Journal of the American Philosophical Association (2016) 310–331 © American Philosophical Association
DOI: 10.1017/apa.2016.7

Free Will and Mental Quausation

ABSTRACT: *The questions of how to understand free will and mental causation are clearly connected, for events of seemingly free choosing are mental events that appear to be efficacious vis-à-vis other events. Nonetheless, the free will and mental causation debates have proceeded largely independently of each other. Here we aim to make progress in determining the mutual bearing of these debates. We first argue that the problems of free will and of mental causation can be seen as special cases of a more general problem of mental 'quausation', concerning whether and how mental events of a given type can be efficacious qua the types they are—qualitative, intentional, freely deliberative—given reasons to think such events are causally irrelevant. We go on to identify parallels between hard determinism and eliminativist physicalism and between soft determinism and nonreductive physicalism, and we use these parallels to identify a new argument against hard determinism and to reveal and motivate a common strategy underlying apparently diverse soft determinist accounts.*

KEYWORDS: ontology, metaphysics, metametaphysics, philosophy of mind, free will, agency, consciousness, physicalism, mental quausation

Introduction

Free will, if such there be, involves free choosing: the ability to choose mentally an outcome (an intention to ϕ, or a ϕ-ing), where the outcome is 'free' in being, in some substantive sense, up to the agent of the choice. As such, it is clear that the questions of how to understand free will and mental causation are connected, for events of seemingly free choosing are mental events that appear to be efficacious vis-à-vis other mental events (e.g., intentions to pet the cat) as well as physical events (e.g., cat pettings). Nonetheless, the debates about free will and mental causation have proceeded largely independently of each other. Discussions of free will have rarely addressed whether and how the prevailing accounts of mental causation might make room for free choosing. And discussions of mental causation have neglected the efficacy of events of free choosing, focusing almost exclusively on the efficacy of qualitative and intentional mental events (pains and color experiences, beliefs and desires) for which freedom is not at issue.

Special thanks to Michael McKenna, Peter van Inwagen, and two referees for this journal, and thanks as well to Rani Anjum, Benj Hellie, Stephen Mumford, David Sosa, and to audience members at the 2013 Norwegian University of Life Sciences conference on the Metaphysics of Free Will, the University of Arizona, the University of Connecticut, and Davidson College.

In this paper, we aim to make progress in determining how the free will and mental causation debates bear on one another. In section 1, we argue that (certain understandings of) the problems of free will and of mental causation can be seen as special cases of a more general problem, concerning whether and how mental events of a given type can be efficacious *qua* the types of events they are—qualitative, intentional, freely deliberative—given their apparent causal irrelevancy for effects of the type in question. Here we generalize what Horgan (1989) identifies as 'the problem of mental quausation'. We then build on this result to identify fruitful parallels between certain positions in each debate. In particular, we argue that there is a parallel between hard determinism and eliminativist physicalism, attention to which suggests a better argument against hard determinism than that typically offered (section 2), and we argue that there is a parallel between soft determinism and nonreductive physicalism, attention to which reveals that apparently diverse soft determinist accounts, like apparently diverse nonreductive physicalist accounts, are implementing a structurally similar strategy for accommodating the causal relevance of the mental events at issue (section 3).

1. Problems and Parallels

1.1 The Problems

Given what's at stake, it would be nice if we could give accounts of free will and of mental causation that accommodate our intuitive beliefs about these phenomena—in particular, that the associated mental events exist and have distinctive causal roles to play—as well as any scientific hypotheses we are not in a position to rule out. But in each case there seems to be a problem with doing so.

The problem with accommodating the intuitive conception of free will has traditionally been posed in terms of the broadly scientific thesis of *Determinism*:

> *Determinism*: Every event is a consequence of the laws of nature and the state of the world at any time.

Determinism admits of interpretations based on entailment between propositions at one time by propositions at any other time and interpretations based on causation between events (see Steward 2015). We are here concerned with a causal reading:

> *Causal Determinism*: Every event (apart from any first events; henceforth we assume this restriction) is a causal consequence of the laws of nature and the state of the world at any prior time.

One might wonder whether *Causal Determinism* can be ruled out, on grounds that quantum mechanics is indeterministic; this would be premature, however, both because quantum mechanics has deterministic interpretations and because its present incompatibility with general relativity suggests that our fundamental theories are still in flux. The possible truth of *Causal Determinism* leads to a

question: If every event *E* (e.g., an intention to ϕ or a ϕ-ing) purportedly caused by a mental event of free choosing *M* is, by *Causal Determinism*, a causal consequence of the laws of nature and prior states, what causal role is left for *M* to play vis-à-vis *E*? The answer, according to (a causal reading of) the consequence argument, is that no role is left:

> If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born; and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us. (van Inwagen 1983: 16)

(Van Inwagen's remarks are directed at an entailment-based interpretation of *Determinism*, but they apply, *mutatis mutandis*, to a causal interpretation: simply insert 'causal' before 'consequences'.) The question and prima facie negative answer constitute what van Inwagen calls 'the problem of free will and determinism', and what we will call, for short, 'the problem of free will'.

The problem with accommodating the intuitive conception of mental causation has traditionally been posed in terms of the broadly scientific hypothesis of *Physical Causal Closure*:

> *Physical Causal Closure*: Every physical effect has a purely physical cause.

The 'physical' entities are typically understood as the (individual or aggregatively complex lower-level) entities that are the target subject matter of physics (see Wilson 2006 and Ney 2008 for discussion of further nuances). Like *Causal Determinism*, *Physical Causal Closure* remains an empirically live possibility. The possible truth of *Physical Causal Closure* gives rise to a question: If every event *E* (e.g., a desire or a bodily movement) purportedly caused by a qualitative or intentional mental event *M* (e.g., a pain or a belief) is, given *Physical Causal Closure*, the effect of a purely physical cause, what causal role is left for *M* to play vis-à-vis *E*? The answer, according to the exclusion argument, is that no role is left. (Note that notwithstanding that events of free choosing are presumably intentional and may also be qualitative, the intentional and qualitative events at issue in the problem of mental causation are standardly those—pains, beliefs, desires—for which free choosing is supposed not to be issue.) One starts by formulating the problem for the case where the effect at issue is itself a physical event:

> Suppose that a certain [mental] event . . . causes a physical event. The causal closure of the physical domain says that this physical event must also have a physical cause. . . . What is the relationship between the two causes, one mental and the other physical? [It might be suggested that] the mental cause and the physical cause are each only a partial cause, and that they together make up one sufficient cause. This surely

> is an absurd thing to say, and in any case it violates the causal closure principle.... Could it be that the mental cause and the physical cause are each an independent sufficient cause of the physical effect? The suggestion is that the physical effect is overdetermined. So if the physical cause hadn't occurred, the mental cause by itself would have caused the effect. This picture is again absurd [given the known broadly synchronic dependence of the mental on the physical, and] violates the causal closure principle as well: in the counterfactual situation in which the physical cause does not occur, the closure principle is violated.... These two ways of looking at the situation are obvious non-starters.... Given that any physical event has a physical cause, how is a mental cause also possible? (Kim 1989: 280–81)

One next observes that *Physical Causal Closure* also threatens to result in the causal exclusion of any mental effects purportedly caused by mental events. As above, mental events clearly broadly synchronically depend on physical events, at least in that any given mental effect M^* will be nomologically necessitated, in the circumstances, by some physical event P^*. By *Physical Causal Closure*, P^* will have some purely physical cause P. But then insofar as P causes P^*, which brings about M^*, it seems reasonable to see P as causing M^*. In that case, what role is there for any mental event M to cause M^*? Here again, the answer seems to be: none.

