Philosophia Christi
Vol. 24, No. 2 © 2022

The Argument from Reason and the Dual Process Reply

Dwayne Moore
University of Saskatchewan

Abstract: The argument from reason states that if naturalism is true, then our beliefs are caused by physical processes rather than being causally based in their reasons, so our beliefs are not knowledge—including the belief in naturalism itself. Recent critics of the argument from reason provide dual process replies to the argument from reason—our beliefs can have both a naturalistic cause/explanation and be caused/explained by its reasons, thereby showing that naturalism can accommodate knowledge. In this paper I consider three dual process replies and conclude that none of them are successful.

Roughly speaking, the argument from reason states that if naturalism is true, then our beliefs are caused by physical processes rather than being causally based in their reasons, so our beliefs are not knowledge—including the belief in naturalism itself, rendering naturalism self-stultifying. Recent critics of the argument from reason provide dual process replies to the argument from reason: our beliefs can have both a naturalistic cause/explanation and be caused/explained by its reasons, thereby showing that naturalism can accommodate knowledge.[1] In this paper I consider three dual process replies, namely, a nonreductive physicalist version, a reductive physicalist version, and an explanatory dualist version. I argue that each version faces trenchant objections, leaving naturalism once again susceptible to self-stultification.

This paper has five parts to it. After outlining the central planks constituting the argument from reason (section 1), I outline three different dual process replies to the argument from reason, as recently put forth by David Kyle Johnson and Peter van Inwagen (section 2). I then argue that the reductive physicalist version of the dual process reply faces the quasation problem and the only physical causation problem (section 3), the nonreductive physicalist version of the dual process reply faces the causal exclusion problem and the coincidental correlation problem (section 4), while the explanatory dualist version of the dual process reply faces the explanatory exclusion prob-

1. See, e.g., David Johnson, "Con: Naturalism Undefeated," in *C. S. Lewis's Christian Apologetics: Pro and Con*, ed. Gregory Bassham (Leiden: Brill, 2015), 91–104; "Reply to Victor Reppert," in *C. S. Lewis's Christian Apologetics*, 113–20; "Retiring the Argument from Reason," *Philosophia Christi* 20 (2019): 541–63; and Peter van Inwagen, "C. S. Lewis' Argument against Naturalism," *Res Philosophica* 90 (2013): 113–24.

lem (section 5). With no viable version of the dual process reply available, naturalism remains susceptible to self-stultification.

1. The Argument from Reason

The argument from reason has a venerable yet controversial history. Ancient hunches from notable philosophers such as Epicurus and Immanuel Kant gesture in the direction of the argument from reason. Epicurus notes that "he who says that all things happen by necessity can hardly find fault with the one who denies that all happens by necessity; for on his own theory this very argument is voiced by necessity."[2] Immanuel Kant also cryptically insists "one cannot possibly think of a reason that would self-consciously receive guidance from any other quarter with regard to its judgments, since the subject would not then attribute the determination of judgment to his reason, but to an impulse."[3] In the twentieth century C. S. Lewis popularized the argument from reason, only to quickly receive what some considered decisive criticism from Elizabeth Anscombe.[4]

In recent years the argument from reason has been articulated differently by numerous philosophers, fashioning the argument from reason into a tapestry of related concerns revolving around the inability of naturalism to deliver certain epistemic features requisite for knowledge. Indeed, Victor Reppert highlights numerous different problems constituting the argument from reason, including a problem with intentionality, a problem with truth, a problem with mental causation, a problem with the relevance of logical laws, a problem with the unity of consciousness during deliberation, and a problem with the reliability of our rational faculties.[5] In this section I provide

2. Epicurus, *Epicurus: The Extant Remains*, trans. Cyril Bailey (Oxford: Clarendon, 1926), 113.

3. Immanuel Kant, *Groundwork of the Metaphysics of Morals*, trans. Mary Gregor and Jens Timmermann (New York: Cambridge University Press), 448.

4. C. S. Lewis, *Miracles: A Preliminary Study* (London: Geoffrey Bles, 1947). For more complete lists of advocates of the argument from reason, both historical and contemporary, see Brandon Rickabaugh and Todd Buras, "The Argument from Reason," *Philosophia Christi* 19 (2017): 382–3; Gregory Bassham, "Introduction: Oxford's Bonny Apologist," in *C. S. Lewis's Christian Apologetics*, 6; and G. E. M. Anscombe, "A Reply to Mr. C. S. Lewis's Argument that 'Naturalism' Is Self-Refuting," *The Socratic Digest* 4 (1948): 7–15.

5. Victor Reppert, "Several Formulations of the Argument from Reason," *Philosophia Christi* 5 (2003): 9–33. It is also worth noting that some direct the argument from reason against physical *determinism*, rather than against *physical* causation. William Hasker, e.g., says: "if physical determinism is true, then the principles of rational inference are superfluous and inoperative . . . and if this is so, then there is nothing to be said for determinism" ("The Transcendental Refutation of Determinism," *Southern Journal of Philosophy* 11 (1973): 182). See also Colson, "The Transcendental Argument," *Southern Journal of Philosophy* 20 (1982): 23; and Lewis, *Miracles*, 2nd ed., 18–20. Even the initial quotations from Epicurus and Kant could be interpreted as being a difficulty with physical determinism, rather than physical causation. The argument from reason against determinism draws support from the fact that what we presently believe would be

a contemporary reconstruction of one especially prominent strand of this tapestry, namely, the problem with mental causation.

I begin with a slight modification to the traditional target of the argument from reason. Whereas the target has often been naturalism, or materialism, I instead calibrate on physicalism, their close kin.[6] Nothing of substance rides on the emphasis on physicalism, but I center on physicalism because it is more commonly emphasized in contemporary literature. This is because materialism focuses on matter, but twentieth-century physics reveals many features of the universe that do not resemble matter as traditionally construed—fields, waves, strings, and so forth. So, most presently focus on physicalism rather than materialism. As Jessica Wilson summarizes: "contemporary physics has reported that the relatively fundamental entities have few, if any, of the characteristics of the material; and thus materialism has been rendered a has-been. Its foundationalist spirit has survived in physicalism."[7] For their own part, naturalists tend to emphasize all the higher-level objects such as rabbits, mountains, minds and societies that all exist within nature, whereas physicalists emphasize the lower-level microphysical objects of physics. However, physicalists often add a supervenience principle to physicalism (as is done below), such that microphysical particles and everything constituted of microphysical particles exist, thereby including higher-level objects existing in the natural world as consistent with physicalism. As Jaegwon Kim puts it: "Physicalism is the doctrine that all things that exist are entities recognized by the science of physics, or systems aggregated out of such entities."[8]

Physicalism can be roughly defined as "the doctrine that there is nothing over and above the physical."[9] If we take away all the physical stuff, there would be no stuff left—no ghosts, no gods, no Cartesian souls. If we make a physical duplicate of the universe, that universe would be a duplicate universe *simpliciter*—no rabbits, minds, or societies would be missing or out of place. Since the physical is all that exists, everything that happens has a complete physical cause, and everything that exists is constituted of physical objects. These points are usually dissolved into the following two emblematic principles of physicalism:

determined by forces before our births in the remote past, so what we believe is not within our control. The problem with this articulation of the argument from reason is that quantum physics suggests that the physical universe is probably indeterministic, thereby undermining a central premise of this version of the argument from reason. However, the central problem that physicalism has with generating knowledge does not rest in the premise that our beliefs are caused by physical determinism. Rather, the central problem that physicalism faces lies in its claim that our beliefs are completely caused by physical causes, regardless of whether those physical causes are deterministic or indeterministic.

6. For naturalism, see Lewis, *Miracles*, 2nd ed., 17. For materialism, see, e.g., Henry E. Allison, "Kant's Refutation of Materialism," *Monist* 79 (1989): 190–209; and James B. Pratt, *Matter and Spirit* (London: Forgotten Books, 1922), 19.

