

Smoke and Mirrors: The Illusion of Metaphysical Zombies

ADAM BRODIE

It is sometimes argued that the conceivability of zombies is reason to believe that phenomenal experience is irreducible to and ontologically independent of physical processes. Zombies, dualists argue, demonstrate that there is no necessary association between phenomenal states and the physical states on which they supervene. However, zombies may be much more problematic than dualists make them out to be. Consider this: are *you* a zombie? It seems like a ridiculous thing to ask, but in pursuing this question we encounter a paradox that forces us toward a unique perspective on the mind-body problem. Of course you're conscious, you think to yourself. You know that you're conscious; you're not a zombie. But how did you come to know this? How did you come to have access to this information?

Here, we find ourselves forced to address the epistemological difficulty that lies at the heart of the mind-body problem: how do we come to have knowledge of the mind, and how do we come to have knowledge of the body? My objective in this paper is to demonstrate that consideration of the mind-body problem from this epistemological perspective supports reductive materialism. I will argue that given the dependence of cognitive processes on the nervous system, we can access only information that is available to the physical systems that constitute our nervous infrastructure. Consequently, if we do in fact have access to information that we are conscious, consciousness must be reducible to physical processes. Finally, I will

Adam Brodie is a senior majoring in philosophy at Arizona State University. His primary interests are in metaphilosophy, philosophy of science, and philosophy of mind. After graduation, Adam intends to pursue an advanced degree in philosophy or cognitive science.

argue that philosophical consideration of the role of mirror neurons in the human brain provides a foundation for a materialist account of our intuitions concerning zombies, lending insight into the relationship between subject-based concepts and the material world.

§1. Phenomenal Consciousness and Metaphysical Zombies

You are phenomenally conscious. If someone pinches your arm, you feel pain. A complex sequence of physical events occurs, but there appears to be something more, a sort of subjective quality.¹ Seemingly, there's something it's like for you to experience pain, something that isn't captured in the description of the sequence of physical events that corresponds with your experience.

Imagine someone who is physically indistinguishable from you, exactly similar in every observable detail, down to the last fundamental particle. For convenience, let's call him Z. When we pinch Z's arm, a sequence of physical events occurs in his body starting in his arm, running up through his spinal cord and through his brain. Since Z is physically indistinguishable from you, the sequence of events that occurs in his body when his arm is pinched is indistinguishable from the sequence of events that occurs in your own body when your arm is pinched. Z doesn't experience pain like you do, however. That's because Z, although qualitatively identical to you physically, is not phenomenally conscious. Z is a zombie.

David Chalmers and his supporters argue that the mind is not reducible to the brain on the grounds that zombies like Z are conceivable. Consideration of zombies demonstrates that a person's physical properties alone do not necessitate any phenomenal properties. After all, Z is exactly like you in every physical respect, yet unlike you, he has no phenomenal experiences. Therefore, the property that distinguishes you from Z, consciousness, cannot be explained in physical terms.

Before discussing Chalmers' argument in greater detail, I'd like to distinguish between the sense in which we'll use the word "consciousness"

¹ The physical process is as follows: pressure applied in the pinch causes gated channels in the outer membranes of damage-detecting nerve cells called *nociceptors* to open. Ions begin to rush in and out of the nociceptors through these channels, collectively constituting the propagation of action potentials across the cells' membranes. When these action potentials reach the nociceptors' axon terminals, chemicals called neurotransmitters are released. The neurotransmitters bond to receptor sites on the surfaces of post-synaptic neurons, opening gated channels in the post-synaptic cells' outer membranes. This allows the action potentials to "jump" to the post-synaptic neurons, which in turn release their own neurotransmitters from their axon terminals onto the membranes of other neurons. This continues in a chain reaction of neural excitation ascending up through your arm, spinal cord, and into multiple parts of your brain. See Hardcastle, Torebjork, and Churchland.

when discussing phenomenal consciousness from another sense in which we may use it.

Phenomenal Consciousness: A person or system is phenomenally conscious if and only if there is something it is like to be that person or system.²

Functional Consciousness: A person or system is functionally conscious of X if and only if that person or system contains a representation of X (X may be an object, property, state of affairs, etc.).³ Functional consciousness is reducible to physical description.

Consider this: Z is presented with a yellow sunflower and asked what color it is. Z's brain is physically just like yours and consequently is capable of representing, distinguishing, and naming colors. Z replies, "Yellow." In this respect, Z is functionally conscious of the color of the sunflower. However, Z, being a zombie, is not phenomenally conscious of the color of the sunflower; there is nothing it is like for Z to observe the flower. For our present purposes concerning metaphysical zombies, it will be helpful to keep this distinction in mind, as the metaphysical possibility of zombies depends on it.

