

# A Gap in Kim's Eliminative Argument for Reductionism

B.D. Mooneyham

**J**aegwon Kim's book, *Physicalism, or Something Near Enough* (2005), addresses important issues for those who wish to understand the mind in the context of a physicalist worldview. His book starts by introducing two difficult concepts for physicalists to explain about the mind, given certain beliefs that go with the physicalist worldview. The two concepts together create an unfortunate dilemma for physicalists, because an explanation of one seems incompatible with an explanation of the other. Having established this dilemma, Kim then proceeds to examine various theories of the mind, evaluating them on the basis of their ability to respond to these difficult concepts. Using this standard, Kim slowly whittles down the options until one remains. In this way, he ends up with a view that, he admits, contains elements that do not appeal to a large number of philosophers, though it is nonetheless the truth that we are stuck with.

In this paper I will show that Kim's argument against a major anti-physicalist view of the mind, using his version of the "pairing problem," does not function in the same way as his argument against nonreductive physicalism, also known as the exclusion argument. Viewed as a stand-alone argument, the pairing problem may not bear the same eliminative weight as the ex-

---

B. D. Mooneyham completed a dual degree at Kansas State University in December of 2008, majoring in Philosophy and Economics. In the fall of 2009, he will attend the University of Chicago Law School, where he will study regulation and jurisprudence. His philosophical interests include philosophy of mind, ethics (especially Kantian ethics), epistemology, philosophy of law, and the history of philosophy.

---

---

clusion argument, because it does not validly lead us to the same conclusion. To set up this discussion, I will introduce the major contemporary theories in the philosophy of mind, then summarize Kim's presentation of the exclusion argument and the pairing problem, and discuss how they differ and why it is important.

The philosophy of mind generally concerns issues and problems pertaining to mental life, including things such as our thoughts, beliefs, desires, perceptions, and sensations: we want to know where these things come from, what they are like, and how we can explain them. To give some brief background, the historical debate about these topics can be divided into roughly two camps: dualists, who hold that the mind is distinct from the body and is the seat of mental life; and non-dualists, who hold that the body and the mind are one and the same, or rather that mental functions are a special kind of bodily function. According to the latter group, "having a mind" does not literally mean there is some special thing you "have" any more than "taking a walk" means that there is some *thing* in the world known as a "walk" that you can take.<sup>1</sup>

On one end of the spectrum between dualists and non-dualists is the view known as substance dualism. This view is commonly associated with René Descartes, although its origins stretch back long before him and it continues to be held among some philosophers today. Substance dualism is the view that human beings are composed of a material body as well as an immaterial "soul" (or mind; I will use these terms interchangeably). On this view, the soul is a completely separate substance from the body, existing independently (that is, not as a derivative of the body) but also somehow very closely connected to it. Substance dualism, one can see, is compatible with such beliefs as life after bodily death, the existence of incorporeal spirits, and so forth. This is not to imply that it is less a philosophical view than a religious view, but is meant to set up the entrenched connection between philosophy of mind and a broader ontology of existent objects.

One will notice that substance dualism is much less com-

patible with worldviews that preclude the existence of immaterial, or spiritual, entities. In the contemporary philosophical realm, the primary criticisms of substance dualism come from those who subscribe to some form of physicalism. Physicalism is roughly the view that everything that exists in the world is made up only of bits of matter and nothing else. Another way of describing physicalism is that everything in the world can be explained exhaustively by using the language of the basic sciences such as chemistry and physics—much of the driving force behind physicalism comes from continued advancements and new discoveries in the sciences, which seem to explain more about the world at every turn. Over the last century, physicalism has become the dominant framework through which most contemporary philosophy of mind is discussed.

Jaegwon Kim writes his book for physicalists, trying to find the position that is most consistent with the principles of physicalism while also being able to explain adequately aspects that we consider necessary to an understanding of the mind. There is serious division even among physicalists about how to characterize the mind. Some people believe that physicalism is compatible with a certain kind of dualism, usually referred to as property dualism. According to property dualism, mental events and mental properties (things such as thoughts, beliefs, and sensations) are categorically distinct from physical events and properties (say, neural or brain states). In other words, one cannot reduce the mental to the physical—even though every substance that exists is purely physical, mental properties also exist. This general position is known as nonreductive physicalism, with property dualism being the most common version. By contrast, reductive physicalists (or reductionists) believe that mental activity *just is* physical activity, nothing more.

Kim aims his initial address at nonreductive physicalists—those who believe that the mental is distinct from the physical—due to the amount of attention and popularity currently enjoyed by these views in the philosophical realm. The two primary issues (*i.e.*, problems for physicalism) that Kim brings up in chapter one of his book are mental causation (how

---

---

one's mind can cause things to happen in one's body and vice versa, or how one mental event can cause another) and consciousness. Kim then presents a simple argument (*viz.*, the exclusion argument) meant to show that under a set of basic assumptions held by physicalists, it becomes difficult to show how mental events can have any causal powers.