The general question, and the prima facie negative answer(s), constitute what Kim calls 'the problem of causal-explanatory exclusion', and what we will call, for short, 'the problem of mental causation'.

1.2 The Generalized Problem of Mental Quausation

In our view, a common theme underlies the problem of free will and the problem of mental causation. More specifically, we see these problems as each being instances of a suitably general version of what we call, following Horgan (1989), 'the problem of mental quausation'.

The problem of mental quausation is, in the first instance, a refinement of the problem of mental causation. The latter problem is sometimes pitched as the problem of how a qualitative or intentional mental event M might be efficacious *at all*, given that any effect E it might be seen as causing is, by *Physical Causal Closure*, already caused by a physical event P. But as Horgan notes, there is a quick route to gaining such efficacy—namely, by identifying M with P, as per the usual reductive physicalist response—that leaves open what is arguably the deeper question underlying the original problem: 'even if individual mental events and states are causally efficacious, are they efficacious *qua* mental?' (1989: 47). As Shoemaker (2000–2001) puts it:

> If what is in fact an instance of a mental property causes something... but does so in virtue of being an instance of a physical property rather than in virtue of being an instance of a mental property,

> then the causal efficacy of the mental does not seem to have been adequately vindicated. (28)

What is needed to vindicate adequately the efficacy of the mental, Horgan suggests, is that the mental be shown to be efficacious *qua* mental—and more specifically, '"*qua F*" where *F* is schematic for a specific type of mental event' (1989, 50). As such, it is mental 'quausation', not mental causation per se, that is most deeply challenged by *Physical Causal Closure*. To wit: how can a mental event *M* be efficacious vis-à-vis an effect *E* in virtue of being qualitative or intentional, given that *E* is causally determined by physical events and associated laws in ways that seem to preclude *M*'s being causally relevant in either qualitative or intentional respects?

Though Horgan's discussion, like that of Shoemaker and others, focuses on mental events for which freedom is not at issue, the deeper concern about whether and how mental events can be causally relevant—that is, efficacious in virtue of or *qua* the distinctive mental types they are—applies also to the events at issue in the problem of free will. In the problem of free will, it is not the mere efficacy of deliberative mental events (as, perhaps, the most proximate causes of intentions or actions in a causally determined chain of events) that is at issue; indeed, that events of choosing are efficacious is typically taken for granted in the debates about free will. Rather, what is at issue is whether events of choosing can be efficacious *qua* free—again, in eventuating in outcomes that are, in some substantive sense, up to the agent—under circumstances where *Causal Determinism* is presumed to be true. More specifically: how can an event of mental choosing *M* be efficacious vis-à-vis an event *E* in virtue of being *free*, given that *E* is causally determined by laws of nature and events antecedent to *M* in ways that seem to preclude *M*'s being causally relevant in this respect?

The parallel here suggests that the problem posed by *Physical Causal Closure* for (nonfree) qualitative and intentional mental events, and the problem posed by *Causal Determinism* for events of seemingly free choosing, are each instances of a suitably general problem of mental quausation:

> *The General Problem of Mental Quausation*. How can a mental event *M* of a given type be efficacious vis-à-vis an event *E* in virtue of being the type of mental event it is, given that there is reason to think that events of *M*'s type are causally irrelevant to the production of events of *E*'s type?

We emphasize that we are not suggesting that there is any interesting parallel between the problem of free will and the *unrefined* problem of mental causation; we are also not suggesting that the problem of free will is an instance of the problem of mental quausation as applied to nonfree qualitative or intentional mental events (as a referee pointed out, there is no clear tension between *Causal Determinism* and mental quausation involving mental causes for which freedom is not at issue). Rather, we are suggesting that both the problem of free will and the refined problem

of mental causation (as directed at nonfree qualitative and intentional mental events) are instances of a suitably general problem of mental quausation, whereby the causal relevance of a mental event of a distinctive (intentional, qualitative, freely deliberative) type is called into question. We also emphasize that our reading of the problem of free will is intended as a basis for fruitful comparison, not as a reconstruction or replacement for every expression of felt tension between *Determinism* (casual or otherwise) and free will.

1.3 Parallels: The Responses

The traditional responses to the problem of free will may be categorized by reference to the following conditional:

> *The Free Will Conditional*. If all events are subsumed by deterministic natural laws, then free mental quausation—the causation of events (e.g., intentions to ϕ, ϕ-ings) by mental choosings *qua* free—does not exist.

Hard determinists maintain that both antecedent and consequent are true; soft determinists maintain that the antecedent is true and the consequent is false; libertarians maintain that both antecedent and consequent are false. The soft determinists at issue here and throughout are what might be called 'positive' compatibilists: compatibilists who aim not just to undermine arguments for incompatibility but to provide a positive conception of the compatibility at issue (see McKenna and Coates 2008 for discussion).

Similarly, the traditional responses to the refined problem of mental causation can be categorized by reference to the following conditional:

> *The Mental Quausation Conditional*. If all physical events are subsumed by physical laws, then qualitative and intentional mental quausation—the causation of physical events (e.g., bodily movements) by mental events *qua* qualitative or intentional—does not exist.

Eliminativist physicalists and epiphenomenalists maintain that both antecedent and consequent are true; reductive and nonreductive physicalists maintain that the antecedent is true and the consequent is false; emergentists and substance dualists maintain that both antecedent and consequent are false.

Since the problem of free will and the refined problem of mental causation may each be seen as special cases of a more general problem, we might expect there to be parallels between the primary responses to each problem, corresponding to parallels in the stances taken toward the components of the corresponding conditional, and we might hope that these parallels would shed light, one way or another, on the viability and/or content of the corresponding positions. Indeed, this is the case (though some whittling down needs to occur). In the next two sections we argue that hard determinism is parallel to eliminative physicalism, soft determinism is parallel to nonreductive physicalism, and that in each case the parallel bears fruitfully on

the debate about free will. Since libertarianism (on the one hand) and emergentism and substance dualism (on the other) introduce special considerations pertaining to singularism about causation, naturalism, agent causation, and the like, we leave treatment of parallels between these views for another occasion.