7. Jessica Wilson, "On Characterizing the Mental," *Philosophical Studies* 131(2006): 62.

8. Jaegwon Kim, *Philosophy of Mind*, 3rd ed. (New York: Routledge, 2011), 11.

9. Justin Tiehen, "Physicalism," *Analysis* 78 (2018): 537.

> *Physical Causal Completeness*: "If a physical event has a cause at *t*, it has a sufficient physical cause at *t*."[10]
>
> *Supervenience:* "Whenever something has a mental property, *M*, at *t*, it does so in virtue of the fact that it has, at *t*, a physical base property, *P*, where *P* necessitates *M*."[11]

Consider Tran, who believes that "Socrates is mortal." His reason for believing that Socrates is mortal is the logical relation between his other beliefs that "All humans are mortal," and "Socrates is human." According to supervenience, Tran's reason here, call it M_1, is determined by some physical state P_1 in his brain. According to physical causal completeness, this physical state P_1 is the sufficient cause of some physical effect P_2 in his brain, which itself serves as the subvenient base of Tran's belief M_2 that "Socrates is mortal."

The contours of the problem are visible from this initial setup. On physicalism, all mental states are completely caused/determined by physical states. So, Tran's belief that Socrates is mortal is completely determined by its physical base, which itself is completely caused by prior physical processes in Tran's brain, seemingly leaving Tran's reason for believing that "Socrates is mortal" causally irrelevant. C. S. Lewis summarizes this point by saying that our reasons for belief are "a trifle" on naturalism, while Reppert concludes that it is a "user illusion" to think our reasons impact our beliefs.[12] After all, physical processes completely cause our beliefs, so our reasons "have nothing to do with it."[13]

This is a pernicious result when supplemented with some widely held epistemological assumptions. To see these assumptions, consider Donald Davidson's famous distinction between reasons for acting and acting for reasons. While hiking Samir notices someone clinging to the edge of a cliff. After Samir starts pulling him up, he notices that the man is actually a bitter enemy, so he thinks about letting him just fall. While contemplating this possibility, Samir's strength runs out, causing him to accidentally lose his grip, which causes his enemy to fall. Samir had reasons for releasing his grip, but he did not release his grip for those reasons. The difference, Davidson insists, is that we act for reasons when those reasons cause us to act: "a person can have a reason for an action, and perform the action, and yet this reason not be the reason why he did it. Central to the relation between a reason and an action it explains is the idea that the agent performed the action *because* he had the

10. Jaegwon Kim, "Mental Causation," in *Oxford Handbook of Philosophy of Mind*, ed. Ansgar Beckermann, Brian P. McLaughlin, and Sven Walter (New York: Oxford University Press, 2009): 38.

11. Ibid., 40.

12. Lewis, *Miracles*, 2nd ed., 25.

13. Victor Reppert, "Pro: The Argument from Reason Defended," in *C. S. Lewis's Christian Apologetics*, 87.

reason."[14] Motivated by these types of considerations, Davidson was influential in establishing the causal theory of action as the widely accepted standard conception of action, according to which actions must be caused by reasons.

A parallel distinction between reasons for belief and believing for those reasons appears in epistemology. One juror believes truly, based on the substantial evidence she heard, that the defendant is innocent. Another juror hears the same evidence but is still undecided. He is a superstitious man, however, so he consults his horoscope, which tells him to see the best in people today, so he also believes truly the defendant is innocent. The superstitious juror has reasons or justification for his true belief—so called, having propositional justification—but his true belief is not based on those reasons—so called, lacking doxastic justification. The difference, according to many epistemologists, is that we believe truly for reasons when those reasons cause us to believe truly. Here is a sample of contemporary epistemologists espousing this view:

> The causal theory is quite intuitive. If my belief that, say, I will not catch my flight is based on my belief that the traffic on the streets is heavy, then I hold the former belief because I hold the latter belief. The same holds for my belief that the traffic is heavy as a result of seeing so many cars on the streets. In both cases my reasons causally explain . . . what I believe and they could do so, it seems, only if they causally sustain the relevant belief.[15]

> There is the kind of reasons that we are referring to when we talk about a person's evidence, and it is here that talk of believed (or, known) propositions is appropriate. Second, there are what I shall call causal reasons for a belief. Causal reasons are, among other things, events or states of the person who has the belief.[16]

> If your reason for forming a certain belief is "represented" by some of your antecedent mental states, then your formation of that belief is—as epistemologists often put it—"based on" those antecedent mental states. Like most contemporary epistemologists, I take this "basing relation" to be a kind of causal relation: for your formation of this new belief to be based on those antecedent mental states, you must have formed that new belief precisely because you were in those antecedent mental states—where this is the 'because' of ordinary causal explanation.[17]

14. Donald Davidson, "Actions, Reasons, and Causes," *Journal of Philosophy* 60 (1963): 691.

15. Hamid Vahid, "Triangulation, Content and the Basing Relation," *Grazer Philosophische Studien* 78 (2009): 233.

16. Marshall Swain, "Justification and the Basis of Belief," in *Justification and Knowledge*, ed. George S. Pappas (Boston: D. Reidel, 1979), 27.

17. Ralph Wedgwood, "The Normative Force of Reasoning," *Noûs* 40 (2006): 661.

These passages reflect the predominant contemporary view that reasons must not only justify true beliefs, but those true beliefs must be based on those reasons, where this basing relation is usually understood as a causal relation.

If true beliefs are not causally based on their reasons, then those true beliefs, even though they have a justification, are not considered knowledge. This is nothing more than an extension of traditional problems of epistemic luck. Sally spins around until dizzy, points in some random direction, and says "Montreal is in that direction." Sally's belief is luckily true, but because her true belief comes about due to luck, it is not considered knowledge. To avoid epistemic luck, true beliefs must be justified as well. Sally asks three truckers who recently returned from Montreal, and familiarizes herself with Canadian maps, so she points and says, "Montreal is in that direction." Sally's true belief is now justified, so it is considered knowledge. But, since Edmund Gettier at least, it is apparent that a true belief can be justified yet still fall short of knowledge due to epistemic luck. John Turri imagines a juror who listens to the ample justification for the defendant's guilt, yet truly believes the defendant is guilty based on a coin flip, so does not know the defendant is guilty.[18] Thus, if true beliefs are not based on their reasons, those beliefs are not doxastically justified, which leads to the failure of knowledge. Here are some articulations of this view:

> A belief's being true and justified is not sufficient for knowledge, . . . one might have a justified true belief that *P* while believing that *P* solely on grounds other than those that actually justify *P*. That is, one might believe that *P* solely for the 'wrong reasons,' even though *P* is justified and true. In such a case, one does not know that *P*. Thus propositional knowledge requires that one's believing that *P* be 'adequately related' to the justifying evidence for *P*. In other words, such knowledge requires that one's believing that *P* be *based on* the justifying evidence.[19]

> And on the assumption that doxastic justification is necessary for *knowledge* . . . basing features principally in causal explanations that make reference to reasons as causes of *why* knowledge comes about. Obviously, such reasons would need to cause states of knowledge *qua* undefeated justifiers above any pertinent threshold for knowledge. In sum, by presenting a reason that is both a justifier (modulo such qualifications) and a cause, basing is a salient component in an epistemically significant, causal explanatory account of why *S* knows *p*.[20]

18. John Turri, "Believing for a Reason," *Erkenntnis* 74 (2011): 383. See also Kathleen Lennon, *Explaining Human Action* (London: Duckworth, 1990), 38; Bill Brewer, "Mental Causation: Compulsion by Reason," *Proceedings of the Aristotelian Society*, supplement 69 (1995): 242; Swain, "Justification," 25; and Gilbert Harman, "Knowledge, Reasons, and Causes," *Journal of Philosophy* 67 (1970): 842.

19. Paul Moser, *Knowledge and Evidence* (New York: Cambridge University Press, 1989), 45.

20. Jesper Kallestrup, "From Epistemic Basing to Epistemic Grounding," in *Well-Founded Belief: New Essays on the Epistemic Basing Relation*, ed. J. Adam Carter and Patrick Bondy (Routledge, 2019), 255.

People who believe truly with justification, but not based on that justification, can be caused to believe via numerous stray arational causal paths—horoscopes, coin flips, prejudice, emotion, and so forth—so those true beliefs are not doxastically justified, hence not known.

