualia generated by my brain. On the hypothesis that one or a few candidates are truly conscious, my boundaries are not those of an ordinary, garden variety, macroscopic object; they are determined not by our ordinary, rough-and-ready standards for being part of an organism or organ (which advert to vague notions like cohesion and functional role). Instead, my boundaries are set by a special, sharp “halo,” a boundary drawn by possession of the precise qualia. The property dualist should admit that this *might* be so. And the resulting view is materialism, fair enough, but it is a kind of *speculative* materialism, not the kind of materialism that finds a thinking person to be just another garden variety physical object of the sort we clothe, or remove surgically, or push around. The precise material object I am becomes a matter of theoretical speculation, determined by laws linking brain activity with a particular physical object or objects, presumably somewhere in the vicinity of my brain.

Suppose the laws select *more* than all of the eligible candidates – including, among the many objects that share my adverbial phenomenal states, some objects with parts that fall just outside all of the garden variety candidates for being this brain or body. In that case, there are larger objects of which this brain or body is a part, and it is just as true of them that they are conscious as it is true of me. Any object that largely overlaps an eligible candidate for being me, and that shares my conscious experience, ought to be a good candidate for being me, so, unless these slightly larger things differ drastically from the slightly smaller things, garden variety materialism is once again false – a halo surrounds a different collection of precise objects from the eligible candidates for being this brain or body. Although words like “brain” and

“body” are not vague terms used to refer to such a thing, we could easily invent some terms, indeed, given the platitude that *I* am the subject of my conscious states, perhaps “I” in my mouth is already a term indeterminate in reference among *these* material objects.

Even if the laws are prodigal, causing many physical objects to be conscious, there remains a kind of magic halo surrounding me (or, rather, around the sum of all the candidates for being me), so long as the perfectly natural qualia are either exemplified, or not, by an object — and this is something one does not find in ordinary, vague, macroscopic objects. The halo remains even in the wildly lucky case of laws that select *just* the eligible candidates for being this brain or body. In a garden variety object, there are not just bits of matter that are neither definitely part of, nor definitely not part of, the object — there are not just things one might call “borderline parts.” There is also no sharp cut-off between the bits of matter that are, and are not, borderline parts. Notoriously, this requires higher-order vagueness; but higher-order vagueness would be obliterated by precise facts about which physical objects have adverbial qualia.

Prodigal laws of qualia production are needed by the adverbialist property dualist, if conscious persons are to have a *chance* of being garden variety bodies or brains. But even then, the chance is slim; it is much more likely that I am a physical object of a different size and shape — one determined by the sizes and shapes of whatever things are caused to have my qualia. A host of overlapping conscious subjects could make “I” a vague term, so that it would be true to say that I am a vague object. Still, my vagueness would be unlike that of garden variety objects, and the resulting metaphysics of persons should count as a form of speculative materialism.

If phenomenal properties are genuinely new and genuinely fundamental, there is little reason — other than our affection for certain familiar, macroscopic, vague objects — to suppose they are produced in abundance and exemplified willy-nilly by a host of subjects that overlap the neurons that are their source. Non-prodigal laws force the materialist to adopt a speculative frame of mind. Brains generate adverbial qualia, which are not exemplified many times over, but rather by just one or perhaps a handful of physical objects. The thesis would not be so bad, were there a heretofore unnoticed kind of physical thing, distinct from the familiar examples of macroscopic objects, suitable to be the true bearer of qualia and the true subject of consciousness. But when one looks around for precisely demarcated physical entities to receive the adverbialist’s phenomenal states, no natural candidates present