2. Hard Determinism and Eliminative Physicalism

Hard determinists (e.g., Holbach 1770) take the same stance toward the *Free Will Conditional* that eliminativist physicalists and epiphenomenalists take toward the *Mental Quausation Conditional*—namely, they assume both that the problematic thesis (*Causal Determinism*, *Physical Causal Closure)* is true and that this fact has negative existential consequences for there being mental quausation of the variety at issue. Indeed, eliminativists and epiphenomenalists go further in assuming that qualitative mental events, in particular, are not at all efficacious (much less efficacious *qua* qualitative). Which, if either, of eliminativism and epiphenomenalism is more parallel to hard determinism?

The answer reflects a dimension of variation in the debate about mental causation that is not present in the debate about free will, concerning whether the efficacy of the mental events at issue is necessary for their existence. Most participants in the debate about mental causation think so, but epiphenomenalists disagree; hence, while they agree with eliminativists in denying the efficacy of (in particular, qualitative) mental events, epiphenomenalists maintain that such mental events nevertheless exist. It seems clear, however, that the existence of acts of mental choosing depends on their being at least potentially efficacious. Even if some mental deliberations do not result in an outcome—perhaps the agent remains undecided—a choosing must be capable of producing an effect: namely, a choice. A mental event in principle incapable of producing any effect would not be a choosing (much less a free choosing), though it might be some other kind of mental event; hence, epiphenomenalism has no clear parallel in the problem of free will.

By way of contrast, hard determinism and eliminativist physicalism are clearly parallel in denying the existence and hence the efficacy (of whatever sort) of the mental events at issue. This parallel is useful in suggesting an objection to hard determinism that seems not to have received any play in the literature and that is also better than one common objection to this position. This common objection charges that hard determinism would be extremely difficult, perhaps even impossible, to put into practice. As Kane (2002: 27) puts it, 'Few thinkers have been willing to embrace [hard determinism] unqualifiedly, since it would require wholesale changes in the way we think about human relations and attitudes, how we treat criminals and criminal behavior, and so on'. This line of thought is not especially compelling, however. If there is no free will, then the fact that it would be onerous to extricate this notion from our understandings of self and others is ultimately orthogonal to whether hard determinism should be believed—or, more to the point, to whether it is true.

An objection to eliminativist physicalism suggests a better objection to hard determinism. Eliminativist physicalists maintain, as per *Physical Causal Closure*,

that every physical event is subsumed by purely physical laws; they moreover maintain that if we cannot explain mental events in physical terms—if an insuperable explanatory gap or seeming incompatibility exists between, for example, qualitative mental events (laws, theories) and physical events (laws, theories)—then the existence of mental events, and thus mental causation, is thereby ruled out (see, e.g., Feyerabend 1963; Churchland 1981; and Churchland 1986). Finally, eliminativist physicalists maintain that there are in-principle barriers to explaining these events in physical terms, thus warranting the eliminativist conclusion. One objection to such eliminativism, simple but powerful, is that we have more reason to believe that mental events and mental causation (and moreover, mental quausation) exist—primarily as a result of our seeming pervasive experience of such—than we have to believe the (highly speculative, theoretical) premises in the eliminativists' arguments to the contrary (for arguments in the ballpark of this objection, see, e.g., Kitcher 1984 and Fodor 1987).

An analogous objection applies to hard determinism: we have more reason to believe that free mental choosing exists—primarily due to our seeming pervasive experience of such—than for believing the (highly speculative, theoretical) premises in the hard determinist's arguments to the contrary—again, that every event is subsumed by natural laws, and that if every event is subsumed by natural laws, free mental choosing is thereby ruled out. The objection to hard determinism from 'epistemological weighting', supposing it goes through, is better than the aforementioned objection, not least because its conclusion is that belief in this view is unjustified, not merely inconvenient.

Hard determinists might respond (as eliminativists respond, *mutatis mutandis*) by telling a story according to which our experience could be just as it is, but the apparent freedom of our choosings an illusion, if these choosings are governed by sufficiently complex laws. As a referee noted, this response resembles a classical skeptical response, where the skeptic undermines an ordinary belief by offering an alternative possibility that seems equally warranted. But in contrast to at least some comparatively compelling skeptical scenarios, any such hard determinist story will likely fail to undercut the objection from epistemological weighting. Consider, for example, a sequence of events in which you choose, in a seemingly completely free way, to throw a piece of chalk into the air. The hard determinist must explain the outcome of this sequence by citing some complex antecedent events, subsumed by natural law, that determine the outcome. But then the strategy for 'explaining away' our experience of free mental choosing relies on a premise that, as before, we have less reason to believe than the claim that free choosings exist; therefore, the objection to hard determinism stands.

3. Soft Determinism and Nonreductive Physicalism

We turn now to considering a response to the problem of free will on which free mental choosing exists, namely, soft determinism, according to which the antecedent of the *Free Will Conditional* (*Causal Determinism*) is true (or possibly

true), but the consequent (denying the existence of free will) is false.[1] Here again two positions—reductive and nonreductive physicalism—take the same stance as regards the *Mental Quausation Conditional*. As previously discussed, however, reductive versions of physicalism, which identify the mental events at issue with physical events, face immediate difficulties in making sense of how mental events can be efficacious *qua* the qualitative or intentional types they are, with the usual strategies being to offer pragmatic or purely epistemic accounts of the desired 'higher-level' relevance. (Hence David [1997: 137] says, in describing Kim's 'conceptual' account of such relevance, 'Kim asks us to embrace a form of eliminativism. . . . Our general concepts of pain, belief, desire, and so on are just that: they are just our concepts; they do not directly mirror any states or properties present in nature'.) Since the deeper concern at issue in the problem of free will is with whether there is a genuinely metaphysical basis for free mental quausation, we will here cut to the chase of considering the parallel between nonreductive physicalism, which standardly aims to make metaphysical sense of distinctive mental efficacy, and soft determinism.

We first identify, in schematic terms, the general strategy that distinguishes the nonreductive physicalist response to the problem of mental quausation, and show that seemingly diverse accounts of nonreductive physicalism aim to implement the strategy, and therefore are more similar than they appear. We then turn to soft determinism (positive compatibilism) and, drawing upon Hawthorne and Pettit's (1996) taxonomy, argue that seemingly diverse soft determinist accounts implement a structurally similar strategy for responding to the problem of free will. To prefigure: in both cases the strategy involves characterizing the target mental event as being associated with only a *proper subset of the (deterministic, physical) causal determinants of the effect at issue* in such a way as to provide a principled basis for the claim that the mental event is efficacious *qua* the type of (qualitative, intentional, freely deliberative) event it is.