The most relevant arational causal source that causes true beliefs while bypassing their justification is the case where physical processes in the brain determine beliefs.[21] This claim has two parts to it. First, that it is physical processes in the brain that determine beliefs, which is established above as endorsed by physicalists. Second, that these physical processes in the brain are arational, so do not serve as reasons for belief. To see this, consider a prototypical physical event in the brain: the firing of neurons. Positively charged sodium ions surge into neurons via ion channels, altering the electrical voltage of resting neurons, causing them to send electrical signals down their axon, eventually releasing chemicals called neurotransmitters toward other neurons, which, in turn, fire as well. This rudimentary articulation suffices to demonstrate that physical processes in the brain involve electrical and chemical processes, but they do not involve logical relations, justifications, or reasons. As Jerry Fodor notes, "I suppose that sooner or later the physicists will complete the catalogue they've been compiling of the ultimate and irreducible properties of things. When they do, the likes of spin, charm, and charge will perhaps appear on their list. But aboutness surely won't; intentionality simply doesn't go that deep."[22]

So strongly held is the view that mental states such as reasons do not appear at the physical level, some so-called *via negativa* physicalists literally

21. Numerous epistemologists mention that physical processes in the brain would not serve as a justificatory base for true beliefs to be counted as knowledge. Hamid Vahid says "one may consider some of the causal ancestors of a perceptual belief, say, certain neurophysiological states of one's brain. Although the perceptual belief is clearly dependent on the pertinent neural state, it is not based on it" ("Triangulation, Content and the Basing Relation," 238). Or again, William Alston says, "not just any kind of causal dependence will do. My belief that *p* is causally dependent on a certain physiological state of my brain, but the former is not based on the latter" ("An Internalist Externalism," *Synthese* 74 (1988): 265). Ian Evans is among many who argue that the basing relation must be a psychological relation: "I do think, however, that the basing relation is a psychological relation, in particular, a relation between mental states. I find it strongly counter-intuitive to suppose that a belief might be based on e.g., a brain tumor, a cup of tea, etc. This suggests a desideratum on theories of the basing relation: *Desideratum*: An adequate theory should have, as an interesting consequence, that beliefs can only be based on other mental states" ("The Problem of the Basing Relation," *Synthese* 190 (2013): 2945). See also Andrew Moon, "All Evidential Basing Is Phenomenal Basing," in *Well-Founded Belief*, 34; Duncan Pritchard, "Epistemological Disjunctivism and Factive Bases for Belief," in *Well-Founded Belief*," 239–40; and Keith A. Korcz, "The Causal-Doxastic Theory of the Basing Relation," *Canadian Journal of Philosophy* 30 (2000): 540.

22. Jerry Fodor, *Psychosemantics* (Cambridge, MA: MIT Press, 1987), 97. Hilary Putnam goes so far as to say that "it is this same mindlessness of nature that makes the action guiding predicates . . . 'is a justified belief' seem 'queer'" (*Reason, Truth and History* (New York: Cambridge University Press, 1981, 211)).

define the physical as anything that is not mental. Here are some articulations of that view:

> A dynamics which introduced forces with immanent purpose, and hence teleological causation at the base level, would not sustain a program maintaining the spirit of physicalism.[23]

> If it were to turn out that to account for certain clearly physical events physicists needed to posit fundamental intentional, or phenomenal, properties, then the resulting theory would not be physical.[24]

> It isn't crucial that you know exactly what a complete physics would include. Much more important is to know what it won't include . . . the sentient, say, or the intentional.[25]

On this view physicalism is actually defined as the view that mental processes are not present at the physical level, so the claim that mental processes such as rationality and reasons are not present in physical brain processing should not be controversial.

To briefly summarize, on physicalism, Tran's true belief that "Socrates is mortal" is determined by physical processes in his brain, which are arational causes of his true belief. But in order for Tran to know that "Socrates is mortal," this true belief must be caused by his reasons, so, on physicalism, Tran does not know that "Socrates is mortal." But physicalism is itself a belief that physicalists hold—say Tran believes that "there is nothing over and above the physical." While Tran may have justification for this belief, his belief is not causally based on this justification. Rather, his belief is determined by physical processes which are arational causes, so he does not know that "there is nothing over and above the physical." Hence, physicalism cannot be known as true, or, in other words, physicalism is self-stultifying.

2. The Dual Process Reply

The most common response to this argument from reason is to grant that physical states cause the belief M_2, say Tran's belief that "Socrates is mortal," but this belief is (also) caused by his reasons M_1. Call this the dual process reply. It preserves the physicalist view that physical states cause and/or determine M_2 while also securing knowledge via the claim that Tran's belief that

23. Keith Campbell, "Critical Notices," *Philosophy and Phenomenological Research* 57 (1997): 224.

24. Barry Loewer, "From Physics to Physicalism," in *Physicalism and its Discontents*, ed. Carl Gillett and Barry Loewer (New York: Cambridge University Press, 2001), 40.

25. David Papineau, "The Rise of Physicalism," in *Physicalism and its Discontents*, 12. See also Wilson, "On Characterizing the Mental," 72; Barbara Montero, "Post-Physicalism," *Journal of Consciousness Studies* 8 (2001): 67; Justin Tiehen, "Physicalism Requires Functionalism," *Philosophy and Phenomenological Research* 93 (2016): 10; David Spurrett and David Papineau, "A Note on the Completeness of 'Physics,'" *Analysis* 59 (1999): 27.

"Socrates is mortal" is causally based in his justificatory reasons M_1 for this belief. While the dual process reply is a common response to the argument from reason,[26] I will focus on two recent articulations of this dual process reply.

Peter van Inwagen provides a dual process response to Lewis's argument from reason. He says that humans can both have beliefs and be physical things, so "there seems to be no reason to deny that that human being's believing certain things might be the *cause* of his or her believing certain other things."[27] This means, in the context of the Lewis and van Inwagen discussion, that physical causes in the remote past determine the reasons M_1 of some human today, which then cause the belief M_2 of that human today. To use van Inwagen's example:

> Phoebe's having the belief that Lewis fought in the First World War was caused by the universe's having been in such-and-such a state many billions of years ago—and that it was also caused by the two belief facts I have already imagined (her having the belief that it says so in *Surprised by Joy*; her having the belief that autobiographies are trustworthy in respect of statements whose falsity is easily detectable).[28]

Phoebe's true belief has a complete physical cause yet is also caused by her justifying beliefs, so her true belief is doxastically justified knowledge. It is therefore possible to give both a physical/causal explanation of Phoebe's belief (that is, Phoebe believes that Lewis served in the First World War because of deterministic physical causes) and a reasons explanation of Phoebe's belief (that is, Phoebe believes that Lewis served in the First World War because of her reasons). So, van Inwagen concludes, "naturalism is consistent with some of our beliefs being grounded in reasoning."[29]

Kyle Johnson provides another robust expression of the dual process reply in a series of papers responding to Victor Reppert's expression of the argument from reason. Johnson thinks the argument from reason presumes that the physical processes determining our beliefs rule out mental processes such as rationality as a causally irrelevant trifle, which is what makes proponents of the argument from reason think that reasons cannot cause beliefs. But he notes that naturalists deny this presumption, preferring instead to say that mental processes metaphysically supervene upon, or are identical with, physical processes, so mental processes are causally relevant (since changing Tran's mental states requires changing Tran's physical states) and necessary

26. G. E. M. Anscombe, "C. S. Lewis' Rewrite of Chapter III of Miracles," in *C. S. Lewis and His Circle*, ed. Roger White, Judith Wolfe, and Brendan Wolfe (New York: Oxford University Press, 2015): 16–17; Theodore Drange, "Several Unsuccessful Formulations of the Argument from Reason," *Philosophia Christi* 5 (2003): 49; and Keith Parsons, "Further Reflections on the Argument from Reason," *Philo* 3 (2000): 97–8.

27. Van Inwagen, "C. S. Lewis' Argument against Naturalism," 121.

28. Ibid., 122.

29. Ibid., 124.

(since Tran's physical states guarantee the presence of Tran's mental states).[30] He explains: "When certain neurophysical states obtain, certain mental events necessarily obtain as well. The mental phenomenon of reasoning is strictly necessary if the conclusion in question is to be drawn . . . therefore, the propositional content is relevant to the conclusion's derivation."[31] Given the fact that the physical processes necessitate the presence of the mental processes, Tran's belief that Socrates is mortal can be explained in terms of his reasons for his belief and also explained in terms of the physical/causal processes giving rise to his belief.