themselves. No cell or molecule or atom in the brain is distinguished in a way that would suggest that it is a better candidate than any of its rivals for being conscious; there seems no precise physical entity in the vicinity that fundamental laws could pick out in virtue of some special physical status, either intrinsic (for example, a special type of particle, atom, or molecule) or extrinsic (for example, a special place in my brain where only one particle, atom, or molecule could be located). Of course, it might be that the precise subjects of phenomenal states are one or many tiny particles selected randomly — in virtue of indeterministic laws, say — from among those in my brain, or that the subjects are randomly chosen larger portions of the matter in my head. Perhaps different regions of my brain are chosen at different times, depending upon the location of the brain activity causing the experience. Speculative materialisms could take many forms; many different algorithms could be proposed to link neural activity with some specific material object or objects, or perhaps with some portion of a field, or even with points of space-time. But, given all the precise objects in the vicinity, there is a “pairing problem”: what natural feature (intrinsic or extrinsic) of a physical entity in the vicinity of my brain could figure in fundamental laws selecting one or several such entities to be the bearer of the newly generated phenomenal property? There is no obvious candidate, so far as I know. Perhaps the special part of the brain has yet to be discovered; therein lies a research program! Perhaps the laws about qualia generation choose physical objects to be experiencers in some indeterministic way, or the laws are strange ones, linking particular neural events with particular physical things but not in virtue of any natural relation that can be seen to hold between the neural events and those particular things.

In any case, property dualism and adverbialism lead the materialist into dark speculations about the true location and physical nature of persons. I do not say that such speculation would be unjustified, or scientifically unfruitful. But I do claim that those willing to engage in it are not in a position to scoff at the speculations of the emergent dualists.

EMERGENT DUALISM: BACK ON THE TABLE FOR THE ADVERBIALIST

Either the fundamental laws of adverbial qualia generation are prodigal, or not. If prodigal, there are many physical objects caused to have each of “my” experiences; in which case, I could believe that I am one of

them or, better, that “I” is ambiguous or indeterminate in its reference to many of them — but I should *not* suppose that the ambiguity lines up with the ambiguity or indeterminacy in our use of words for garden variety objects belonging to biological kinds. If the laws are less prodigal, and more choosy — so that neural activity causes only one instance, or only a few instances, of each quale — the materialist can hardly pretend to know the size and shape of a conscious person; as a matter of empirical fact, there are no promising candidates for being the unique, conscious physical object in the vicinity of a human brain or body.

The emergent dualist is bound to point out that another possibility remains: the possibility that, as in other circumstances in which a new fundamental property is exemplified, the phenomenal states come with a new subject. And of course this is exactly what the dualist believes to be the case. Once there is neural activity sufficient to generate consciousness, a subject for that consciousness is also generated. Given the perfect naturalness of the properties that are newly instantiated, one should suppose that any subject of such properties is itself as natural in kind as a fundamental particle.

The details of the mechanism by which brains generate souls remain, admittedly, as speculative as the search for a special conscious particle or a precisely demarcated conscious chunk of brain matter. Perhaps there is some minimal level of neural activity that could be identified as *the* sustaining cause of the soul. Perhaps, for every brain-and-soul pair at every time, there is a single pattern of neural firing that is responsible for the soul’s overall phenomenal state, then and there. I suppose that the following hypothesis is more likely: that many overlapping sets of events occur in the brain, none of which is the minimal cause of the soul’s ongoing existence, nor the single cause of its overall phenomenal state. With many overlapping patterns of neural firing, each lawfully sufficient for the existence of a soul with the same phenomenal states, there could still be just one soul, its existence and phenomenal state simply *overdetermined*. There need be no vagueness about which activities in the brain generate the subject of consciousness — in fact, on this supposition, *many* precise (and largely overlapping) events are equally responsible — nor about how many subjects there are.

Emergent dualism is clearly not the only coherent way to combine property dualism with adverbialism about the most natural phenomenal states. But, given the unlikelihood that the laws of qualia-generation choose just the macroscopic candidates that have captured *our* attention, garden variety materialism is extremely unlikely to be true. And more speculative forms of materialism become quite bizarre, so long

as no precise, physically special parts of the brain (or special fields or special sets of space-time points inside the brain) present themselves. The substance dualist alternative is to suppose that phenomenal states come with their own natural kind of subject, like new fundamental particles. Property dualists ought to accept this as a genuine possibility — a speculative hypothesis worth taking seriously, especially if there are no promising leads in the search for a physical alternative.²³

who provided feedback on the ideas in this paper. A special thanks to the co-authors and editors of this book who provided absolutely invaluable help with dialectical and expository issues. The graphics in this chapter were generously released by their creators into the public domain, and may be downloaded from www.wikimedia.org.