Before continuing, a methodological caveat and a terminological clarification. First, the caveat: while by way of motivating the nonreductive physicalist and soft determinist strategies, we present these in broadly positive terms—for example, as being promising so far as satisfying certain desiderata is concerned—the structural parallel we aim to draw here does not depend on the success of these strategies or the accounts implementing them. Indeed, given that the strategies in the cases of nonreductive physicalism and soft determinism are structurally similar, reasons to think the strategy problematic in one case might serve as reasons to think the strategy problematic in the other (for pessimistic assessments of the nonreductive physicalist strategy to come, see, e.g., Audi 2012, Morris 2013, and Bernstein 2016). What is most important for our purposes is that these two seemingly different types of accounts are in fact structurally similar, in ways that we will now try to articulate;

1 Nothing important in what follows turns on whether a soft determinist takes *Causal Compatibilism* to be actually or (qua compatibilist) merely possibly true. That said, our discussion does not bear on soft determinists whose rejection of causation or causal production—Leibniz and perhaps also Hume—would entail their rejection of a causal reading of *Determinism*.

whether these accounts are ultimately correct is (way) beyond the purview of this paper.

Second, the clarification: while the problems of free will and of mental causation are often presented as involving events, some discussions of these problems reference other ontological categories; in particular, some nonreductive physicalist responses to the (refined) problem of mental causation are presented in terms of properties, reflecting the truism that what an object or event can enter into causing ultimately depends on what properties are associated with the event or object: for example, it is in virtue of being massy, not magnetic, that a magnet falls to earth, and it is in virtue of being magnetic, not massy, that a magnet attracts steel; and so on. (Note that even nominalists about properties will accept and aim to accommodate this truism via their preferred ontology.) As such, in what follows we will sometimes speak neutrally of mental and physical 'features' as covering, *mutatis mutandis*, events, properties, or states.

3.1 The Nonreductive Physicalist's 'Proper Subset Strategy'

Nonreductive physicalists aim to respond to the problem of mental causation by identifying a specific relation between mental and physical features that preserves the reality, distinctness, and efficacy of mental features, compatible with the physicalist commitment to *Physical Causal Closure*, and without inducing an unacceptable ('firing squad' or 'double rock throw') variety of overdetermination. Call an account of the relation between mental and physical features intended to satisfy these desiderata 'nonreductive realization'. A number of accounts of nonreductive realization have been proposed, including:

- Functional realization (Putnam 1967; Fodor 1974; Papineau 1993; Antony and Levine 1997; Melnyk 2006). Mental feature types are functional types that are implemented, on a given occasion, by tokens of physical feature types.
- Mereological realization (Shoemaker 2000–2001; Clapp 2001). Mental feature types/tokens are proper parts of physical feature types/tokens.
- The determinable-determinate relation (Yablo 1992; Wilson 2009). Mental feature types/tokens are determinables of physical feature types/tokens.

On the face of it, the relations appealed to in these accounts are quite diverse, and there is much to say about whether a specific such relation makes sense as applied to the mental/physical case. But as Wilson (1999 and 2011) argues, this seeming diversity hides a structural commonality, reflected in each of these relation's arguably guaranteeing satisfaction of a certain 'proper subset' condition on the powers of the mental feature vis-à-vis its physical base feature, as follows:

> *Subset Condition on Causal Powers* (*SCCP*): The token powers of a realized mental feature *M* on a given occasion are a non-empty proper subset of the token powers of the physical feature *P* realizing *M* on that occasion. (Wilson 2011: §1.2; see also Wilson 1999: 45)

The notion of 'power' at issue here may be understood in metaphysically neutral terms, as primarily reflecting the truism that what an entity can cause depends on how it is (again: it is in virtue of being massy, not magnetic, that a magnet falls to earth; and so on). Reflecting this truism, the condition presupposes that broadly scientific features be associated, perhaps only contingently, with ('have', 'bestow') powers to produce certain effects. In addition, the condition presupposes that a given feature will typically be associated with powers to produce a variety of effects, depending on the circumstances of its instantiation.

We will shortly present reasons that support taking the aforementioned accounts to satisfy *SCCP*. First, though, it is worth sketching, in a schematic way, how satisfaction of this condition is supposed to satisfy the nonreductive physicalist's desiderata (see Wilson 1999, 2009, and 2011 for detailed defense of the following claims). To start, a realization relation satisfying *SCCP* appears to guarantee the reality and distinctness of mental features, as per nonreduction: if *M* has a (nonvacuous) proper subset of powers of the physical base feature *P* that realizes *M* on a given occasion, then *M* exists and is distinct from *P*, in accordance with Leibniz's law. Such a relation also appears to conform to physicalism, in that satisfaction of *SCCP* blocks the possibility of robust emergentism or other routes to *M*'s physical unacceptability. Next, such a relation appears to avoid problematic overdetermination: if the relation between *M* and *P* satisfies *SCCP*, then for any effect produced by both *M* and *P* on a given occasion, only one power is manifested: there is only one causing, not two.

Finally, such a relation accommodates a sense in which *M* is causally relevant—that is, efficacious *qua* the type of (qualitative or intentional) mental feature it is. Here the idea is that there are two ways for a mental feature *M* to be causally relevant as compared to its base feature *P*. One way is for *M* to be associated with a *new* power—a power that *P* doesn't have. The nonreductive physicalist denies that mental features are causally relevant in this (robustly emergent) sense. A second way reflects that *M* is associated with a *distinctive power profile* consisting, as per *SCCP*, of a proper subset of the powers associated with *P*. The underlying basis for the powers-based strategy is the claim that *M*'s causal relevance does not require that *M* have a distinctive power: it is enough that *M* have a distinctive set (collection, plurality) of powers.

How is having a distinctive power profile supposed to provide a basis for *M* to be efficacious *qua* the type of mental feature it is? One case for taking distinctive power profiles to suffice for causal autonomy appeals to difference-making or 'proportionality' considerations in cases where *M* (or *M*'s type) is multiply realizable (see, e.g., Yablo 1992, Antony and Levine 1997, Wilson 1999, and Shoemaker 2000–2001). Suppose *M* is a state of feeling thirsty, which causes a desire for a glass of water (effect *E*). Now suppose that *M* (or another instance

of *M*'s type) were to have been realized by *P′* rather than *P*, in circumstances relevantly similar to those in which *M* caused *E*. Would *E* (or another instance of *E*'s type) still have occurred? Intuitively, yes, since the only powers that matter for the production of *E* are the powers associated with *M*; powers differing between *P* and *P′* (e.g., to produce different readings on a neuron detector) are irrelevant for the production of *E*. That *M*'s distinctive power profile contains just those powers relevant or 'proportional' to the production of *E* provides a principled reason for taking *M*'s efficacy vis-à-vis *E* to be distinctively different from that of *P*.