§2. The Hard Problem and the Argument from the Conceivability of Zombies

The *hard mind-body problem* (or simply the "hard problem") concerns the relationship between the mind, which is the subject of phenomenal experience, and the physical body on which it supervenes. In virtue of what natural principles do phenomenal experiences supervene on physical processes? The hard problem asks us to bridge the explanatory gap between the physical and the phenomenal. There seems to be something it is like to be in your current physical state, but why? Human beings seemingly could have gotten along without phenomenal consciousness; natural selection should have selected only for mechanisms of functional consciousness. Couldn't the world have lacked phenomenally conscious subjects altogether?

² A *system* is an arrangement of causally related components. For our present purposes, we'll want to keep open the possibility that physical systems are capable, on their own, of supporting phenomenal consciousness.

³ A *representation* is simply "something (an event or process) that stands in for and carries information about what it represents, enabling the system in which it occurs to use that information in directing. . . behavior" (Bechtel 334).

Chalmers argues that such a world is metaphysically possible. *Zombie world* is a possible world with a physical history identical to our world's physical history, but it completely lacks phenomenally conscious subjects. For every phenomenally conscious person in the history of our world, there is a corresponding zombie counterpart in the history of zombie world, each zombie living out his unconscious life in perfect parallel with his conscious counterpart's. For instance, right now in zombie world your own zombie counterpart is reading a paper identical to this one.⁴

Such a world certainly seems conceivable, and according to Chalmers, this leads us to reject materialism as a possible solution to the hard problem (142).⁵ Chalmers' argument in its most simple formulation is this:

Argument from the Conceivability of Zombies

- (1) Zombie world is conceivable
 - (2) If zombie world is conceivable, then zombie world is metaphysically possible.⁶
 - (3) If zombie world is metaphysically possible, then materialism is false.
- ∴ (4) Materialism is false.

Concluding that materialism is false, Chalmers insists that phenomenal consciousness emerges from the physical and is ontologically distinct from it.

Chalmers' argument is admittedly quite plausible at first blush. Phenomenal consciousness doesn't seem to do anything in the world that the laws of physics don't already account for; we can imagine every event and every process ever to have occurred in the world occurring no differently in the absence of phenomenal consciousness. This has led many philosophers to endorse epiphenomenalism or some other sort of mind-body dualism that treats the mental and physical as ontologically distinct categories. Before we reject materialism on the grounds of the conceivability of zombies, however, we'll want to critically revisit what we think we know about zombies.

⁴ It's currently reading a footnote exactly like this one, actually.

⁵ By *materialism* I simply refer to the view that the mental is not ontologically distinct from the physical, that all phenomenal experiences are reducible to physical processes, and that folk psychological concepts like beliefs, desires, and even minds are elements of a fundamentally flawed conceptual framework and should be abandoned in scientific and philosophical discourse.

⁶ For a discussion of how conceivability entails possibility see Chalmers, "The Two-Dimensional Argument against Materialism." For the purposes of consistency, I will maintain Chalmers' linguistic conventions concerning metaphysical modality throughout the course of the essay. However, these conventions do not weigh heavily on the central argument of the paper (§§3-4).

§3. The Argument from Ignorance

Chalmers' argument from the conceivability of zombies begins with the assumption that zombies are conceivable. Zombies, however, are defined in contrast to conscious subjects. Accordingly, the notion of zombies is dependent on our capacity to distinguish them from phenomenally conscious subjects. Since we have access only to our own phenomenal consciousness, our conception of zombies depends entirely on our knowledge that we are distinguishable from zombies. The concept of zombies would be altogether incoherent if it turns out that we cannot determine whether we ourselves are actually zombies. Imagine I ask you if you're a zombie. Your intuitive response might be something like, "No, of course not." Further, you would want to say that you're not a zombie because you *aren't* a zombie; after all, you know that you're conscious. But what's happening in zombie world as I'm asking you if you're a zombie? My zombie counterpart asks your zombie counterpart the very same question, "Are you a zombie?" Your zombie counterpart responds, "No, of course not." Remember, it is physically indistinguishable from you; it produces a response identical to your own as a consequence of the natural laws that govern physical processes both in our world and in zombie world. But your zombie counterpart can't report that it's not a zombie because it isn't a zombie; it *is* a zombie! If your response is really informed by your phenomenal consciousness, how do we account for your zombie counterpart's response?