These assumptions are:

1. Mind-Body Supervenience, according to which no occurrence or change of mental properties or events can exist without a corresponding physical property or event—to say that the mental “supervenies” on the physical is not merely to say that they are in constant conjunction; it means that the mental *depends on* the physical in some way;
2. The Causal Closure of the Physical Universe, which states that for any physical event that has a cause, that cause will be physical—this is a basic principle used especially in scientific explanations under physicalism; and
3. A Causal Exclusion Principle, which entails that there can be no more than one sufficient cause for a given event, unless that event is a “genuine case of causal overdetermination.”

The last of the three assumptions Kim believes to be a “general metaphysical constraint”<sup>2</sup> and not simply an assumption of physicalism. These three beliefs, coupled with the nonreductive view of mental events, give rise to the problem of mental causation that Kim tries to establish.

Roughly translated, the problem amounts to this: the physicalist must believe (according to supervenience) that any time a mental event, such as a thought, occurs, there exists a corresponding physical event which necessarily accompanies the mental event—say, some brain state S. That mental event could not exist without its physical “supervenience base.” The physical supervenience base generates, or causes, the supervenient mental event. If this is the case, there seems to be little room to

---

---

say that one mental event can “cause” another, because for any given mental event, it will have a physical supervenience base that is sufficient for its existence. This makes mental-to-mental causation an illusion.

We might be tempted to say that a mental event can “cause” another by causing the physical supervenience base for that mental event—this would be an example of mental-to-physical causation. But again, according to supervenience, the former mental event must have its own physical supervenience base, without which the former mental event could not exist. Because of this dependency of the mental on the physical, this former physical event (which generated the former mental event) now has a strong case to be considered the true cause of the latter physical event (which generated the latter mental event). Moreover, according to physical causal closure, the latter physical event must have a physical cause, and according to the causal exclusion principle, there must only be one cause (since this does not seem to be a genuine case of causal overdetermination). So we have a physical event (with a supervenient mental property) which causes another physical event (generating its own supervenient mental property), but the mental events do not stand in a causal relationship to one another.

This argument takes away mental-to-mental causation and mental-to-physical causation, leaving, at best, physical-to-mental causation. Thus we are faced with a situation in which mental events appear to be stripped of causal powers, leading to epiphenomenalism—roughly, the view that mental events merely appear to be causally efficacious when in fact they are causally impotent. This unfortunate result could be overcome on a reductionist view, because in that case, the mental event *just is* the physical event, so positing mental events as causal influences would not result in causal overdetermination. However, as Kim points out, many philosophers resist this concession to reductionism. For one thing, reductionism seems to compromise our sense of the mental as something distinct and important in our understanding of ourselves.

This leads into the next major problem for physicalists:

---

---

the problem of consciousness. Philosophers of mind are challenged to explain how such a phenomenon as consciousness could arise in a universe that strictly consists of bits of matter. Kim argues that this mystery can only be solved if consciousness is reducible,<sup>3</sup> but for a variety of reasons, consciousness seems to evade reductionist strategies. Part of this has to do with the phenomenal quality of certain experiences, such as the sensation of redness or the "hurt" of pain, for which it is hard to find physical correlates.

Of course, if consciousness cannot be reduced, this constrains our ability to use reductionism to preserve mental causation. The dilemma should now be clear: physicalists seem forced to either accept a genuine view of mental causation by adopting reductionism and thereby risk being unable to account for consciousness, or reject reductionism and sacrifice mental causation. This is a hard place for physicalists, because both mental causation and consciousness seem to be key components in our understanding of our own mentality. Kim endorses reductionism as the answer to this dilemma because he thinks that the idea of mental causation is too important to leave out of a theory of the mind. He says, "[T]he possibility of human agency, and hence our moral practice, evidently requires that our mental states have causal effects in the physical world."<sup>4</sup> He then goes on to identify different forms of reductionism and present arguments for the best among these options.

But before he moves on to reductionism, he acknowledges that there seems to be an important third choice in this matter. Given the uncomfortable dilemma that has been presented to physicalists, some may take this to be evidence that physicalism is an inadequate worldview under which to construct a theory of the mind, and that one should instead embrace the major alternative—substance dualism. Recognizing this inclination, Kim offers an argument meant to show that substance dualism fares no better (and perhaps worse) than nonreductive physicalism on the issue of mental causation.