Another case for taking the having a distinctive power profile to be a basis for *M*'s being causally relevant, which applies even if *M* is only singly realized, reflects that the distinctive power profiles of features may be associated with distinctive systems of laws governing the features. Plausibly, systems of laws track causal joints in nature—in particular, they track more or less abstract causal joints, with fundamental physical laws tracking a causal joint that is highly specific and sensitive to microphysical details; and special science laws, ranging from chemistry to biology and psychology and beyond, tracking causal joints that are increasingly abstract and insensitive to lower-level details. Correspondingly, *M*'s distinctive power profile, associated with psychology (among other special sciences), reflects the existence of a distinctively psychological (qualitative or intentional) causal joint in nature—one that tracks a causal level that is broadly independent of microphysical details, whether or not *M* is multiply realizable. As such, even if every token power of *M*, on a given occasion, is identical with a token power of its physical realizer *P* on that occasion, *M* can be causally efficacious *qua* the (qualitative, intentional) mental event it is, in virtue of *M*'s distinctive power profile tracking the comparatively abstract causal level associated with psychology.

Again, the suggestion here is that there are two ways for a mental feature (event, property, state) *M* to be causally relevant—two ways for a mental feature to be efficacious *qua* the type of mental feature it is—as compared to the physical feature upon which the mental feature depends. One way—emphasized by Kim and others—is for *M* to be associated with a new power to produce the effect—a power that its physical realizer *P* doesn't have; here the distinctive efficacy reflects facts that are broadly intrinsic to the effect's production. Positing such a new power would violate *Physical Causal Closure* and as such is incompatible with physicalism. Another way—the way at issue in the powers-based strategy for accommodating nonreductive physicalism—is for *M* to be associated, either nomologically or essentially, with a collection of powers that are relevantly proportional to the effect in the ways indicated by difference-making considerations and special-science laws. Here, the distinctive efficacy reflects, in part, facts (concerning which power profiles are associated with which properties) that are broadly extrinsic to the production of the effect.

We turn now to canvassing reasons for thinking that each of the aforementioned specific accounts of nonreductive realization are plausibly seen as satisfying *SCCP* and so are structurally similar in that each is implementing a proper subset strategy.

3.1.1 ***Functionalist Accounts.*** First, consider functionalist accounts, according to which realized feature types are second-order types associated with causal roles

that, on a given occasion, are implemented by tokens of realizer types. A causal role is just a collection of powers. Hence, if *M* is of a functional type, then, on any given occasion, every token power of *M* will be numerically identical with a power of the base state *P* that implements *M*'s causal role on that occasion. Moreover, functionalist accounts are typically motivated by and thus presuppose the multiple realizability of the mental types in question: recall the analogy to computer systems, whereby a given piece of software may be implemented by many different hardware systems. As such, the powers of a functionally realized type will be a proper subset of each of its realizing types. This proper subset relation between powers will plausibly hold between token powers of the instantiated types, in which case *SCCP* will be satisfied.

3.1.2 ***Mereological accounts.*** Second, consider a broadly mereological account of realization, according to which, as Shoemaker (2000–2001: 81) says, 'the instantiation of a realizer property entails, and might naturally be said to include as a part, the instantiation of the functional property realized' (see also Clapp 2001). A mereological account of realization is promising for nonreductive physicalist purposes insofar as proper parts of more fundamental features are distinct from and yet in a sense nothing over and above these features, and proper parts may be efficacious without inducing overdetermination, as when both you and your eye cause a wink (as discussed in Paul 2002). Does a mereological account satisfy *SCCP*? It must do so, if a given mental feature *M* is to be guaranteed to be causally efficacious (at all); for on some accounts of features the latter have noncausal aspects ('quiddities'), and in that case the holding of a proper parthood relation between mental *M* and physical *P* is compatible with *M*'s being epiphenomenal.

In fact, Shoemaker and Clapp each assume that features are essentially and exhaustively constituted by their associated powers (as per Shoemaker 1980), in which case the holding of a proper parthood relation between *M* and *P* will entail satisfaction of *SCCP*. Moreover, Shoemaker, like Wilson (1999 and 2011), sees satisfaction of *SCCP* as key to nonreductive realization, with a mereological view being (given his preferred account of properties) a natural way of ensuring satisfaction of this condition.

Along lines similar to those discussed above, Shoemaker (2000–2001: 78–79) takes satisfaction of *SCCP* to be motivated by considerations of multiple realizability:

> Where the realized property is multiply realizable, the conditional powers bestowed by it will be a proper subset of the sets bestowed by each of the realizer properties.

When any such feature is multiply realized, its realizing types will share all the powers of the realized type, but they will differ with respect to further powers. This type-level proper subset relation between powers will plausibly hold between token powers of the instantiated types, in which case *SCCP* will be satisfied.

3.1.3 ***Determinable-based Accounts.*** Finally, consider accounts of nonreductive realization in terms of the determinable/determinate relation, the relation of increased specificity paradigmatically holding between colors and their shades. Yablo (1992:259) suggests that taking mental features to be determinables of physical determinates avoids problematic overdetermination:

> [W]e know that [determinables and determinates] are not causal rivals. This kind of position is of course familiar from other contexts. Take for example the claim that a space completely filled by one object can contain no other. Then are even the object's parts crowded out? No. In this competition wholes and parts are not on opposing teams.... Determinables and their determinates, like objects and their parts, are guaranteed to be on the same team.

Yablo anticipated that many might find his suggestion puzzling, but as suggested by Wilson (1999), a determinable-based account can be rendered less puzzling by attention to the powers of the features (types and tokens) involved (see also Wilson 2009).

To start, consider a patch that is red, and more specifically, scarlet. Sophie the pigeon has been trained to peck at any red patch; she is presented with the patch, and she pecks. The patch's being red caused Sophie to peck—after all, she was trained to peck at red patches. But the patch's being scarlet also caused Sophie to peck—after all, to be scarlet just is to be red, in a specific way. Nonetheless, Sophie's pecking was not problematically overdetermined. Plausibly, this is because each token power of the determinable red instance is numerically identical with a token power of its determining scarlet instance. Moreover, given Sophie's training, she would have pecked even had the patch been a different shade of red—burgundy, say. This is not true, however, of Sophie's cousin Alice, who has been trained to peck *only* at scarlet patches. This suggests that the determinable type *red* has fewer powers than its determinate types. More generally, since broadly scientific determinables are associated with distinctive sets of powers and are typically 'multiply determinable', the powers of determinable types will typically be a proper subset of the powers of their determinate types. This relation will plausibly hold, in turn, between token powers of determinable and determinate instances; hence, a determinable/determinate account of realization satisfies *SCCP*.