To begin with, we need to reflect on what happens when someone asks you if you're a zombie. It's not enough that someone simply asks if you're a zombie in general. You can answer that question simply by referring to your commonsense belief that you, generally speaking, are phenomenally conscious. We're not interested in whether you're simply disposed to attribute consciousness to yourself, but whether you can confirm that you're genuinely conscious by appeal to your own privileged access to the contents of your conscious experience. We want to know what happens when someone asks you if you're a zombie at this very moment. Right now, at this precise instant, are you phenomenally conscious? This is novel information and in order to know this you must introspect, or at least engage in some cognitive process that results in *propositional knowledge* of your own consciousness. Nigel Thomas points out that "there must be *some* sort of cognitive process that takes me from the *fact* of my consciousness to my (true) *belief* that I am conscious" (173). If there isn't such a cognitive process, then your belief that you are conscious is spurious. The following argument builds on Jaegwon Kim's "supervenience argument" to establish that if phenomenal experience is irreducible, then such a process cannot exist and, in turn, that knowledge of an irreducible phenomenal consciousness is impossible (13-21).

Argument from Ignorance

- (1) The mind supervenes on the brain such that for any given mental event or process, there is an ontologically prior brain event or process.
 - (2) If you know that you're phenomenally conscious, there must be some cognitive process that takes you from the fact that you are phenomenally conscious to your belief that you are phenomenally conscious.
 - (3) Given (1), in order for there to be a cognitive process that takes you from the fact that you are phenomenally conscious to the belief that you are phenomenally conscious, phenomenal qualities must affect the brain.
 - (4) Irreducible phenomenal qualities don't affect the brain.
 - (5) Given (2), (3) and (4), if phenomenal qualities are irreducible, then you don't know that you're phenomenally conscious.
 - (6) If you don't know that you're phenomenally conscious, you cannot know that you're not a zombie.
- ∴ (7) If phenomenal qualities are irreducible, you cannot know that you are not a zombie; zombies are impossible.⁷

§4. Breaking Down the Argument

Let's take some time to go through this argument in greater detail, considering each premise individually.

(1) *The mind supervenes on the brain such that for any given mental event or process, there is an ontologically prior brain event or process.*

It has been apparent since antiquity that damage to the brain produces mental defects. In the second century CE the great Roman physician Galen demonstrated that motor control and sensation of the body depended on the integrity of the nervous infrastructure innervating it.

⁷ As noted in §2, I employ Chalmers' conventions concerning metaphysical modality: zombies are 1-conceivable but not 2-conceivable, and consequently 2-impossible—impossible in the same sense that water's being identical to anything other than H₂O is impossible (146). While water's being identical to anything other than H₂O is impossible in virtue of the reducibility of water to H₂O, the impossibility of zombies is the consequence of mind-body supervenience and the subsequent dependence of epistemic access on the functional properties of the body.

Galen speculated that other mental processes, such as cognition and emotion, were grounded in the activity of the brain—the “great king” of the nervous system. Western science has subsequently accumulated substantial evidence in support of Galen's intuition. For example, modern technology has shown that mental processes exhaustively correlate with particular processes in the brain. Most recently, the development of fMRI (functional Magnetic Resonance Imaging) has allowed neuropsychologists to observe in real time what regions of a patient's brain are active. This equipment allows neuropsychologists to converse with their patients or to have them perform specially designed tasks while simultaneously tracking the storm of electrical activity swirling around their brains.

While correlation does not in itself entail dependence one way or another, the mental's dependence on the physical is widely considered to be the most attractive explanation. The physical world exhibits astoundingly regular causal relationships. In turn, physical events can almost always be explained sufficiently in terms of prior physical events.⁸ Mental events, however, often cannot be sufficiently explained exclusively in terms of other mental events. Rather, mental events can be explained most consistently by reference to the physical events that accompany them. As a consequence, the correlation between the mind and the body is best explained by positing that the mind depends on the body.

(2) *If you know that you're phenomenally conscious, there must be some cognitive process that takes you from the fact that you are phenomenally conscious to your belief that you are phenomenally conscious.*

Something happens between the time you hear a question and the time you articulate your answer. It is no different when someone asks you if you're a zombie. There is some cognitive process that must take place before you produce an answer. If you are phenomenally conscious, then this process must take you from the fact that you are phenomenally conscious at this moment to your belief that you are phenomenally conscious.