Notes to Chapter 7: From Experience to Experiencer

- 1 Franz Brentano, *Psychology from an Empirical Standpoint* (London and New York: Routledge, 1995), p. 159.
- 2 George Graham, “Self-Consciousness, Psychopathology, and Realism about the Self,” *Anthropology and Philosophy* 3 (1999): 533–9.
- 3 Given certain metaphysical views, they *might* affect such a shift, but not just any change in one’s self-conception is guaranteed to bring about a change in the thing to which one refers using first-person pronouns.
- 4 Thomas Nagel, *The View From Nowhere* (New York and Oxford: Oxford University Press, 1986), pp. 41–2. The emphasis is Nagel’s.
- 5 René Descartes, *The Philosophical Writings of Descartes*, vol. 2, trans. by John Cottingham, Robert Stoothoff, and Dugald Murdoch (Cambridge: Cambridge University Press, 1984), p. 13.
- 6 See Roderick M. Chisholm, “Which Physical Thing Am I? An Excerpt from ‘Is There a Mind-Body Problem?’,” in Peter van Inwagen and Dean Zimmerman (eds), *Metaphysics: The Big Questions*, 2nd edn. (Malden, MA: Wiley-Blackwell, 2008), pp. 328–33; and Philip L. Quinn, “Tiny Selves: Chisholm on the Simplicity of the Soul,” in Lewis Hahn (ed.), *The Philosophy of Roderick M. Chisholm*, (LaSalle, IL: Open Court, 1997), pp. 55–67.
- 7 I suppose one *could* say that the presence of a soul in our case acts as a sort of “consciousness magnet,” pulling mentality away from the physical substance — say, the brain — to which it would otherwise be “attracted.”
- 8 William Hasker, *The Emergent Self* (Ithaca, NY: Cornell University Press, 1999).
- 9 Richard Swinburne, *The Evolution of the Soul*, rev ed. (Oxford: Clarendon Press, 1986), chapter 10.
- 10 Hasker, *The Emergent Self*, Chapter 8; and Swinburne, *The Evolution of the Soul*, chapter 15.
- 11 I know of only one line of thought that would support such a bizarre view: the idea that persons are to their bodies as programs are to the computers that run the programs. Now there may be more and less plausible ways to understand this idea. But, on one way of construing the nature of a *program*, it is a set of rules, independent of the particular computers running them. Programs, so understood, are more like mathematical entities: abstract objects, existing outside space and time. But it is hard to take this picture seriously. If programs

are abstract things, like functions defined over numbers, then they are unchanging, and they exist no matter what the world is like. But surely *persons* are not unchanging things that exist no matter what the world is like.