To sum up: a wide variety of apparently diverse accounts of nonreductive realization aim to implement a proper subset strategy, either implicitly or explicitly, and so at a structural level are more similar than they appear. To be sure, whether a specific such relation is suitably posited as holding between mental and physical events is a further question. But in any case, to the extent that these relations are considered promising for purposes of providing a nonreductive physicalist response to the problem of mental quausation, this is due primarily to their appearing to ensure satisfaction of the proper subset condition expressed in *SCCP* and hence to their appearing to accommodate the efficacy of mental events *qua* the qualitative

and intentional types of events they are, in a way compatible with *Physical Causal Closure*.

3.2 The Soft Determinist's Proper Subset Strategy

We now turn to the problem of free will, understood as the problem of whether and how events of mental choosing can be efficacious *qua* free (as previously, in being, in some substantive sense, 'up to the agent') and to showing that soft determinist (positive compatibilist) responses to the problem implement a strategy that is structurally similar to that implemented by nonreductive physicalists in response to the problem of mental quausation.

Hawthorne and Pettit's (1996) taxonomy of compatibilist strategies serves as a useful basis of operations. They start by noting:

> All compatibilists agree that every choice has antecedents and . . . that this fact puts freedom of choice in doubt. How can a choice be made freely if it is the product of independent antecedents? The response they make is that some possible antecedents are better than others from the point of view of free choice and that a choice is free to the extent that its antecedents, or at least its relevant antecedents, satisfy the inherently vague condition of leaving it up to the agent. (191)

In schematic form:

> *X* chooses freely to ϕ if and only if the relevant antecedents of the choice leave the ϕ-ing up to *X*. (191)

Here we suppose that a choice to ϕ (the outcome of an event of choosing) may be either an intention to ϕ or a ϕ-ing.

Hawthorne and Pettit identify three main compatibilist accounts of what it is for a choice to be 'up to an agent', associated with the notions of freedom as underdetermination, ownership, and responsibility, respectively. The accounts vary to some extent as regards which causal antecedents are supposed to be relevant to establishing whether the choice was up to the agent in the intended sense. But as we will shortly argue, on all three conceptions the relevant causal antecedents will be a proper subset of the set (collection, plurality) of casual antecedents that in fact determines the outcome of the choosing.

More specifically, we will argue that each of the following accounts plausibly imposes satisfaction of a certain 'proper subset' condition as key to each account's strategy of response to the problem of free will. To start, reflecting endorsement of *Causal Determinism*, the compatibilist accepts the following condition on the causal antecedents of any outcome of a free mental deliberation:

> *Condition on Causal Antecedents* (*CCA*): The total causal antecedents of an event of free choosing *M* completely determine the outcome of *M* (e.g., a choice to ϕ).

As a first pass, the compatibilist strategy requires that a free mental choosing *M* satisfy the following proper subset condition:

> *Subset Condition on Causal Antecedents* (first pass): The relevant causal antecedents {*C*} of a free mental choosing *M* are (i) a nonempty proper subset of the total causal antecedents of *M*, which (ii) satisfy the condition of leaving the outcome of *M* up to the agent.

Moreover, as we'll shortly argue, if the compatibilist's strategy of identifying the relevant causal antecedents of *M* is to make sense of the idea that these antecedents leave the choice up to the agent, then the relevant antecedents must more specifically satisfy the following (final pass) proper subset condition, *SCCA*:

> *Subset Condition on Causal Antecedents* (*SCCA*): A free mental choosing *M* resulting in a choice to ϕ satisfies the following: (i) *M* has relevant causal antecedents {*C*} that are a nonempty proper subset of the total causal antecedents of *M*, and (ii) it is possible that a choosing *M′* of the same type as *M* occurs, having relevant antecedents {*C′*} of the same type as {*C*}, but where the *total* antecedents of *M′* are such as to determine completely the outcome of *M′* (as per *CCA*) as either a choice not to ϕ or as the absence of a choice to ϕ.

3.2.1 ***Freedom as underdetermination.*** On underdetermination accounts, a choice to ϕ is the result of a free choosing *M* iff *M* could have resulted in a choice not to ϕ. How could this be, given that the choice to ϕ was determined, as per *CCA* (encoding *Causal Determinism*)? The underdetermination approach proceeds by identifying a subset {*C*} of the causal antecedents of the choice to ϕ relative to which it was left open whether or not *M* would result in a choice to ϕ:

> Taken as a whole, the antecedents of any choice will necessitate that choice under a deterministic picture and compatibilists of this stripe must take the relevant antecedents to be a subset of the totality. But which subset? (Hawthorne and Pettit 1996: 193)

The relevant subset of antecedent events will include the choosing event, along with events tracking the standing beliefs and desires of the agent at the time of choosing and events tracking whether the choosing took place under conditions of physical restraint, threat, etc. (cf. Ayer 1954). From broader perspectives, the relevant antecedents might also include events tracking cultural influences, past trauma, or other psychological, social, psychiatric, neurological, etc., conditions holding of the agent. In general:

> The line will be that an agent is free to the extent that the antecedents that can or have to be countenanced in that perspective leave the choice underdetermined. . . . To be free, if you like, is to be free relative to that stance. (Hawthorne and Pettit 1996: 193–94)

Here the relevant antecedent events must be a *proper* subset of the causal antecedents that, as per *CCA*, completely determine the choice, since only if the subset is proper is there any hope that the subset of antecedents will leave that choice underdetermined. Moreover, the assumption that the subset of antecedents leaves the choice underdetermined plausibly entails (indeed, has as its content) that it is possible that a choosing M' of the same type as M occurs, having relevant antecedents $\{C'\}$ of the same type as M's relevant antecedents $\{C\}$, but where the *total* antecedents of M' are such as to completely determine the outcome of M' as either a choice not to ϕ or the absence of a choice to ϕ.

This last entailment just is *SCCA*. As such, a 'freedom as underdetermination' form of compatibilism explicitly implements a proper subset strategy, characterizing a mental choosing M as associated with a proper subset of its causal antecedents, and using this association to accommodate M's being distinctively efficacious, *qua* free, vis-à-vis the ensuing choice.

3.2.2 ***Freedom as Ownership.*** A second compatibilist approach takes freedom to be a matter of ownership:

> The ownership line takes a choice to be up to an agent to the extent that it is not due to anyone or anything other than the agent themselves; it is a choice that the agent owns, a choice with which the agent identifies, and not something forced upon them. Suppose that the relevant antecedents in the adjudication of free will are taken to be . . . beliefs and desires. [Then] an agent ϕs freely just in case their beliefs and desires combine to lead—or at least lead in 'the right way' (see Davidson 1963) to their ϕ-ing. (Hawthorne and Pettit 1996: 194)

Underdetermination by the relevant antecedents is not explicitly required here since an agent could own or identify with completely determined intentions. But, Hawthorne and Pettit argue, if the 'ownership' line is to be viable, it will have to ensure underdetermination by these antecedents, and so satisfaction of *SCCA*.