(3) *Given (1), in order for there to be a cognitive process that takes you from the fact that you are phenomenally conscious to the belief that you are phenomenally conscious, phenomenal qualities must affect the brain.*

Here is the crux of the argument. It starts off with a very simple idea: the brain is a representational information-processing organ. Information

⁸ Since explanation requires appeal to preceding states, the first event to occur in the world constitutes an unexplainable event. This is not a problem; by definition the first event falls outside the domain of explainable phenomena. Other events may be unexplainable in this sense; on the level of quantum particles, prediction is probabilistic, and accordingly, events are only explainable probabilistically. Mental processes, however, correlate with phenomena on the cellular level. These phenomena are explainable according to biological, chemical, and physical laws. Accordingly, for our current purposes, we may consider the physical world to be causally closed.

about the external world is collected by a variety of sensory neurons within your body. This information is represented in the states of your nervous system and in this way is made available to higher-level processing mechanisms. Consider touching a piece of paper. Mechanoreceptors in your fingertips become activated to various degrees according to how much pressure is exerted on them by the paper. This activation pattern carries information about the texture of the paper. This information, preserved as signals from your fingers' mechanoreceptors, is relayed to your brain, where the texture of the paper is represented as an activation pattern in your somatosensory cortex. This representation interacts with other parts of your brain, allowing you to move your fingers so as to turn the page without tearing it and to talk meaningfully about what it is like to feel like paper. Information about the internal states of your body, including the states of other parts of your nervous system, is represented in this manner as well. Take, for instance, the neurons in your hypothalamus that serve to represent low blood sugar. When enough of these neurons become activated simultaneously, you become hungry.

Our nervous systems are capable of representing a wide range of information, allowing us to behave in an organized, purposeful way. We manage to walk around, navigating our environment without tumbling over obstacles, and we reliably seek food before our bodies starve to death. But in order for any type of information to inform our behavior, it must be compatible with our nervous systems; it must be encodable as a pattern of neural activation. This encoding of information in activation patterns is a physical effect on our nervous systems.

Since we talk about the phenomenal qualities of experiences, we must have access to information characterizing those experiences. In order to have access to that information, these phenomenal qualities must in some way affect our nervous systems such that they may be represented within our brains. It is only when representations of phenomenal qualities are developed that our nervous systems can respond and generate the behavior to report on them.⁹

Some epiphenomenalists argue that it isn't necessary that phenomenal qualities affect the brain in order for us to have representations of them. They argue that the physical states on which those phenomenal

⁹ Murat Aydede and Güven Güzeldere give a much more rigorous information-theoretic analysis of phenomenal concepts in this sense: "If we want to talk about sensory concepts carrying information about experiences, we have to treat experiences as information-generating sources *on their own*—even when much of the information thus generated at the sensory level nomologically depends on the elimination of possibilities at a source beyond them, i.e., in the world. . . [e.g.,] when we token RED in response to a ripe tomato, our [phenomenal] concept [carries information not about the tomato, but rather] about which neurophysiological property is instantiated in the relevant part of our visual cortex [in response to viewing the tomato]" (223).

qualities supervene could be doing the causal work in informing the nervous system of the phenomenal state. There's a critical problem with this argument, however. Imagine driving to work. You see a stop sign at an upcoming intersection. You observe the stop sign; it's octagonal and red-colored. You know that the stop sign is red in virtue of the color sensitive photoreceptors in your eyes. They become selectively activated in a pattern that represents the color of the perceived object. But then you take a moment to philosophize—not only do you know that there's a red stop sign out in the world in front of you, but you also know something about the phenomenal experience of seeing the stop sign. It's not just that it's red, there's also something it's like to see the redness of the stop sign, and you know what that's like. It should be clear that we are considering two distinct phenomena. Not only are they conceptually distinct, but they are temporally distinct as well. We must consciously attend to our phenomenal experience in order to perceive it as such. At first, we simply see a red stop sign, but then we've turned our attention to the phenomenal experience of redness. Somehow, the phenomenal experience of redness has influenced the course of the cognitive processes occurring in your brain.

If we think of our brains as representational systems, then this would be analogous to the distinction between representing something out in the world and representing a system representing something out in the world. There is a higher degree of representation at work when we consider our phenomenal experience as such, and according to mind-body supervenience, there is a physical substrate of that higher degree of representation. If the epiphenomenalist is right, then even if we are conscious, we only say we're conscious for the same reason the zombie does—because of the way the brain represents itself. We say things like "I know I'm conscious," because we know that we're conscious, not just because of what our brains are doing. This requires that phenomenal consciousness affect the brain.

(4) *Irreducible phenomenal qualities don't affect the brain.*

Physical events almost always occur in a regular way. We can sufficiently explain almost every physical event in terms of preceding physical events; the physical world is causally closed. Brain states can be sufficiently explained in terms of preceding brain, nervous system, and environmental states. Consequently, there's no room for phenomenal qualities to influence physical states of the brain.