While Johnson and van Inwagen offer their own particular versions of the dual process reply, I want to briefly summarize and expand upon three central aspects of the dual process reply that they refer to and expresses the dual process reply more generally.[32] First, there is a reductive physicalist ver-

30. Johnson, "Con," 95, 96.

31. Ibid., 101.

32. Johnson also argues that naturalism is compatible with numerous other models of the mind-brain relation as well, including Thales's panpsychism, Russellian monism, Chalmers's naturalistic dualism, and emergentism, so long as mental states exist within the spatiotemporal universe rather than existing in souls outside of the natural universe ("Reply to Victor Reppert," 116; "Retiring the Argument from Reason," 548–52). His expansive definition of naturalistically acceptable mental states is partly due to his broad understanding of physical causal completeness: "causal closure is simply the thesis that nothing from beyond the natural world causally effects it . . . but since naturalists can maintain that mentality is a part of nature, the naturalist can think mentality is causally operative (even at the basic level) without denying causal closure" ("Retiring the Argument from Reason," 552). This view is sometimes called broad physical causal completeness (David Papineau, "Must a Physicalist Be a Microphysicalist?," in *Being Reduced: New Essays on Reduction, Explanation, and Causation*, ed. Jakob Hohwy and Jesper Kallestrup (New York: Oxford University Press, 2008), 127–9; Jaegwon Kim, "Does the Problem of Mental Causation Generalize?," *Proceedings of the Aristotelian Society* 87 (1997): 293), according to which all physical effects have a sufficient broadly physical cause, where a broadly physical cause includes any object or property existing within the spatiotemporal universe and dependent upon microphysical objects and properties, including trees, rabbits, mountains, and possibly minds, and social structures. Without judging whether broad physical causal completeness is sufficient for satisfying naturalism, physicalists typically consider microphysical causal completeness (sometimes called the completeness of physics) to be central to physicalism as well, according to which all physical effects have sufficient microphysical causes, where a microphysical cause includes the objects and properties of a completed physics. Jaegwon Kim, e.g., says: "If you reject this principle, you are ipso facto rejecting the in-principle completability of physics . . . it is safe to assume that no serious physicalist could accept such a prospect" (*Mind in a Physical World* (Cambridge, MA: MIT Press, 1998), 40; cp. David Papineau, *Philosophical Naturalism* (Blackwell: Oxford, 1993), 16; Barbara Montero, "What Does the Conservation of Energy Have to Do with Physicalism?," *Dialectica* 60 (2006): 393; Brian McLaughlin, "The Causal Closure of the Physical and Naturalism," in *Oxford Handbook of Philosophy of Mind*, ed. B. McLaughlin, Ansgar Beckermann, and Sven Walter (New York: Oxford University Press, 2008): 53–4; Augustin Vicente, "On the Causal Completeness of Physics," *International Studies in the Philosophy of Science* 20 (2006): 149–71. Physicalists endorse microphysical causal completeness because the leading argument for physicalism (the so-called causal argument for physicalism) requires the truth of microphysical causal completeness. The causal argument states that physics is complete, so all effects have complete physical causes, so everything that has causal power is physical, so there exist no causally potent nonphysical minds. As David Yates says, "the 'completeness of

sion of the dual process reply, according to which Tran's reasons M_1 are identical with his causally efficacious physical processes P_1 that cause P_2, which is identical with Tran's belief M_2 that "Socrates is mortal." On this model, physicalism is true by virtue of Tran's reasons being causally efficacious physical processes in his brain, while Tran knows that "Socrates is mortal" because his belief is causally based on his reasons.

Second, there is also a nonreductive physicalist version of the dual process reply, according to which Tran's reasons M_1 supervene upon, hence are necessitated by, his causally efficacious physical processes P_1 that cause P_2, which subvenes Tran's belief M_2 that "Socrates is mortal." On this view, physicalism is true by virtue of Tran's reasons supervening on the causally efficacious physical processes in his brain that cause his beliefs, while Tran knows that "Socrates is mortal" because his belief is necessarily preceded by his reasons which also cause his belief.

Third, there is a dual explanation version of the dual process reply, according to which the causal processing that occurs between Tran's reason M_1, which is identical with or supervenes upon the physical cause P_1, and Tran's belief M_2, which is identical with or supervenes upon the physical effect P_2, can be given both a physical explanation and a mental explanation. On this view, physicalism is true by virtue of there being a complete physical explanation of Tran's belief that "Socrates is mortal," while Tran knows that "Socrates is mortal" because there is a reasons explanation of Tran's belief that "Socrates is mortal."

3. Reductive Physicalism and the Argument from Reason

In the next three sections I outline why none of these dual process replies are ultimately successful, beginning with the dual process reply that adverts to reductive physicalism. According to this model, Tran's reasons M_1 are identical with the physical state P_1 in Tran's brain that is the sufficient physical cause of P_2, which is identical with Tran's belief M_2 that "Socrates is mortal." Thus, physicalism is true because Tran's belief that "Socrates is mortal" only has a sufficient physical cause, while Tran's belief is causally based on his reasons, since his reasons are identical with the physical cause of his belief.

There are three problems with this reductive physicalist model. First, there are notorious difficulties associated with delivering the required iden-

physics' is the key premise in the causal argument for physicalism" ("Emergence, Downwards Causation and the Completeness of Physics," *Philosophical Quarterly* 59 (2009): 110). Or, as David Papineau frames it: "if the completeness of physics is right, and all physical effects are due to physical causes, then anything that has a physical effect must itself be physical" ("The Rise of Physicalism," 8). To abandon microphysical causal completeness would be to render false the foundational premise in the leading argument in support of physicalism. For this reason, physicalists do not usually make this move. But, if microphysical causal completeness is accepted, it excludes broadly physical causes when paired with a causal exclusion principle.

tity, including the multiple realizability of mental states and the discernibly distinct natures between mental states and physical states. One relevant example of the latter difficulty occurs when physicalists define the physical level as a nonmental level of physical forces such as charge and gravitational attraction, which rules out the possibility that the physical level is identical with the mental. As Justin Tiehen explains: "Such [*via negativa*] theories say that certain entities (properties, events, etc.) are both mental and physical. . . . This amounts to saying those entities are both mental and not mental—a contradiction."[33] It is difficult to see how, after emphasizing the fact that physical states are arational, physicalists can now say that physical states are rational. Partly for this reason, many physicalists turn to nonreductive physicalism where physical states subvene, but are not identical with, mental states.

Even if mental states can be identified with physical states, two other problems remain. First, the mental quasation problem, which first emerged as an objection to Donald Davidson's anomalous monism. Anomalous monism secures mental causation by taking mental events to be identical with physical events, though these events can be described in mental vocabulary or physical vocabulary.[34] Critics raise the so-called mental quasation problem against anomalous monism, according to which events cause *qua*, or in virtue of, their properties. Green pears tip the scale in virtue of their mass, not in virtue of their greenness. Similarly, events cause in virtue of their physical properties (to secure physical causal completeness), leaving their mental properties as causally irrelevant.[35] So, even though Tran's belief that "Socrates is mortal" is identical with physical state P_1, Tran's belief is caused by physical state P_1, in virtue of its arational physical properties such as its pattern of neural firing, not in virtue of the fact that it is a reason for his belief. So, Tran does not believe that "Socrates is mortal" in virtue of the reasons he has, so Tran's belief is not causally (or, quasally) based on his reasons, so he does not know that "Socrates is mortal." Though he has reasons for believing that "Socrates is mortal," he does not hold his belief because of the reasonableness of those reasons, rather he holds his belief because of the electrochemical machinations of his neural processing.[36]

33. Justin Tiehen, "Physicalism," *Analysis* 78 (2018): 543; see also Tiehen, "Physicalism Requires Functionalism," 4; Stoljar, "Physicalism," in *Stanford Encyclopedia of Philosophy*, ed. Edward Zalta, http://plato.stanford.edu/archives/fall2009/entries/physicalism, section 4.5; and Montero, "Post-Physicalism."

34. See, e.g., Donald Davidson, "Thinking Causes," in *Mental Causation*, ed. John Heil and Alfred Mele (Clarendon: Oxford, 1993), 187; "Mental Events," in *Experience and Theory*, ed. Lawrence Foster and J. W. Swanson (Amherst, MD: University of Massachusetts Press, 1970).

