

Free Will and Mental Quausation*

Sara Bernstein[†] and Jessica Wilson[‡]

January 14, 2016

Introduction

Free will, if such there be, involves free choosing: the ability to mentally choose 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 free will and mental causation debates have proceeded largely independently of each other. Discussions of free will have rarely addressed whether and how the going 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.

Here we aim to make progress in determining how the free will and mental causation debates bear on one another. In §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 may be efficacious, *qua* the types of event 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,

*Penultimate version, forthcoming in the *Journal of the American Philosophical Association*. Thanks to Benj Hellie, David Sosa, and two anonymous referees for this journal, and special thanks to Michael McKenna, for valuable comments and suggestions that greatly improved this paper.

[†]Department of Philosophy, Duke University; sara.bernstein@duke.edu

[‡]Department of Philosophy, University of Toronto; jessica.m.wilson@utoronto.ca

which suggests a better argument against hard determinism than that typically offered (§2), and we argue that there is a parallel between soft determinism and non-reductive physicalism, which illuminates a common structural basis and associated strategy for positively accommodating the compatibility of free will and determinism (§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 that we are not in position to rule out of court. 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.¹

Determinism admits of different interpretations.² We are here concerned with a reading involving causal determination of present or future events by prior states (broadly construed), as follows:

Causal Determinism: Every event³ is a causal consequence of the laws of nature and the state of the world at any prior time.

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:

¹One might wonder whether *Determinism* can be ruled out of court, on grounds that quantum mechanics, which is likely true, is indeterministic; but this would be premature, both because there are deterministic interpretations of quantum mechanics (see, e.g., Bohm 1952), and because the present incompatibility of quantum mechanics and general relativity suggests that we are not yet in position to infer to the likely truth of quantum mechanics. Even if *Determinism* is actually false, one might be interested in understanding what bearing the truth of *Determinism* would have on free will.

²See Steward's (2015) for discussion, in which she distinguishes between interpretations based on entailment between propositions at one time by propositions at another time, and those based on causation between events.

³Apart from any first events; henceforth we assume this restriction.

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.⁴ (van Inwagen 1983, 16)

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’.⁵

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 physical⁶ effect has a purely physical cause.⁷

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 intentional⁸ mental event M is, by *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. 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 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, 280–1)

⁴Van Inwagen’s remarks are directed at an entailment-based reading of *Determinism*, but they apply, *mutatis mutandis*, to a causal reading (simply insert ‘causal’ before ‘consequences’).

⁵Here we depart, again for convenience, from van Inwagen’s preferred (2008) usage, in which ‘the problem of free will’ adverts to a constructive dilemma for free will skepticism (the claim that no one has free will).

⁶There are nuances concerning how to define the ‘physical’ into which we cannot enter here; see Wilson 2006 and Ney 2008 for further discussion.

⁷As with *Determinism*, the actual truth of *Physical Causal Closure* remains empirically controversial; but even if this thesis is actually false one might be interested in understanding what bearing its truth would have on the viability of mental causation.

⁸Mental choosings are also intentional (and perhaps also qualitative) events; but in discussing the problem of mental causation, by ‘qualitative’ and ‘intentional’ we’ll have in mind the experiences and intentional states (e.g., pain states, beliefs) typically at issue in these contexts, for which free choosing is not at issue.

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 a 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? I.e., do the mental types (properties) tokened by mental events and states have the kind of relevance to individual causal transactions which allows these properties to figure in genuine causal explanations? (47)

Shoemaker (2000/2001) similarly expresses the deeper concern:

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 adequately vindicate 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” (50). As such, it is mental ‘qua causation’, not mental causation per se, that is most deeply challenged by the possible truth of *Physical Causal Closure*. How could a mental event *M* be efficacious *vis-à-vis* an effect *E*, in virtue of being qualitative or intentional, given that *E* was causally determined by physical events and associated laws in ways that seem to preclude *M*’s being causally relevant in either of these respects?

Though Horgan’s discussion, like that of Shoemaker and others, concerned mental events for which freedom is not at issue, the deeper concern about whether and how mental events can be seen as 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 free will debates. 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. How could an event of mental choosing *M* be efficacious *vis-à-vis* an event *E*, in virtue of being *free*, given that *E* was 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?