Why is this so? To start, note that if, for example, you are brainwashed with beliefs and desires leading to your choice to ϕ, this intention cannot be seen as the effect of free mental choosing. A well-known response (see Frankfurt 1971) requires that choices result from desires that the agent X desires, at the second order, to have and be moved by. The brainwashing problem will re-arise, however, unless 'the action issues from desires that the agent has some measure of second-order control over' (Hawthorne and Pettit 1996: 195). As O'Connor (2002) puts it:

> We can . . . imagine external manipulation consistent with Frankfurt's account of freedom but inconsistent with freedom itself. . . . One might

> discreetly induce a second-order desire in me to be moved by a first-order desire—a higher-order desire with which I am satisfied—and then let me deliberate as normal. Clearly, this desire should be deemed 'external' to me, and the action that flows from it unfree.

These considerations indicate that one needs to ensure somehow that the formation of second-order desires is up to the agent, and the natural compatibilist approach will be to restrict attention to a proper subset of the antecedents determining the desire—for example, those relevant to whether the agent's choosing was constrained by other persons, or by other psychological, social, psychiatric, neurological, etc., conditions. In other words, to accommodate the needed control of second-order desires on a 'freedom as ownership' picture, a proper subset of the antecedents of the choice to ϕ must be specified relative to which it was underdetermined that the agent had the second-order desires the agent actually had, hence underdetermined that the agent would identify with the first-order beliefs and desires leading to the agent's choice to ϕ, hence underdetermined that the agent's choosing would result in a choice to ϕ. Such underdetermination in turn entails (indeed, has as its content) satisfaction of *SCCA*.

More generally, on a 'freedom as ownership' form of compatibilism we see implementation of a proper subset strategy, whereby a mental choosing *M* is associated with a proper subset of its causal antecedents, in service of making room, as per *SCCA*, for *M* to be causally efficacious, *qua* free, vis-à-vis the ensuing choice.

3.2.3 ***Freedom as responsibility.*** On the 'freedom as responsibility' approach, a choice to ϕ is the result of a free choosing *M* iff the agent of the choosing could be held responsible for the outcome of *M*. The criteria for an agent's being responsible might advert to prevailing systems of law and morality, or (as per Strawson 1962) to the participant or reactive attitudes characteristic of human interactions. As Hawthorne and Pettit (1996: 197) point out, this approach also requires that the choice at issue be underdetermined:

> To hold an agent responsible in certain choices is to think that it is not inevitable either that they get things right or that they get them wrong—either that they do well or that they do ill—and so it is to believe that there is a sense in which they could have done otherwise.

The relevant antecedents in this case would then include those relevant to determining whether the agent was deliberating under conditions where the agent would, by the lights of the prevailing system of law, morality, or interaction, be held responsible for the outcomes of the choosings. Again, these might cite events tracking whether various physical, psychological, neurophysiological, etc., conditions or constraints were in place antecedent to or concurrent with the choosing.

Here again, the account (i) identifies a relevant proper subset of *M*'s causal antecedents and (ii) requires that, relative to these antecedents, the outcome of *M*

could have been different, as a way of making sense of *M*'s being efficacious, *qua* free. In other words, a 'freedom as responsibility' form of compatibilism implements a proper subset strategy, encoded in satisfaction of *SCCA*.

To sum up: All three soft determinist/positive compatibilist accounts require that the effect or outcome of free mental choosings be underdetermined, in the sense that relative to a proper subset {*C*} of the total set of causal antecedents of the choice to ϕ, the outcome of a mental choosing *M* could have been different from what it actually was, notwithstanding that the outcome of *M* was completely determined relative to the total set of *M*'s antecedents. To be sure, soft determinists/positive compatibilists disagree about which antecedents of a given mental choosing *M* are relevant to tracking whether the outcome of *M* is appropriately 'up to the agent', but they agree that the relevant antecedents satisfy the proper subset condition expressed in *SCCA*. As such, a variety of apparently diverse compatibilist accounts are structurally more similar than they at first appear.

3.3 Motivating Compatibilism in Light of the Parallel

Nonreductive physicalist and soft determinist/positive compatibilist positions each respond to their respective problems by characterizing the mental event at issue as associated with a proper subset of the 'causal determinants' of the effect. To be sure, the determinants are not the same: in the one case, they are powers; in the other case, they are causal antecedents. Nonetheless, the structural similarity in strategies is clear: in each case, associating the mental event *M* with (only) the relevant proper subset of causal determinants is supposed to provide a basis for showing that *M* is causally efficacious vis-à-vis the effect *E* in question, *qua* the type of mental event that *M* is.

This structural similarity is also useful as a case in point of how connecting the debates about free will and those about mental causation may be fruitful. For it suggests how soft determinists might respond to a concern about their strategy—namely, the concern that identification of a given subset of causal antecedents won't make sense of how a choosing *M* can be efficacious *qua* free, since the mere presence of a subset of antecedents doesn't establish that *M* 'selects' or 'determines' the outcomes of the choosing.

To start, the soft determinist, like the nonreductive physicalist, will grant that *M*'s relevance vis-à-vis the effect at issue doesn't proceed by way of *M*'s having a distinctive power: just as *Physical Causal Closure* blocks taking a qualitative or intentional mental event *M* to have a new power (that would be strong emergence, not physicalism), so too does *Causal Compatibilism* block taking a mental event of choosing *M* to have a new power (that would be libertarianism, not soft determinism).

Even so, just as the nonreductive physicalist has alternative ways of motivating the distinctive causal relevance of qualitative and intentional mental events—either as tracking difference-making considerations (if the physical realizer had been slightly different, I would still have been thirsty) or as tracking a distinctive comparatively abstract psychological level of causal grain—so too the soft

determinist may maintain that even in the absence of new powers to 'select' or 'determine' outcomes, *M* may be causally relevant to those outcomes, either in tracking difference-making considerations (if the causal antecedents of my choice had been slightly different, I would still have chosen as I did) or as tracking a distinctive broadly psychological level of causal grain. The nonreductive physicalist's distinctive but nonproductive form of causal relevance appears to be, *mutatis mutandis*, just what the soft determinist needs.

4. Closing Remarks

We have argued for the following claims:

1. The problems of free will and of mental causation may each be seen as instances of a general problem of mental quausation, as follows:

   > How can a mental event *M* of a given type be efficacious vis-à-vis an event *E* in virtue of being the type of mental event it is, given that there is reason to think that events of *M*'s type are causally irrelevant to the production of events of *E*'s type?