- 12 I express my reservations about one traditional argument for dualism in “Two Cartesian Arguments for the Simplicity of the Soul,” *American Philosophical Quarterly* 28 (1991): 217–26.
- 13 Paul M. Churchland, *Matter and Consciousness* (Cambridge, MA: MIT Press, 1985), pp. 12–13.
- 14 An extreme form of act-object theory is presupposed in G. E. Moore’s “The Refutation of Idealism”, in *Mind* 12 (1903): 433–53. For defense of the act-object theory, see Frank Jackson’s *Perception* (Cambridge, MA: Cambridge University Press, 1977).
- 15 Classic defenses of adverbialism are found in C.J. Ducasse, *Nature, Mind, and Death* (La Salle, IL: Open Court, 1951), Chapter 13; and Roderick M. Chisholm, “The Theory of Appearing,” in Max Black (ed), *Philosophical Analysis* (Englewood Cliffs, NJ: Prentice-Hall, 1963), pp. 97–112.
- 16 Jackson, *Perception*, p. 59.
- 17 C. D. Broad, *Scientific Thought* (London: Routledge and Kegan Paul, 1923), pp. 245–55.
- 18 I say more in defense of these claims in “From Property Dualism to Substance Dualism,” *Proceedings of the Aristotelian Society*, supplementary volume 84: (page numbers missing). The kind of argument that I think leads the act-object property dualist to sense data can be found in Howard Robinson, *Perception* (London: Routledge, 1994), Chapter 6, pp. 119–50.
- 19 Here, I slightly repurpose a comparison due originally to Peter van Inwagen; see his *Material Beings* (Ithaca, NY: Cornell University Press, 1990), p. 238.
- 20 For important versions of such a theory, see Kit Fine, “Vagueness, Truth and Logic,” *Synthese* 30 (1975): 265–300; and Van McGee and Brian McLaughlin, “Distinctions without a Difference,” *Southern Journal of Philosophy* (Supplement) 33 (1994): 203–50. David Lewis endorses such an account in his *On the Plurality of Worlds* (Oxford: Basil Blackwell, 1986), p. 244.
- 21 See William Robinson, *Understanding Phenomenal Consciousness* (Cambridge: Cambridge University Press, 2004), pp. 207–26.
- 22 See David Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996),
- 23 For criticisms and questions, I am grateful to participants in conferences hosted by the Ursinus College, University of Nottingham, and the University of Geneva, and to the Metaphysics and Philosophy of Religion reading group at Rutgers. Special thanks are due to Mark Baker, John Hawthorne, Philipp Keller, Daniel Nolan, Ted Sider, Timothy Williamson, and Leopold Stubenberg. Also, in retrospect, I can see that I owe a great debt to Peter

Unger's book, *All the Power in the World* (New York: Oxford University Press, 2005).

Notes to Chapter 8: Souls Beastly and Human

- 1 For my account of this case I am relying on Aram Vartanian, "Trembley's Polyp, La Mettrie, and Eighteenth-Century French Materialism," in Philip P. Wiener and Aaron Noland (eds), *Roots of Scientific Thought: A Cultural Perspective* (New York: Basic Books, 1957), pp. 497–516. Page references in the text are to this essay.
- 2 "Along with its powers of locomotion, contraction and extension, eight or ten arm-like projections at its mouth-end could seize whatever prey came their way, which was then conveyed to the stomach and digested." *Ibid.*, p. 497.
- 3 *Ibid.*
- 4 *Ibid.*, p. 498.
- 5 *Ibid.*, pp. 504–5.
- 6 *Ibid.*, p. 503.
- 7 For more on the background for Trembley's discovery, see Aram Vartanian, *Diderot and Descartes: A Study of Scientific Naturalism in the Enlightenment* (Princeton: Princeton University Press, 1953), Chapter 4; and John P. Wright and Paul Potter (eds), *Psyche and Soma: Physicians and metaphysicians on the mind-body problem from Antiquity to Enlightenment* (Oxford: Clarendon, 2000), Chapters 10 and 11.
- 8 Vartanian, "Trembley's Polyp, La Mettrie, and Eighteenth-Century French Materialism," p. 508.
- 9 *Ibid.*
- 10 Our science today only partly agrees with this. Modern biology does see plants and animals as differing at a fundamental level; polyps are unambiguously animals, and there are no "animal-plants." But the common ancestry of all living things is affirmed, which readily lends itself to interpretation in terms of a single "scale of nature" if we are so inclined.
- 11 Vartanian makes a strong case for the importance of the polyp to La Mettrie and to later French materialism.
- 12 Jaegwon Kim, *Physicalism, Or Something Near Enough* (Princeton: Princeton University Press, 2005).
- 13 I should state that Kim would not necessarily agree with the reasons I am giving here; for his reasons his own writings should be consulted.
- 14 Nor is this limited to contemporary materialism; remember La Mettrie's reference to "certain physical causes . . . to which the whole chain of this vast universe is so necessarily bound and subjected that anything that occurs could not have not occurred."