35. See, e.g., Terence Horgan, "Mental Quasation," *Philosophical Perspectives* 3 (1989): 51; Ernest Sosa, "Mind-Body Interaction and Supervenient Causation," *Midwest Studies in Philosophy* 9 (1984): 277; Ted Honderich, "The Argument for Anomalous Monism," *Analysis* 42 (1982): 63.

36. Davidson responds to this complaint with a simple cause strategy, which says the cause cannot be divided into properties so it does not cause in virtue of one property over the other:

Reductive physicalists also face the only physical causation problem. On reductive physicalism, Tran's belief that "Socrates is mortal" only has physical process P_1 as its cause. Unlike nonreductive physicalism, emergentism, or substance dualism, Tran does not also have a distinct reason M_1 that causally contributes to his belief that "Socrates is mortal"; rather there is nothing but the physical process P_1. The physical process P_1, however, is an arational process involving electrochemical forces in neurons. So, Tran's belief that "Socrates is mortal" only has an arational physical process involving electrochemical forces in neurons as its cause. Since Tran's belief that "Socrates is mortal" is causally based completely on arational physical processes, Tran does not know that "Socrates is mortal."

It is tempting to appeal to the identity in order to avoid this result: but P_1 *is* Tran's reasons, which cause his belief that "Socrates is mortal," so he knows that "Socrates is mortal"! But the identity does not help. The damage is already done once the reductive physicalist acknowledges that Tran's belief that "Socrates is mortal" has only one cause, which is an arational physical process P_1. Arational causal processes cannot ground knowledge, and Tran's belief is caused only by this arational causal processes. The fact that Tran's reasons are—if possible—identical with his arational brain processing does not magically render his arational electrochemical brain processing rational, rather it only guarantees that Tran's reasoning is nothing but arational electrochemical processing in his brain.[37]

"It is events that have causes and effects . . . it makes no literal sense to speak of an event causing something as mental, or by virtue of its mental properties, or as described in one way or another" (Davidson, "Thinking Causes," 13). But, all identity theories inescapably face the quasation problem. Identity theorists say that the physical state is a mental state, which means that "the state's being physical" is true and "the state's being mental" is true, which is to ascribe mental properties and physical properties to the state. Since the state has mental and physical properties, we can ask always ask: "In virtue of which property does the state cause the effect?" Contrast this with the nonreductive physicalist or the eliminative reductionist, where the physical state is only physical and has no mentality to its essence. On these views "the state's being physical" is true, but "the state's being mental" is false, so the physical state does not have both mental properties and physical properties, so the quasation question does not get off the ground. But, on mind-brain identity theories, the quasation problem can always, by very definition, get off the ground.

37. Johnson appeals to a computer metaphor to demonstrate that rationality can be realized in physical mechanisms. He says: "If reasoning truly is a non-mechanical process, it should not be possible to program a computer to figure out anything—especially something as complicated as chess moves and *Jeopardy!* responses" (Johnson, "Reply to Victor Reppert," 118; cp. "Con," 97; and "Retiring the Argument from Reason," 562). To be sure, software is realized in hardware, but this analogy breaks down in several ways. First, the computer metaphor suggests a nonreductive relation between software logic and its hardware realizer, as the informational level is distinct from its realizing hardware, so at best this analogy points toward the nonreductive physicalist model discussed in section 4. Second, there is no parallel completeness principle available on the computer metaphor. That is, there is no "hardware causal completeness principle" that says all outputs have sufficient hardware causes. Rather, hardware constantly receives input from the software level and from users, indicating that the hardware alone, without any software influence or user input, is not a sufficient cause of its outputs. Since the problem for physicalism begins when physical causal completeness renders mental causes otiose, and there is no parallel

4. Nonreductive Physicalism and the Argument from Reason

Given the failure of the reductive physicalist version of the dual process reply to overcome the argument from reason, it is worth considering the nonreductive physicalist version. This is especially worthwhile given the fact that the last problem with the reductionist model revolves around the fact that Tran's belief that "Socrates is mortal" only has an arational physical cause and lacks a distinct reason M_1 as a cause. Nonreductive physicalism offers a distinct reason M_1 as a cause of Tran's beliefs, so perhaps nonreductive physicalism avoids the argument from reason. On first glance, it does. Tran's belief M_2 that "Socrates is mortal" is caused by his reason M_1 which necessarily precedes M_2, satisfying the requirement that Tran's belief M_2 be causally based on his reasons M_1, which establishes that Tran knows that "Socrates is mortal."

There are several reasons to think that nonreductive physicalism does not offer an effective dual process reply to the argument from reason either. First of all, nonreductive physicalism is presently taking fire, with the following principle of causal exclusion serving as the ammunition:

> *Causal Exclusion*: "No single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of causal overdetermination."[38]

principle with computer hardware threatening the efficacy of software, the analogy is not relevantly similar. If anything, computers are examples of downward causation, where the software level continuously influences the lower-level hardware. Third, even if the hardware alone caused outputs, this would only be the case because an outside source, namely, humans, engineered the hardware to do so, once again showing that lower-level processes, by themselves, do not deliver their outputs.

38. Jaegwon Kim, *Physicalism, or Something Near Enough* (Princeton, NJ: Princeton University Press, 2005), 42; cp. Rickabaugh and Buras, "The Argument from Reason," 394. For Kim, genuine cases of causal overdetermination occur when two independent causal processes converge on the same effect (Kim, *Physicalism, or Something Near Enough*, 48). A barn that burns down by the simultaneous occurrence of a lightning strike and a lit match dropping on the hay is a case of genuine overdetermination. Nonreductive physicalists agree with Kim that the supervenience relation ensures that the two causes are not independent, so the two causes would not be genuine cases of causal overdetermination, so exclusion pressures apply to nonreductive physicalist mental causation. The two most relevant arguments for the causal exclusion principle are the parsimony argument and the necessity argument. The parsimony argument states that as a general scientific value we should "get by with the fewest possible entities" (Kim "Mechanism, Purpose, and Explanatory Exclusion," *Philosophical Perspectives* 3 (1989): 98), so if one cause is sufficient, we should get by without positing additional causes, so we should exclude additional causes. The necessity argument states if one cause is sufficient for causing an effect (i.e., one cause is all that one needs to cause the effect), then a second cause is not necessary as a cause of the effect (i.e., the second cause is not also needed to cause the effect), so it can be excluded (Kim, "Mechanism, Purpose, and Explanatory Exclusion," 82; and *Mind in a Physical World*, 44–5).

According to the causal exclusion principle, since M_2 has P_1 as its sufficient physical cause, M_1 is excluded from causing M_2.[39] And if M_2 is not caused by M_1, M_2 is not causally based on its reason M_1, so M_2 cannot be known.

There are several replies that the nonreductive physicalist can make to these exclusion pressures, the first two of which are suggested by Johnson.[40] First, Johnson says M_1 is causally relevant to the occurrence of M_2 by virtue of the fact that the supervenience relation entails that a change in M_1 necessitates a change in P_1, which alters P_2, which alters M_2, so M_1 is causally relevant to the occurrence of M_2. This response trades on an outdated and incomplete understanding of the supervenience relation. Consider Donald Davidson's similar suggestion that mental properties are causally relevant to their effects because the supervenience relation shows that changes to mental properties of the cause implies changes to physical properties of the cause, which alters the effect.[41] Davidson's critics reject this notion of causal relevance because it mischaracterizes the determinative direction of the supervenience relation. On supervenience, changes to mental properties only imply changes to phys-

39. This is a simplistic formulation of the difficulty. More thoroughly, M_1 is excluded from causing M_2 because the base P_2 is a sufficient determinant of M_2. Perhaps M_1 causes M_2 via causing P_2? M_1 is excluded from causing P_2 because P_1 is a sufficient cause of P_2. Hence, the simplistic result: M_2 is excluded from causing M_1 because the complete physical process running from P_1 to P_2 to M_2.