According to Kim, the classic criticism of Cartesian substance dualism has come in the form of a protest that it is hard to

see how two very different substances, one immaterial and one material, could causally interact with one another. As Kim points out, this objection really amounts to more of a demand for an explanation than an actual demonstration that substance dualism is false. Kim thinks the physicalist can find a more persuasive argument against the dualist. The central problem of immaterial souls concerning mental causation, Kim believes, lies not in the differences between mental and physical substances but in the inherent nonspatiality most commonly attributed to immaterial souls under substance dualism. The objection can be best illustrated by what Kim calls the pairing problem, which employs a scientific explanation of causal interactions. It is meant to undermine substance dualism's ability to account for mental causation.

The pairing problem can be summarized with the following observations. When we talk about causal relations between physical objects, the way that we pair a specific cause with its particular effect is to identify spatiotemporal properties of the two objects—physical coordinates, points of contact, *etc.* When one billiard ball hits another, or a flame heats a pot of water, there is a physical interaction. By identifying these spatiotemporal relations, we can distinguish the exact physical cause from other objects in the vicinity, because two distinct objects cannot exist in the same space at a given time. The same cannot be said, supposedly, about immaterial souls. If souls really are nonphysical, it does not seem to make sense to point to a location and say that one's soul is "there."

According to Kim, this lack of defined spatial relations for souls makes it virtually impossible to attribute any causal activity to them, because they cannot be paired with anything else on the basis of physical relations. As far as we know, he says, causal relationships require some sort of spatiotemporal reference:

Causal relations must be selective and discriminating, in the sense that there can be two objects with identical intrinsic properties such that a third object causally acts on one but not the other, and,

---

---

similarly, that there can be two intrinsically indiscernible objects such that one of them, but not the other, causally acts on a third object.... This calls for a principled way of distinguishing intrinsically indiscernible objects in causal situations, and it is plausible that spatial relations provide us with the principal means for doing this.<sup>5</sup>

This pairing problem allegedly shows that immaterial souls cannot participate in causal relations, either with bodies or with other souls. Kim concludes that "the very idea of immaterial, nonspatial entities precludes them from entering into causal relations; in fact, I think that the very idea of such objects may well be incoherent and unintelligible."<sup>6</sup>

Essentially, the pairing problem exposes substance dualism as a form of epiphenomenalism. Consequently, if positing immaterial souls provides no helpful solution to the problem of mental causation, then we should ask why they should be believed to exist; they would be unnecessary entities in our theory of the mind, explanatorily useless, and so should be eliminated. This leads us back to physicalism, where Kim believes he has steered us to accept reductionism as the only appropriate answer to the problem posed by mental causation.

But has he jumped to this conclusion too quickly? Kim argues that we must assess whether substance dualism "fares better" on the issue of mental causation. Can one logically move from the premises Kim offers to the conclusion he wants? I contend that the explanatory step in which Kim justifies his essential premise that "our idea of causation requires that the causally connected items be situated in a space-like framework"<sup>7</sup> is missing, which leaves a hole for the substance dualist to rejoin.

This hole is important, because Kim's entire argument throughout the book depends on elimination of the alternatives—he tirelessly concedes that the view we are "left with," namely a kind of functional reductionism, has its own imperfections (particularly, that reductionism may not be able to account fully for conscious, qualitative mental experiences), but that none



---

---

of the alternative views stand up to criticism regarding mental causation. Presuming the importance of mental causation for our understanding of ourselves, we should, *ceteris paribus*, prefer a theory that can account for mental causation. Although he is careful to provide positive arguments for reductionism, recognizing that preserving mental causation does not automatically make a theory the right one,<sup>8</sup> these positive arguments become less compelling if it can be shown that reductionism is not the only view that is consistent with mental causation.

Kim's argument against substance dualism is different from his argument against non-reductionism (*viz.*, the exclusion argument), because in the latter case he provides principles that a physicalist must (or at least *should*) hold and shows that these principles are logically inconsistent with the idea of mental causation if one takes a nonreductive view of the mental. In the case against substance dualism, all that he shows is that if we are to understand how a soul causally interacts with a body, it cannot be the same way that two bodies interact. But this is, perhaps, to be expected. Why, the substance dualist might ask, should we expect souls to behave according to the exact same rules that govern bodies?<sup>9</sup> What Kim wants to have as a premise in his argument is that causal relations can only exist if they model the interactions between two physical entities; in other words, for souls to have causal powers, they must be brought into the physical realm, in which case they seem to lose their distinctive non-physical quality. However, this premise begs the question against the substance dualist because it assumes physicalism, but the truth of substance dualism entails the falsity of physicalism.