Some may argue that phenomenal qualities affect the brain in the sense that some physical states are caused both by prior phenomenal states and by prior physical states, thereby endorsing overdetermination. Even if this were the case, it wouldn't undermine the argument from ignorance. This is so for the same reason that rendered the epiphenomenalist's

criticism impotent: in order for us to develop distinct ideas of the physical events occurring in our brains and the phenomenal events occurring in our minds, it is necessary that the two cause effects on the nervous system that are discernible from one another. We would need to develop a neural representation of the physical event and a distinct neural representation of what it is like to experience that physical state. If phenomenal qualities don't determine any brain states independently, then it is impossible for experience to modify the brain in such a way as to produce distinct concepts of phenomenal experiences and physical processes. Accordingly, property dualist accounts that posit either epiphenomenalism or overdetermination to describe the causal role of the mind are vulnerable to the argument from ignorance.

(5) Given (2), (3) and (4), if phenomenal qualities are irreducible, then you don't know that you're phenomenally conscious.

Just a bit of logic: Given (4), the consequent of (3) is false. By modus tollens, the antecedent of (3) is false. This means that the consequent of (2) is false, and given this, by modus tollens, the antecedent of (2), "You know that you're phenomenally conscious," is false.

(6) If you don't know that you're phenomenally conscious, you cannot know that you're not a zombie.

As we mentioned earlier, zombies are defined by a lack of phenomenal consciousness. Since we want to continue to say that we know that we are conscious, consciousness must be reducible. So assuming mind-body supervenience, it must be the case that phenomenal experience is either reducible or essentially unknowable. Either way, we generate the conclusion that zombies are impossible.

§5. The Problem with Mind-Body Dualism

The argument we've been discussing is grounded in the idea that we are restricted in what we may talk effectively about by what sort of information is available to our nervous systems. Before we conclude our discussion of how this leads us to reductive materialism, I would like to take a moment to discuss an alternative view.

It is still possible that consciousness is, in fact, irreducible. Zombies may be a genuine possibility. It is my opinion, however, that this is simply an unattractive and cumbersome position. In the preceding sections I argued that we should accept reductive materialism because, given mind-body supervenience, we are restricted in the information we have access to by the physical properties and sensitivities of our nervous systems.

It is possible though, despite the apparent correlation between our phenomenal experiences and the observed phenomena in our brains, that the activity within our minds is not exhaustively underpinned by the physical processes occurring in our brains. That is, it is possible that the mind does not entirely supervene on the body. It may be that when we consider the phenomenal qualities of our experiences, we, as mental subjects, are doing something independently of our brains. Perhaps when we reflect on what it is like to see the color red, or to feel pain, we are acting beyond the limitations of our brains, which are embedded in a material world.

This is certainly possible. However, as I mentioned earlier, it is my opinion that this is a poor explanation for our observations. If the phenomenal qualities that we believe characterize our experiences were irreducible, then we would be forced to conclude that information is processed outside of the causally-closed material world and then somehow informs our behavior in the material world. Consider the question that we started with: "Are you a zombie?" You may adhere to an account on which zombies are a coherent concept and respond, "No, I'm sorry, I'm not a zombie." It is possible that the question was carried into your ears as disturbances in the air, transformed into a neural activation pattern, processed and interpreted in your brain, sent out to your immaterial mind, which then processed the question in its own right and sent a message back to your brain, which was then somehow transformed into the motor commands that coordinated your verbal response, "No, I'm sorry, I'm not a zombie." However, it simply seems implausible that this could actually be the case.

This sort of reasoning is often used to criticize the position of substance dualism. In considering our "Are you a zombie?" inquiries, however, we find that a strong case can be made that so-called "property dualist" theories also implicitly require some sort of mind-body interaction loophole in an otherwise regular physical world. If we affirm that we are conscious and maintain that consciousness is an ontologically independent property, we must conclude that our response behavior is influenced by this independent property, consequently endorsing the same sort of interactionism that the substance dualist supports. It is possible, though, that the mind-body problem is really as mysterious and intractable as some of these dualist accounts claim, but I find the materialist perspective to be much more plausible. But what exactly does this materialist account look like? And how in a material world can subjects of phenomenal experiences exist?

§6. The Zombie in the Mirror

Before we can endorse materialism, it is paramount that we address the problem of explaining our habits and intuitions concerning phenomenal consciousness and metaphysical zombies. After all, it certainly still seems like we can conceive of zombies. And more importantly, we frequently talk about what it's like to be this or that person, or to do something or other. Furthermore, it's often a lot more than just talk; it really seems like we know, in some sense, what others' phenomenal experiences are like. But if knowledge of irreducible phenomenal consciousness is conceptually impossible, why do we talk and think the way we do?