40. It is also common for nonreductive physicalists to argue that the causal exclusion principle relies upon a generative model of causation, according to which causes push, pull, or otherwise add causal oomph to their effects. In place of this generative model of causation, some nonreductive physicalists propose alternative models of causation, with the counterfactual analysis and the difference-making analysis being the leading contenders. Barry Loewer ("Comments on Jaegwon Kim's *Mind and the* [*sic*] *Physical World*," *Philosophy and Phenomenological Research* 65 (2002): 655–62) is a leading advocate of solving the causal exclusion problem via a counterfactual analysis. According to this view, *C* is a cause of *E* if the nearest possible world where *C* does not occur is a world where *E* does not occur. Since the nearest possible world where no physical cause *P* occurs is also a world where the mental cause *M* does not occur, and the effect *E* does not occur either, it is the case that both the mental event *M* and some physical event *P* is a cause of *E*, eluding exclusion pressures. Kim rejects this counterfactual analysis, arguing that counterfactual dependency can be present while causation is absent. The bullet hole in the wall is counterfactually dependent on both the gun shot and the sound the gun shot makes, but the former, not the latter, is the real cause (Kim "Causation and Mental Causation," in *Contemporary Debates in Philosophy of Mind*, ed. B. P. McLaughlin and J. D. Cohen (Oxford: Blackwell, 2007), 227ff.). Christian List and Peter Menzies are the most prominent advocates of the difference-making account, according to which *C* is a cause of *E* if the nearest possible worlds where *C* does not occur are also worlds where *E* does not occur, and the nearest possible worlds where *C* occurs are also worlds where *E* occurs (Christian List and Peter Menzies, "Nonreductive Physicalism and the Limits of the Exclusion Principle," *Journal of Philosophy* 106 (2009): 475ff.). One proposal they make is that the nearest worlds where *M* occurs but the physical cause *P* doesn't occur are worlds where some other realizer P_2 occurs, and *E* still occurs. This allows room for the mental cause to be included, but at the expense of making *P* fail to be a cause of *E* (since, the counterfactual "Had *P* not occurred, *E* would not have occurred" is false). Suffice it to say, without diving into the details of these views of causation further, modifying the metaphysics of causation does not straightforwardly dissolve exclusion pressures.

41. Davidson, "Thinking Causes," 14.

ical properties because it is the physical properties that determine the existence and nature of the mental properties, leaving mental properties as accidental by-products of lower-level determinative processes rather than causal contributors. For example, the sound of a gunshot may mistakenly appear causally relevant to the appearance of a hole in the wall because changing the sound implies the bullet was not fired, which implies the hole does not appear in the wall, but the appearance of causal relevance is misleading, since it is the gunshot that determines both the sound and the hole in the wall, leaving the sound as causally irrelevant for the hole in the wall.[42]

Johnson also suggests the supervenience relation renders Tran's reasons M_1 necessary as causes for his belief that "Socrates is mortal" since P_1 metaphysically necessitates his reasons M_1, so Tran's reasons must precede his beliefs. This response suffers from several objections. First, that P_1 metaphysically necessitates M_1 only implies that M_1 is necessarily present prior to M_2 occurring, it does not imply that M_1 is a cause of M_2. There are plenty of examples of metaphysically necessary supervenient epiphenomenal properties. Imelda brings two dollars to the store with the intent of buying all the apples she can. Apples are fifty cents each, so Imelda buys four apples, but the apple's being fifty cents metaphysically necessitates the apple's being less than sixty million dollars, though this does not cause Imelda to buy the four apples. Or, to use another example, Sophie the pigeon is trained to peck at red things. She sees a red triangle and so she pecks at it, but the redness of the triangle also metaphysically necessitates that the triangle is red-or-infrared, though she does not peck in virtue of this disjunctive property.[43]

Not only can Tran's reasons be necessarily present without being causes of his belief, there is reason to think that Tran's reasons would be necessarily present but not be causes of M_2. The reason is the familiar causal exclusion problem yoked together with physical causal completeness: the physical pro-

42. See, e.g., Ernest Sosa, "Davidson's Thinking Causes," in *Mental Causation*; and Brian McLaughlin, "On Davidson's Response to the Charge of Epiphenomenalism," in *Mental Causation*, 27–40.

43. Victor Reppert uses an unfortunate example to make this point. He considers a baseball that breaks a window, where this baseball was once hit by Luis Gonzales (Reppert, "Reply to David Kyle Johnson," in *C. S. Lewis's Christian Apologetics*, 109–10). Reppert concludes that, as the baseball's being hit by Luis Gonzales is not causally relevant to the smashed window, so mental properties are not causally relevant to our beliefs. Johnson correctly responds that the mass of the baseball does not metaphysically subvene the property of once being hit by Luis Gonzales, so the analogy is not relevant to the case where physical properties metaphysically subvene mental properties (Johnson, "Retiring the Argument from Reason," 560–1). However, the metaphysical supervenience of M_1 on P_1 still does not establish that M_1 is causally relevant to some future effect. To return to the baseball analogy, the baseball's mass metaphysically necessitates that the baseball is heavier than one milligram, and that the ball exists within the spatiotemporal universe, and that the ball is the thing that it is, and that the ball has mass or does not have mass, but none of these properties are causally relevant to the window breaking. The reason is that P_1 metaphysically necessitates the presence of M_1, but this says nothing about whether or not M_1 causally contributes to future effects.

cess from P_1 to P_2 to M_2 completely determines that M_2 occurs, so there is no work left for M_1 to do, so it can be excluded from causing M_2. Some nonreductive physicalists reject the causal exclusion principle here, arguing that Tran's belief that "Socrates is mortal" has both a sufficient physical cause and a distinct mental cause, thereby overdetermining Tran's belief that "Socrates is mortal." Besides the difficulties associated with abandoning exclusion, this move is ill-advised because it leads to the failure of physical causal completeness. Physical causal completeness is the view that P_1 can cause Tran's beliefs all by itself. If P_1 and M_1 both cause Tran's beliefs, then P_1 does not cause Tran's beliefs all by itself. Moreover, it is metaphysically impossible for P_1 to cause Tran's beliefs all by itself, given that M_1 is necessarily present as a cause of Tran's beliefs—P_1 cannot be all the causation needed for Tran's beliefs while M_1 is also needed as a cause of Tran's beliefs. The nonreductive physicalist gains M_1 as a cause of Tran's beliefs only by sacrificing P_1 as the complete cause of Tran's beliefs, a steep price that nonreductive physicalists are not likely willing to pay.[44]

The nonreductive physicalist also faces the coincidental correlation problem. Since the physical process from P_1 completely determines Tran's belief that "Socrates is mortal," his belief that "Socrates is mortal" could still occur, regardless of how his mental life unfolded. If Tran believed "mice like cheese" and "pineapples are sweet," he could still conclude that "Socrates is mortal," since this conclusion is completely determined to occur by P_1. Most nonreductive physicalists, including Johnson, will immediately protest here: the fact that P_1 necessitates M_1 guarantees that the appropriate reason state M_1 occurs prior to Tran's belief that "Socrates is mortal." This response fails, for two reasons. First, if needed, we can advert to the impossible worlds to regenerate the difficulty for nonreductive physicalists.[45] In the nearest impossible world where P_1 occurs, but gives rise to the belief that mice like cheese instead of the belief that Socrates is human, Tran would still believe that Socrates is mortal, revealing that Tran's belief is not based on his reasons.[46] Secondly, this discovery points to a larger difficulty for nonreductive physicalism. Namely, given that Tran's belief that "Socrates is mortal" is completely determined by the arational process from P_1 to P_2, Tran would hold this belief no matter what his prior beliefs are—he could (counterpossibly) believe that "mice like cheese" or that "Socrates is human," yet that does not change the effect. Despite the fact that he could (counterpossibly) believe that "mice like cheese," he ends up believing the exactly appropriate belief that "Socrates

44. See Dwayne Moore, "Mental Causation, Compatibilism, and Counterfactuals," *Canadian Journal of Philosophy* 47 (2017): 36, for further elaboration on this issue.

45. See, e.g., Jens Bjerring, "On Counterpossibles," *Philosophical Studies* 168 (2014): 327–53; and Berit Brogaard and Joe Salerno, "Remarks on Counterpossibles," *Synthese* 190 (2013): 639–60.