The premise mentioned above will be rejected by substance dualists because nothing in their worldview relies on the assumption of physical explanations being exhaustive. Thus, the pairing problem does not produce the same kind of inconsistency as the exclusion argument. Even if the dualist admits that causal relations between spiritual and corporeal entities are wholly mysterious, that does not *ipso facto* discredit her view. Presumably, the *concept* of a soul that causes certain events to occur within the body is neither incoherent nor particularly enig-

---

---

matic. It becomes difficult when one attempts to explain *how* this process might work. Still, the possibility of there being an explanation for mental causation is different from our ever knowing this explanation. Kim has only attacked the latter, not the former.

Many substance dualists will claim that the mystery of mind-body interaction does not preclude its possibility, and if we have independent reasons for thinking that immaterial objects exist (say, philosophical arguments for the existence of God), then physicalism may not offer a "better" explanation of mentality, because as Kim himself shows throughout the book, physicalism seems unable in principle to account for all of the important components of mentality, either by reducing everything or by positing mental properties. In response to this claim, Kim needs to either offer additional arguments that substance dualism is logically incompatible with mental causation (not merely that it lacks a physical explanation), or he needs to offer arguments meant to show the implausibility of the existence of immaterial entities, *based on principles to which dualists and physicalists alike are committed*. The lack of a physical explanation alone does not get him there.

One example of an argument which, if sound, *would* entail the logical inconsistency of substance dualism (given a principle that the dualist ought to accept) might be the following. The scientific community broadly supports what is now called the first law of thermodynamics, or the conservation of energy principle. According to this principle, the amount of energy within a closed system does not change; it is neither created nor destroyed. But on the substance dualist's view of mental causation, the soul causally interacts with the body, which means that there must be some kind of energy exchange between these entities. Energy somehow departs from the soul, and then *enters* the physical domain, where it causes the motion of particles, and vice versa. This, of course, violates the first law of thermodynamics. Now we have a new dilemma for the substance dualist: either accept substance dualism at the cost of abandoning (or at least modifying) the first law of thermodynamics, or abandon

substance dualism and adopt some kind of physicalist view.

Note the similarity between this argument and Kim's exclusion argument. It is thus surprising that Kim relies on a weaker argument, the pairing problem, to eliminate substance dualism from the realm of viable options in the philosophy of mind. This, as I have shown, produces a gap in his argument: Kim cannot argue that reductionism is the only option unless he truly *eliminates* the other options. But to show the inconsistency that he wants, he has to show that some view is inconsistent with principles that *both parties* accept, as he did in the exclusion argument.

As I said earlier, one of the distinctive features of the philosophy of mind is the degree to which it is tied to additional issues relating to the ontology of existent objects. If one believes that God or other spiritual entities exist (more importantly, if one has a reasoned view of their existence), then physicalism (especially reductive physicalism) concerning the mental may not be a compelling view. Conversely, one who subscribes to ontological physicalism cannot accept the existence of immaterial substances, and thus cannot use them to explain the mind. Any arguments attempting to span this ontological gap must appeal to shared principles. This may require more careful work than what Kim offers here.<sup>10</sup> After all, his book is addressed to physicalists, who would presumably share many of his ontological intuitions. Nevertheless, without this appeal, at least so far as the above examination shows, Kim's elimination of substance dualism remains incomplete.

---

---

**NOTES**

<sup>1</sup> This analogy comes from Kim, *Philosophy of Mind*, 6.

<sup>2</sup> Kim, *Physicalism, or Something Near Enough*, 22.

<sup>3</sup> *Ibid.*, 27, 29.

<sup>4</sup> *Ibid.*, 9.

<sup>5</sup> *Ibid.*, 85.

<sup>6</sup> *Ibid.*, 92.

<sup>7</sup> *Ibid.*, 84.

<sup>8</sup> He concedes, "Whether or not the mental can be reduced to a physical base is an independent question that must be settled on its own merits. Those of us who believe in mental causation should hope for a successful reduction. But again this is only a wish; it doesn't make reducibility real or reductionism true" (Kim, *Physicalism, or Something Near Enough*, 161).

<sup>9</sup> This question brings up the problem of determinism for reductive physicalism. If all our mental activity strictly obeys physical laws, then free will and rationality appear to be threatened. This issue calls for greater treatment than can be provided here, but it nevertheless seems, in principle, to justify the substance dualist's concern for keeping souls *out* of the physical realm.

<sup>10</sup> Even my argument above using the first law of thermodynamics needs more explanation and detail in order to be considered conclusive. The goal was rather to introduce a kind of argument which would better suit Kim's purposes. Other philosophers have offered arguments along these lines.

**WORKS CITED**

Kim, Jaegwon. *Philosophy of Mind*. Cambridge, MA: Westview Press, 2006.

Kim, Jaegwon. *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press, 2005.