Imagine you see a friend stub his toe, holler in pain, and hop around on his good foot for a few seconds. It's likely that you would feel perfectly comfortable claiming that you know what his experience is like; you know how he feels, at least in some sense. But why do we want to say these sorts of things? What do we mean when we talk about what experiences are like?

An explanation for our linguistic habits and conventional ways of thinking might emerge from philosophical interpretation of recently developed theories of mirror neurons. Mirror neurons are neurons in the brain that become activated both when some action is performed and when another agent is observed performing that same action.¹⁰ For instance, there is a collection of mirror neurons in your brain that become activated both when you wave your right hand as if to say hello and when you see someone else wave his right hand in the same way. With mirror neurons, multimodal, subject-dependent representations can be developed, and in turn we can appreciate the behavior of others by reference to ourselves. It is easy to imagine how such mechanisms could be useful for social, goal-representing animals like human beings.

But that's only half of the story. We also have systems that allow us to mirror others' emotional states. This has profound implications on how we perceive events involving other people. What happens when you see someone step on a piece of broken glass? If you're any bit as squeamish as I am, you find yourself wincing and curling your toes. Neurons in the visceral motor system in your brain act in the same way that they would have if you had stepped on the glass. You don't feel the pain, exactly, but you do

¹⁰ See Rizzolatti and Sinigaglia, *Mirrors in the Brain: How Our Minds Share Actions and Emotions*; Gallese, "Embodied Simulation: From Neurons to Phenomenal Experience"; Gallese, Keysers, and Rizzolatti, "A Unifying View of the Basis of Social Cognition"; and Gallese and Metzinger, "The Emergence of a Shared Action Ontology: Building Blocks for a Theory."

experience the emotional response that accompanies such an injury.¹¹ In turn, you exhibit behaviors characteristic of that emotional response.

What I am suggesting is that mirror neurons provide us with the capacity to interpret events involving other persons using the same neural machinery that informs us of the states of our own bodies. In turn, we come to think of the world as being inhabited by mental subjects who are relatable to ourselves. Our notions of ourselves as mental subjects are then reinforced when we compare our own experiences with the experiences we attribute to others. As a matter of fact, it seems likely that we use our mirror and visceral motor systems in a somewhat autobiographical role as well; for instance, we may use these "mirror mechanisms" when we reflect on our own past experiences. This is supported by the fact that we remember our own injuries in a fashion strikingly similar to the way we perceive injury to others; we simulate the emotional aspect of pain but are incapable of simulating the sensory element. So we talk about and even know, in some sense, what our own experiences and the experiences of others are like. But note that this is not knowledge of irreducible phenomenal qualities. Rather, it remains embedded in the physical neural network; it's the sort of knowledge that zombies can have.

We can explain our idea of zombies in terms of mirror systems as well. Normally when we think of persons we think of them in terms of ourselves using a rich set of neural resources. That is to say that generally, when we conceive of other people, we use the same sensory and motor representations that we use when representing our own bodily states and planning our own movements. When we conceive of zombies, however, we are merely thinking of what a human being looks and sounds like; we don't attribute to them the mirrored motor and sensory states that we would attribute to others whom we take to be conscious. A zombie, then, is simply a person whom we've turned our mirrors away from.

If mirror neurons do perform this function, the way we think about others' phenomenal experiences is a product of our brains' capacities to represent our own bodies and the bodies of others. This theory is supported by our experience with people with autism. Autism is speculated to be the consequence of mirror neuron dysfunction.¹² We should expect, then, that autistic persons would have radically different ideas of what it means to

¹¹ Pain actually involves two parallel processes: a "sensory discriminative" process and an emotional, or "affective-motivational," process. Though we mirror the affective-motivational processing of others when we observe them sustaining an injury, we do not mirror their sensory discriminative processing. Consequently, and perhaps fortunately, we do not experience pain, per se, when witnessing injury to others. For more on how pain is processed, see Hardcastle. For more on how this compares to the processing of the pain of others see Frith, pp. 149-51.

¹² See Hadjikhani.

be conscious or to have a mind. Not surprisingly, autistic persons tend to treat other people as if they were unconscious, often finding it difficult to understand that other people have mental lives.