This parallel suggests that the problem posed by *Physical Causal Closure* for 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 qua causation:

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; nor are we suggesting that the problem of free will is an instance of the problem of mental qua causation as applied to qualitative or intentional mental events, since there is no clear tension between *Causal Determinism* and mental qua causation involving mental causes for which freedom is not at issue.⁹ Rather, we are suggesting that both the problem of free will and the refined problem of mental causation (as directed at qualitative and intentional mental events) are specific instances of a suitably general problem of mental qua causation, whereby the causal relevance of a mental event of a distinctive—intentional,

⁹As a referee correctly pointed out, *Causal Determinism* as such does not seem to exclude mental causation, or even mental qua causation (at least, by non-free mental events), if, e.g., natural laws include psychological laws.

qualitative, freely deliberative—type is called into question by the holding of certain theses which either might be or might have been true. 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* (of whatever variety) 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 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 hold that both antecedent and consequent are true; soft determinists (effectively: ‘positive’ compatibilists¹⁰) take the antecedent to be true and the consequent to be false; libertarians hold that both antecedent and consequent are false.

Similarly, the traditional responses to the refined problem of mental causation can be categorized by reference to the following mental causation 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 hold that both antecedent and consequent are true; reductive and non-reductive physicalists hold that the antecedent is true and the consequent is false; emergentists and substance dualists hold 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 towards 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

¹⁰In general, compatibilists maintain, as Lewis (1981) puts it, that “soft determinism may be true”. Beyond this commitment, individual compatibilists may aim simply to undermine arguments for incompatibility, or moreover to provide a positive conception of the compatibility at issue (see McKenna and Coates 2008 for discussion). While the parallel with non-reductive physicalism we are soon to draw has bearing on compatibilism generally, it has special interest for illuminating positive compatibilist positions of the sort we discuss in §3; and insofar as nothing deep turns on whether these positive accounts are actually or merely possibly implemented (that is: on whether *Causal Determinism* is assumed to be actually or merely possibly true), we won’t much distinguish between soft determinism and positive compatibilism. Thanks to a referee for discussion of these nuances.

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 non-reductive physicalism, and that in each case the parallel bears fruitfully on the free will debate.¹¹

2 Hard determinism and eliminative physicalism

Hard determinists (e.g., Holbach 1770) take the same stance *vis-à-vis* the free will conditional that eliminativist physicalists and epiphenomenalists take towards the mental causation 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 causation of the variety at issue. Indeed, eliminativists and epiphenomenalists go further, in assuming that qualitative mental events are efficacious at all (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 mental causation debate that is not present in the free will debate, concerning whether the efficacy of the mental events at issue is necessary for their existing. Most participants to the mental causation debate think so, but epiphenomenalists disagree; hence while they agree with eliminativists in denying the efficacy of (in particular, qualitative) mental events, they 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.

In 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 play in the literature, and which is better than one common objection to this position.

The common objection charges that hard determinism would be extremely difficult, perhaps even impossible, to put into practice. As Kane (2002) puts it, “Few thinkers have been willing to

¹¹Since libertarianism (on the one hand), and emergentism and substance dualism (on the other) introduce special considerations requiring detailed treatment, pertaining to singularism about causation, naturalism, agent causation, and the like, we leave treatment of parallels between these views for another occasion.

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” (27). This objection 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 argument. 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, e.g., 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, they 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 seeming pervasive experience of such—than we have to believe the (highly speculative, theoretical) premises in the eliminativist’s arguments to the contrary.¹²

An analogous objection applies to hard determinism: we have more reason for believing that free mental choosing exists—primarily as a result of 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 unlike at

¹²For arguments in the ballpark of this objection, see, e.g., Kitcher 1984 and Fodor 1987.

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 I 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, which 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; so the objection to hard determinism stands.