2. Attention to the *Free Will Conditional* and the *Mental Quausation Conditional* indicates that there are parallels between two of the main positions in each debate, in particular, between hard determinism and eliminativist physicalism and between soft determinism (positive compatibilism) and nonreductive physicalism.
3. The parallel between hard determinism and eliminativist physicalism is useful in suggesting an 'epistemic weighting' objection to hard determinism, according to which acceptance of hard determinism is not just pragmatically infeasible but moreover unjustified.
4. The parallel between soft determinism/positive compatibilism and nonreductive physicalism is useful in revealing that apparently diverse soft determinist/positive compatibilist accounts, like apparently diverse nonreductive physicalist accounts, are structurally alike in implementing 'proper subset' strategies for resolving their respective problems of mental quausation.
5. This structural similarity is also useful in suggesting how soft determinists/positive compatibilists may respond to concerns that identification of a given subset of causal antecedents won't make sense of how a choosing *M* could be efficacious *qua* free, by observing (following an available nonreductive physicalist response to a similar concern raised against their proper subset strategy) that considerations of difference-making and/or distinctive levels of

causal grain show how a distinctive set of causal antecedents can be associated with a distinctive form of causal relevance.

Other parallels remain to be drawn between the 'incompatibilist' positions in each debate—most promisingly, between robust emergentism and libertarianism.

SARA BERNSTEIN
UNIVERSITY OF NOTRE DAME
sbernste@nd.edu

JESSICA WILSON
UNIVERSITY OF TORONTO
jessica.m.wilson@utoronto.ca

References

Antony, Louise M., and Joseph M. Levine. (1997) 'Reduction with Autonomy'. *Philosophical Perspectives*, 11, 83–105.
Audi, Paul. (2012) 'Properties, Powers, and the Subset Account of Realization'. *Philosophy and Phenomenological Research*, 84, 654–74.
Ayer, Alfred J. (1954) 'Freedom and Necessity'. In Alfred J. Ayer, *Philosophical Essays* (London: Macmillan), 271–84.
Bernstein, Sara. (2016) 'Overdetermination Underdetermined'. *Erkenntnis*, 81, 17–40.
Churchland, Patricia S. (1986) *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge, MA: MIT Press.
Churchland, Paul M. (1981) 'Eliminative Materialism and the Propositional Attitudes'. *Journal of Philosophy*, 78, 67–90.
Clapp, Lenny. (2001) 'Disjunctive Properties: Multiple Realizations'. *Journal of Philosophy*, 98, 111–36.
David, Marian. (1997) 'Kim's Functionalism'. *Philosophical Perspectives*, 11, 133–48.
Davidson, Donald. (1963) 'Actions, Reasons, and Causes'. *Journal of Philosophy*, 60, 685–700.
Feyerabend, Paul K. (1963) 'Mental Events and the Brain'. *Journal of Philosophy*, 40, 295–96.
Fodor, Jerry. (1974) 'Special Sciences (Or, The Disunity of Science as a Working Hypothesis)'. *Synthese*, 28, 77–115.
Fodor, Jerry. (1987) *Psychosemantics*. Cambridge, MA: MIT Press.
Frankfurt, Harry G. (1971) 'Freedom of the Will and the Concept of a Person'. *The Journal of Philosophy*, 68, 5–20.
Hawthorne, John, and Philip Pettit. (1996) 'Strategies for Free Will Compatibilists'. *Analysis*, 56, 191–201.
Holbach, Paul-Henri. (1770) *The System of Nature; or, Laws of the Moral and Physical World*. New York: Franklin.
Horgan, Terence. (1989) 'Mental Quausation'. *Philosophical Perspectives,* 3, 47–76.
Kane, Robert. (2002) 'Introduction'. In Robert Kane (ed.), *The Oxford Handbook of Free Will* (Oxford: Oxford University Press), 3–41.
Kim, Jaegwon. (1989) 'The Myth of Nonreductive Materialism'. *Proceedings and Addresses of the American Philosophical Association*, 63, 31–47.
Kitcher, Patricia S. (1984) 'In Defense of Intentional Psychology'. *Journal of Philosophy*, 81, 89–106.
McKenna, Michael, and Justin Coates. (2008) 'Compatibilism'. In Edward Zalta (ed.), *Stanford Encyclopedia of Philosophy*.

Melnyk, Andrew. (2006) 'Realization-based Formulations of Physicalism'. *Philosophical Studies*, 131, 127–55.
Morris, Kevin. (2013) 'On Two Arguments for Subset Inheritance'. *Philosophical Studies*, 163, 197–211.
Ney, Alyssa. (2008) 'Defining Physicalism'. *Philosophy Compass*, 3, 1033–48.
O'Connor, Timothy. (2002). 'Free Will'. In Edward Zalta (ed.), *Stanford Online Encyclopedia of Philosophy*.
Papineau, David. (1993) *Philosophical Naturalism*. Oxford: Basil Blackwell.
Paul, L. A. (2002) 'Logical Parts'. *Nous*, 36, 578–96.
Putnam, Hilary. (1967) 'Psychological Predicates'. In W. H. Capitan and D. D. Merrill (eds.), *Art, Mind, and Religion* (Pittsburgh, PA: University of Pittsburgh Press), 37–48.
Shoemaker, Sydney. (1980) 'Causality and Properties'. In Peter van Inwagen (ed.), *Time and Cause* (Dordrecht: Reidel), 109–35.
Shoemaker, Sydney. (2000–2001) 'Realization and Mental Causation'. In Bernard Elevitch (ed.), *Proceedings of the 20th World Congress in Philosophy*, (Cambridge, UK: Philosophy Documentation Center), 23–33. Published in revised form in Gillet, C. and Loewer, B. (eds.), *Physicalism and its Discontents*. Cambridge, UK: Cambridge University Press.
Steward, Helen. (2015) 'What is Determinism?'. *Flickers of Freedom* blog. Available at: http://philosophycommons.typepad.com/flickers_of_freedom/2015/07/what-is-determinism.html.
Strawson, Peter F. (1962) 'Freedom and Resentment'. *Proceedings of the British Academy*, 48, 1–25.
van Inwagen, Peter. (1983) *An Essay on Free Will*. Oxford: Clarendon Press.
Wilson, Jessica M. (1999) 'How Superduper Does a Physicalist Supervenience need to Be?' *The Philosophical Quarterly*, 49, 33–52.
Wilson, Jessica M. (2006) 'On Characterizing the Physical'. *Philosophical Studies*, 131, 61–99.
Wilson, Jessica M. (2009) 'Determination, Realization, and Mental Causation'. *Philosophical Studies*, 145, 149–69.
Wilson, Jessica M. (2011) 'Non-reductive Realization and the Powers-based Subset Strategy'. *British Journal for the Philosophy of Science*, 94, 121–54.
Yablo, Stephen. (1992) 'Mental Causation'. *The Philosophical Review*, 101, 245–80.