§7. Subjects in a Material World

There's no natural law that says that our brains must be designed to make it easy to understand how the world is. To say that human beings are social animals is an understatement. The evolutionary development of our brains was driven not only by the pressure to accurately represent the world, but also by the advantages of representing a rich, complex social world. Consequently, our brains aren't rigged up to make doing metaphysics easy. Think about it—what exactly is your mind? Is it identical to your body? Maybe it's identical to your brain. Maybe it's just some of your brain. We can certainly talk this way if we want to, and at times it may feel natural to use the word “I” to refer to any of these things. At the same time, brains are not, strictly speaking, how we think of minds. To us, minds are unified and persistent and preserve numerical identity over time. Our brains, on the other hand, are dynamic, constantly changing and turning material over within complex, distributed electrochemical networks. Further, there's nothing that requires our minds to be in the same location as our brains, either spatially or temporally.¹³ The closer you look at your brain, the harder it becomes to relate to it. So what does a mind reduce to?

As I've been suggesting, our brains perform certain functions that are social in character. While much of the language we use when describing the world is subject-independent, we also have socially driven language that is subject-dependent, or subjective. This includes our language for discussing concepts such as persons and feelings, the sorts of things that are grounded in the assumption that the world is inhabited and experienced by subjects. We are, of course, socially driven animals, and accordingly, these sorts of concepts are very important to us. It doesn't follow, however, that we should seek to map these subjective concepts onto what we consider our best available conceptual framework for explaining events in the world. But perhaps we don't need to reduce the mind in a conventional sense of reduction in order to be materialists. While subjective concepts may not be directly reducible into the framework of physics, subjective concepts and the concepts of physics purport to explain phenomena in a common domain. Consequently, any explanation that employs subjective concepts can be replaced with an explanation utilizing only the conceptual framework of physics. As it turns out, the concepts of physics allow us to make much

¹³ See Rick Grush's thought experiment.

more precise and more accurate predictions than our folk subjective concepts do. The mental, then, can be *reconstructed* in terms of the physical. It is in this sense that the mind is reducible.

So we can have our cake and eat it too. We can continue to enjoy our world, full of the minds of friends, family, and others. But we can be materialists at the same time, appreciating how we come to see subjects all around us in a material world.¹⁴ It should be noted that this capacity to model the bodies of others and in turn recognize other mental subjects is not equivalent to consciousness, at least not in a certain sense. However, it may be helpful to break the problem of consciousness into two distinct sub-problems: first, how is it that a “world comes to appear to us,” and second, how is it that we come to recognize ourselves in that world as instances of a general class of “conscious subjects”?¹⁵ As long as we are concerned with the phenomenon of conceiving of metaphysical zombies, it is the latter problem that we must address. As we continue to learn how our brains produce the idea of minds, we can begin to determine what other physical systems might share this capacity. We may find that it's often this more specific capacity that we're really interested in when we ask whether some animal or machine is conscious. This is especially the case when we are concerned with the mental lives of other highly social animals, animals that we are often compelled to relate to, such as chimpanzees, dogs, cetaceans, and other higher mammals. And while these questions are yet to be rigorously investigated within comparative psychology, given the behaviors exhibited by these animals, it may be fair to conjecture that they too share the sort of social “mirrors” that enable the discrimination of mental subjects.

§9. Concluding Remarks

It may appear that the view advanced in this paper is largely motivated by developments in neuroscience. Many philosophers of mind argue that neuroscience cannot inform our investigation of the mind-body problem. They claim that neuroscience purports only to explain the operation of the

¹⁴ It may be helpful to compare our capacity to “see” mental subjects in the world to our capacity to see colors: while we see the world in color, we now understand that the world is composed of essentially colorless constituents, and although this is still contentious, it seems unlikely that we should find anything satisfying or insightful in carrying out the reduction of colors to collections of light reflectance properties (in other words, the set of all conceivable objects that produce in you the sensation of some particular color have nothing in common other than their capacity to reflect light in such a way, given certain conditions, as to produce in you such a sensation).

¹⁵ This phrasing comes from German philosopher of mind Thomas Metzinger, who has produced a particularly eloquent analysis of the problem of consciousness informed by contemporary neuroscience; see Metzinger, *Being No One: The Self-Model Theory of Subjectivity* and *The Ego Tunnel: The Science of the Mind and the Myth of the Self*.

brain and nervous system; it has no role in determining the relationship between those physical systems and the phenomenal experiences that seem to accompany such systems. In some sense, I agree with this sentiment. I do not think that an appreciation of neuroscience is necessary to investigate the philosophy of mind. However, the view advanced in this paper does not rest so much on contemporary neuroscience as on a thorough commitment to mind-body supervenience. While an appreciation of neuroscience is by no means necessary for an understanding of the position in this paper, it may be helpful in the sense that it becomes easier to accept materialism once one begins to see how the brain might have the resources to comprehensively underpin the activity of the mind.