46. For a similar view, see Alvin Plantinga, *Where the Conflict Really Lies* (New York: Oxford University Press, 2011), 334–9.

is human," which corresponds precisely to the conclusion that "Socrates is mortal"—what an amazing coincidence.[47] Given that Tran's beliefs arise from P_1, regardless of whether he has the appropriate beliefs or not, why would he have the exactly appropriate beliefs instead of one of billions of other beliefs?[48]

5. Explanatory Dualism and the Argument from Reason

Explanatory dualism is the final version of the dual process reply to the argument from reason. The explanatory dualist version posits either a nonreductive physicalist or reductive physicalist metaphysic, but then replies to the argument from reason by positing both a rationalizing explanation and a physical explanation of the same event. Thus, Tran's belief that "Socrates is mortal" has both a rational explanation in terms of his reasons for this belief and a complete physical explanation in terms of the neural processes causing him to hold this belief.[49]

This version departs from the realm of metaphysics toward the realm of explanation—a lamentable move that should be retired for several reasons. First, it ultimately generates no novel solutions but only repeats the argumentation already present at the level of causation. Second, it introduces needless complications pertaining to the nature of explanations that are not germane. Third, and most importantly, it has already been established that knowledge

47. For a similar argument, see Stewart Goetz, "The Argument from Reason," *Philosophia Christi* 15 (2013): 61.

48. This coincidental correlation problem is sometimes leveled against qualia epiphenomenalism: given that mental states do not cause physical effects, it is surprisingly coincidental that mental states are exactly commensurate with physical effects (Daniel Hutto, "A Cause for Concern," *Philosophy and Phenomenological Research* 59 (1999): 389; Michael Pauen, "Feeling Causes," *Journal of Consciousness Studies* 13 (2006): 150; Dwayne Moore, "Physical Effect Epiphenomenalism and Common Underlying Causes," *Dialogue* 51 (2012): 414–15). It is easy to level this objection against epiphenomenalism because epiphenomenalists believe that the supervenience relation from P_1 to M_1 is not metaphysically necessary, so it is possible for P_1 to subvene any other mental state. Nonreductive physicalists claim immunity to this problem since they endorse the metaphysically necessary supervenience relation, so it is not possible for P_1 to subvene any other mental state. The problem still applies to nonreductive physicalists, for the two reasons raised above. The typical physicalist explanation for the appropriate mental correlations is that P_1 necessitates M_1 and causes physical effects of type P_2 to determine M_2, resulting in stable, predictable macro-generalizations from M_1 to M_2 (Jaegwon Kim, "Why There Are No Laws in the Special Sciences," in *Essays in the Metaphysics of Mind* (New York: Oxford University Press, 2010), 300). But the question is not Why are there stable, predictable macrogeneralizations? The question is Why are there appropriately correlated stable, predictable macrogeneralizations? Why does P_1 predictably necessitate the appropriate M_1 rather than predictably necessitating some inappropriate belief that "mice like cheese," given that the conclusion M_2 is going to occur once P_1 occurs, no matter what prior beliefs occur.

49. Advocates of this explanatory dualist solution include: Drange, "Several Unsuccessful Formulations of the Argument from Reason," 49; Parsons, "Further Reflections on the Argument from Reason," 97–8; van Inwagen, "C. S. Lewis' Argument against Naturalism," 119–23; Johnson, "Retiring the Argument from Reason," 559–60.

requires causal basing, not simply reasons explanations, so establishing that Tran's belief has both a reasons explanation and a physical explanation would still not suffice to deliver the required causal component to knowledge.

Even if explanatory dualism is a viable dual process reply, there are still reasons to think it fails as a response to the argument from reason. It is common to respond to this explanatory dualism by appealing to a parallel principle of exclusion in the explanatory realm. For example, Reppert says: "in arguing that the mentalistic account of Gödel's mental activity and the physicalistic one cannot both be true, I am applying Jaegwon Kim's principle of explanatory exclusion."[50] Similarly, Rickabaugh and Buras claim: "For any belief *b*, if *b* has a complete mechanistic explanation, then *b* cannot also have any rational explanation."[51] Tran's belief has a complete physical explanation, so we can exclude the reasons explanation, and without the reasons explanation, Tran's beliefs do not have reasons explaining the belief, so is not knowledge.

Johnson points out a difficulty with this initial response. Namely, Kim's principle of explanatory exclusion technically allows for multiple explanations of the same event, so long as a dependency relation between the explanations is established. Here is Kim's principle of explanatory exclusion:

> *Explanatory Exclusion*: "There can be no more than a single complete and independent explanation for any one event."[52]

As Johnson notes, Kim claims that either an identity relation or a supervenience relation can establish ontological dependency relations among two explanations. Thus, on both reductive physicalism and nonreductive physicalism, the reasons explanation for Tran's belief that "Socrates is mortal" is dependent upon the physical explanation for Tran's belief (since Tran's reasons are identical with or supervenient upon the physical processes P_1 in his brain). So, the reasons explanation is not in danger of exclusion, as there can be multiple dependent explanations of Tran's belief. As Johnson concludes, "human minds are dependent upon their physical substrate: the brain. . . . Mental explanations for human behaviour are thus dependent upon their neuronal explanations. Therefore, by the Explanatory Exclusion Principle, both are correct and ultimate."[53]

Rather than legitimizing the explanatory model of the dual process reply, Johnson's response merely highlights both a historical and conceptual inconsistency in Kim's version of explanatory exclusion. Kim's principle of explanatory exclusion, which allows dependent explanations to be included as explanations of events, was forged during a time when he also allowed de-

50. Reppert, "Extending the Debate on the Argument from Reason: A Further Response to David K. Johnson," *Philosophia Christi* 20 (2018): 530.

51. Rickabaugh and Buras, "The Argument from Reason," 388; cp. Goetz, "The Argument from Reason," 61–2.

52. Kim, "Explanatory Realism, Causal Realism, and Explanatory Exclusion," *Midwest Studies in Philosophy* 12 (1988): 233.

53. Johnson, "Retiring the Argument from Reason," 559.

pendent supervenient causes to be included as causes of events. Before long Kim emphasized his now famous causal exclusion principle, which began excluding dependent supervenient causes of events, and in so doing he rejects his earlier attempts at including dependent causes. He realized his earlier attempts at including dependent supervenient causes was "an empty verbal ploy . . . a gimmick with no meaning."[54]

Using the same reasoning, the inclusion of additional dependent explanations beyond an already complete explanation is also a gimmick with no meaning. If the physical explanation is, in principle, complete, then dependent mental explanations are not ultimately necessary. As Kim modified his thinking to exclude dependent mental causes, so a similar amendment is warranted for excluding dependent mental explanations. Here is an amended principle of explanatory exclusion:

> *Revised Explanatory Exclusion*: There can be no more than a single complete explanation of any one event.[55]

Why is this amendment warranted for explanatory exclusion? As the principle of causal exclusion is supported by, among other things, the parsimony argument and necessity argument (see note 38), so this amended version of explanatory exclusion is supported by a parsimony argument and necessity argument.[56] The principle of parsimony excludes any cause or explanation

54. Kim, *Physicalism, or Something Near Enough*, 62; cp. Kim, *Mind in a Physical World*, 37. Kim's principle of explanatory exclusion is developed in "Explanatory Realism, Causal Realism, and Explanatory Exclusion" and "Mechanism, Purpose, and Explanatory Exclusion," published in 1988 and 1989, respectively. Kim's sympathy for including dependent supervenient causes during this time is evident in his *Supervenience and Mind* (New York: Cambridge University Press, 1993), 106–7, originally published in 1984.

55. Dwayne Moore, "Explanatory Exclusion and Mental Explanation," *Philosophical Psychology* 29 (2016): 390–404, defends a similar principle in greater detail. Indeed, a number of authors highlight the similarities that causal exclusion has with explanatory exclusion; see Andre Fuhrmann, "Causal Exclusion without Explanatory Exclusion," *Manuscrito* 25 (2002): 182; and Sophie Gibb, "Explanatory Exclusion and Causal Exclusion," *Erkenntnis* 71 (2009): 5. Fuhrmann, e.g., says that Kim thinks that "there is essentially only one exclusion problem which may both be cast in terms of explanation as well as in terms of causation" ("Causal Exclusion without Explanatory Exclusion," 184). Kim even goes so far as to blend his two principles together under the heading of causal/explanatory exclusion at times; see *Supervenience and Mind*, 281, 291; and "Mechanism, Purpose, and Explanatory Exclusion," 44.