The more we understand about how our brains work, the better we will be able to explain the way we perceive the world, and in turn the better equipped we will be to understand the world itself. As we continue to make progress in our efforts to understand the world, we will inevitably come to interpret our own experiences in new ways. Perhaps we are already prepared to reinterpret our experiences imagining zombies. Perhaps they're not really possible, and what once seemed unexplainable was really nothing more than smoke and mirrors.

Works Cited

- Aydede, Murat and Güven Güzeldere. "Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness." *Nous* 39.2 (2005): 197-255.
- Bechtel, William. "Representations: From Neural Systems to Cognitive Systems." *Philosophy and the Neurosciences*. Ed. William Bechtel, et al. Oxford: Blackwell, 2001. 332-48.
- Chalmers, David. "The Two-Dimensional Argument against Materialism." *The Character of Consciousness*. Ed. David Chalmers. New York: Oxford UP, 2010. 141-91.
- Churchland, Patricia. *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press, 1986.
- Frith, Christopher. *Making Up the Mind: How the Brain Creates Our Mental World*. Malden, MA: Blackwell, 2007.
- Gallese, Vittorio. "Embodied Simulation: From Neurons to Phenomenal Experience." *Phenomenology and the Cognitive Sciences* 4 (2005): 23-48.
- Gallese, Vittorio, Christian Keysers, and Giacomo Rizzolatti. "A Unifying View of the Basis of Social Cognition." *TRENDS in Cognitive Sciences* 8.9 (2004): 396-403.
- Gallese, Vittorio and Thomas Metzinger. "The Emergence of a Shared Action Ontology: Building Blocks for a Theory." *Consciousness and Cognition* 12 (2003): 549-71.
- Grush, Rick. "Locating Subjects of Experience in the Natural Order." Research Lecture. University of California, San Diego. 6 Apr. 2006.
- Hadjikhani, Nouchine. "Mirror Neuron System and Autism." *Progress in Autism Research*. Ed. Paul Carlisle. New York: Nova Science Publishers, 2007. 151-66.
- Hardcastle, Valerie. "The Nature of Pain." *Philosophy and the Neurosciences: A Reader*. Ed. William Bechtel, et al. Oxford: Blackwell, 2001. 295-311.
- Kim, Jaegwon. *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton UP, 2005.
- Metzinger, Thomas. *Being No One: The Self-Model Theory of Subjectivity*. Cambridge, MA: MIT Press, 2004.
- . *The Ego Tunnel: The Science of the Mind and the Myth of the Self*. New York: Basic Books, 2009.
- Rizzolatti, Giacomo and Corrado Sinigaglia. *Mirrors in the Brain: How Our Minds Share Actions and Emotions*. Trans. Frances Anderson. New York: Oxford UP, 2008.

Works Cited

- Rumelhart, David, et al. "Schemata and Sequential Thought Processes in PDP Models." *Parallel Distributed Processing, Vol. 2: Psychological and Biological Models*. Ed. Jerome Feldman, Patrick Hayes, and David Rumelhart. Cambridge, MA: MIT Press, 1986. 7-57.
- Thomas, Nigel. "Zombie Killer." *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*. Ed. Stewart Hameroff, Alfred Kaszniak, and Alwyn Scott. Cambridge, MA: MIT Press, 1998. 171-77.
- Torebjork, Erik. "Nociceptor Activation and Pain." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 308.1136 (1985): 227-334.

Works Consulted

- Bainbridge, David. *Beyond the Zonules of Zinn: A Fantastic Journey Through Your Brain*. Cambridge, MA: Harvard UP, 2008.
- Chalmers, David. "Consciousness and Its Place in Nature." *Philosophy of Mind: Classical and Contemporary Readings*. Ed. David Chalmers. New York: Oxford UP, 2002. 247-72.
- Gazzaniga, Michael. "Why Can't I Control My Brain? Aspects of Conscious Experience." *Cognition, Computation, and Consciousness*. Ed. Masao Ito, Yasushi Miyashita, and Edmund Rolls. New York: Oxford UP, 1997. 69-79.
- Hinton, G. E., J. L. McClelland, and David Rumelhart. "Distributed Representations." *Parallel Distributed Processing, Vol. 1: Foundations*. Ed. Jerome Feldman, Patrick Hayes, and David Rumelhart. Cambridge, MA: MIT Press, 1986. 77-109.
- Squire, Larry. *Memory and Brain*. New York: Oxford UP, 1987.