56. It is possible to object that the principle of parsimony is properly applied to metaphysical entities, such as events and objects (i.e., ontological simplicity), but is improperly applied to theoretic entities, such as explanations and hypotheses (i.e., syntactic simplicity). In response, parsimony is for Kim and most others a general principle that applies equally to theoretical posits. After all, parsimony, in Ockham's originative words, is a generic principle of simplicity: "Plurality must never be posited without necessity" (William Ockham, *Quaestiones et Decisiones in Quattuor Libros Sententiarum Petri Lombardi* (1495), 1.27.2). Within the contemporary sphere, parsimonious theories are applauded because of both their grammatical elegance and their increased likelihood of being true (given they have fewer propositions required to be true); see Michael Friedman, "Explanation and Scientific Understanding," *Journal of Philosophy* 71 (1974): 5–19; and Jaegwon Kim, "Explanatory Knowledge and Metaphysical Dependence," *Philosophical Issues* 5 (1994): 51–69.

that is not necessary. Once an event already has a complete explanation, additional explanations, no matter if they are dependent explanations, are not in principle necessary, so can theoretically be excluded. To return to the previous examples, the complete explanation for why Imelda bought four apples is because she had two dollars to spend. If her grandchild prattles on and continues to explain that Imelda bought four apples because she had two-dollars-or-less-than-sixty-million-dollars to spend, and because she had two-dollars-or-less-than-fifty-million-dollars to spend, these are not in principle necessary explanations, so can theoretically be excluded, even though they are dependent on the first explanation. The complete explanation for why Sophie pecks at the red triangle is that she was trained to peck at red things. We need not add: "Also, Sophie pecks at things that are red-or-infrared, which explains why she pecks at the red triangle." If we allowed multiple dependent explanations beyond the complete explanation, then there could be no end to explanations: Sophie also pecks because the triangle is red-or-colored, and because the triangle is red-or-brown-or-green. The principle of parsimony counsels us that we can exclude these surplus explanations.[57]

With respect to Tran's belief that "Socrates is mortal," physical causal completeness has a parallel principle in the explanatory realm called physical explanatory completeness, according to which all events have a complete physical explanation. As Kim says: "when the causal relation provides a sufficient cause, the explanans can also be said to be complete and sufficient."[58] So,

57. There are several reasons why it is common to accept multiple explanations of the same effect. First, explanations provide epistemic illumination, and multiplicities of explanations may be necessary to provide epistemic illumination to different people in different contexts with different background knowledge. To borrow an example from Fuhrmann ("Causal Exclusion without Explanatory Exclusion," 184), those lacking the requisite background may not gain epistemic illumination from the explanation stating that "The baby in the tub cried because the mean kinetic energy of the H_2O molecules passed a certain threshold." Thus, the explanation stating that "The baby in the tub cried because the water was too hot" may be epistemically necessary, so not excludable. While additional explanations may provide illumination, they are not *in principle* necessary. If a chemist were present, the first explanation would have sufficed, and the second explanation would not have been necessary, so could have been left out. Second, there seems to be no problem with describing the same event in multiple different ways: the red rose can also be called "the rose that is Sophie's favorite color." As Arnodottir and Crane explain, "Any occurrence can be explained in countless ways, and there is no incompatibility between any physical explanation of an event and a mental explanation of the same event. So there seems to be little plausibility to the idea that one explanation 'excludes' another" (Steinvor Arnodottir and Tim Crane, "There Is No Exclusion Problem," in *Mental Causation and Ontology*, ed. E. J. Lowe, Sophie Gibb, and R. D. Ingthorsson (New York: Oxford University Press, 2013), 256). This is, of course, true. It is *possible* to explain the same event in many different ways, but this does not imply that we *need* to explain the same event in many different ways, so it is still true that there *can be* (i.e., it is possible for there to be) no more than a single complete explanation. The fact that mental explanations are theoretically excludable—not that we must exclude them, but that we could exclude them—is the problem.

58. Kim, "Mechanism, Purpose, and Explanatory Exclusion," 234. Or, as David Papineau frames it: "prior physical factors will always suffice to give us as full an explanation of that result as is possible" (Papineau, "Arguments for Supervenience and Physical Realization," in *Super-*

Tran's belief that "Socrates is mortal" has a complete physical explanation in terms of its physical causes and determinants. Since Tran's belief has a complete physical explanation, additional mental explanations are not in principle necessary, even if they are dependent explanations, so can theoretically be excluded. The result is that the reasons explanation citing Tran's reasons for believing that Socrates is mortal can be excluded as an explanation for his belief that Socrates is mortal. Since the reasons explanation can be excluded, the explanatory version of the dual process reply fails—there is no longer a complete physical explanation and a distinct reasons explanation of Tran's belief that "Socrates is mortal."[59]

Johnson sometimes suggests solving the threat of explanatory exclusion by invoking an extensional model of explanatory individuation. On this model, the mental explanation states the same explanation as the physical explanation by virtue of the fact that they explain the same event. Thus, there is only one explanation of the event, and explanatory exclusion pressures do not arise when there is only one explanation of the event. He says, "each instance of a mental event is identical to some instance of a brain event; thus, on both theories, every explanation in terms of a brain event is, *ipso facto*, an explanation in terms of a mental event."[60]

While Kim also considers such a model at times, it is dismissed by critics because it incorrectly implies that two epistemically inequivalent descriptions can state the same explanation.[61] Ausonio Marras, for example, points out that "the earthquake caused the collapse of the building" and "the event that caused the collapse of the building caused the collapse of the building" cannot state the same explanation, as one is epistemically informative and the other is epistemically vacuous. On the extensional model, however, "anyone in possession of the one explanation is thereby in possession of the other.

venience: New Essays, ed. E. E. Savellos and U. Yalcin (New York: Cambridge University Press, 1995), 228–9; cp. Gibb, "Explanatory Exclusion and Causal Exclusion," 209; Fuhrmann, "Causal Exclusion without Explanatory Exclusion," 186).

59. Reppert and Johnson sometimes discuss these issues by introducing the term "ultimate explanation." For Reppert an ultimate explanation is a nonmental, physical explanation that completely explains the event (Reppert, "Pro," 77–8; and "Reply to David Kyle Johnson," 107). Reppert says mental explanations of actions are "washed out" (Reppert, "Pro," 78) by these ultimate physical explanations. Johnson takes mental events to be identical with or to supervene upon physical events, so physical explanations and mental explanations are both ultimate explanations of those events (Johnson, "Con," 95; and "Reply to Victor Reppert," 115). But the problems physicalists face can be formulated without invoking "ultimate explanations," so I bypass this lacuna in the discussion. The fundamental problem is that the physical explanation of an event is a complete explanation of the event (by physical explanatory completeness), and there can be (i.e., it is possible for there to be) no more than a single complete explanation of the event (by revised explanatory exclusion), so the mental explanation of the event can be excluded as an explanation of the event.

60. Johnson, "Reply to Victor Reppert," 117.

61. See, e.g., Kim, "Explanatory Realism, Causal Realism, and Explanatory Exclusion," 226–33.

And this, I believe, is counter-intuitive."[62] Solving the explanatory exclusion problem by appeal to an extensional model of explanatory individuation, therefore, is not viable. Perhaps this is why Johnson usually prefers to treat the mental explanation and physical explanation as two distinct but dependent explanations.[63] But we have seen that this does not work either.

In conclusion, knowledge requires not only true beliefs to be justified by reasons, but also that true beliefs be causally based on those reasons. In order for a physicalist to know physicalism is true, the belief that physicalism is true must not only be true, and be supported by reasons, but the belief must also be caused by those reasons. Physicalists have difficulty demonstrating how the belief is caused by reasons, given that the belief is completely caused by arational physical processes. The two most common physicalist models—reductive physicalism and nonreductive physicalism—both face difficult obstacles in showing how reasons cause beliefs. Reductive physicalism faces the quasation problem and the only physical causation problem. Nonreductive physicalism faces the causal exclusion problem and the coincidental correlation problem. Appealing to reasons explanations of beliefs that already have complete physical explanations does not help either. Without any viable models of the dual process reply remaining, physicalism remains in a self-stultifying position.

62. Ausonio Marras, "Kim's Principle of Explanatory Exclusion," *Australasian Journal of Philosophy* 76 (1998): 443.

63. See, e.g., Johnson, "Con," 93; and "Retiring the Argument from Reason," 559.