3 Soft determinism and non-reductive 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, but the consequent (denying the existence of free will) is false.¹³ Here again two positions—reductive and non-reductive physicalism—take the same stance as regards the refined mental causation 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.¹⁴ Since the deeper concern at issue in the problem of free will is with whether there is a genuinely metaphysical basis for free mental causation, we will here cut to the chase of considering the parallel between non-reductive physicalism, which standardly aims to make metaphysical sense of distinctive mental efficacy, and soft determinism.

We first identify, in schematic terms, the general strategy which distinguishes the non-reductive physicalist response to the problem of mental causation, and show that seemingly diverse accounts of non-reductive physicalism aim to implement the strategy, and so are more similar than they appear. We then turn to compatibilism, and drawing upon Hawthorne and Pettit’s (1996) taxonomy,

¹³Alternatively, for those positive compatibilists who are neutral on the actual truth of *Causal Determinism*: according to which the antecedent of the free will conditional (*Causal Compatibilism*) might be true, and the consequent (denying the existence of free will) false. Again, nothing important for the parallel we aim to draw here turns on whether a given compatibilist takes *Causal Compatibilism* to actually or more possibly true. That said, our discussion does not bear on compatibilists whose rejection of causation or causal production between events—Leibniz, and perhaps also Hume—would entail their rejection of a causal reading of *Determinism*.

¹⁴Hence David (1997) 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”.

argue that seemingly diverse compatibilist 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 caveat: while by way of motivating the non-reductive physicalist and soft determinist strategies we present these in broadly positive terms—e.g., as being promising as regards the satisfying of certain desiderata—the parallel we aim to draw here does not depend on the success of these strategies or associated accounts. Indeed, given that the strategies in the cases of non-reductive physicalism and soft determinism are parallel, reasons to think the strategy problematic in one case might well serve as reasons to think the strategy problematic in the other. What is most important for our purposes is that these two seemingly different types of accounts are in fact deeply parallel, in ways that we will now try to articulate; whether these accounts are ultimately correct is (way) beyond the purview of this paper.

3.1 The non-reductive physicalist’s ‘proper subset strategy’

Non-reductive physicalists aim to respond to the problem of mental causation by identifying a specific relation between mental and physical events (more generally: 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.¹⁵ Call an account of the relation between mental and physical features intended to satisfy these desiderata ‘non-reductive realization’. A number of accounts of non-reductive 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 base feature types.
- Mereological realization (Shoemaker 2000/2001, Clapp 2001). Mental feature types/tokens are proper parts of base feature types/tokens.

¹⁵Not everyone agrees that the schematic strategy to be shortly discussed can satisfy all these desiderata; for pessimistic assessments, see, e.g., Audi 2012, Morris 2013, and Bernstein forthcoming.

- The determinable-determinate relation (Yablo 1992, Wilson 2009). Mental feature types/tokens are determinables of base 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 features *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.¹⁶ (Wilson 1999, 45; 2011, §1.2)

We will shortly present reasons supporting 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 desiderata. To start, a realization relation satisfying *SCCP* appears to guarantee the reality and distinctness of mental features, as per non-reduction: if M has a (non-vacuous) proper subset of powers of the physical base feature P that realizes it on a given occasion, then M exists and is distinct from P , by 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.¹⁷ 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.

¹⁶The notion of ‘power’ at issue can be taken to be metaphysically neutral, reflecting mainly the truism that what a broadly scientific particular can do (cause) depends on how it is—that is, on what features it has: it is in virtue of being massy, not magnetic, that a magnet falls to earth; in virtue of being magnetic, not massy, that a magnet attracts steel; and so on. Reflecting such truisms, we may say that broadly scientific features are associated with (‘have’, ‘bestow’) powers to produce the effects; in addition, a given feature will typically be associated with powers to produce a variety of effects, depending on the circumstances of its instantiation. Philosophers differ over the nature of powers, their role in causation, and whether they are essential to or exhaustive of the features that have them. As we’ll observe below (note 17), no commitment on these or other controversial issues is required in order to implement the powers-based strategy to follow.

¹⁷One might be concerned (following Melnyk 2006 and 2008) that *SCCP*, even if necessary for non-reductive realization, isn’t sufficient for this, on grounds that if powers are only contingently associated with features, satisfaction of *SCCP* won’t ensure that P metaphysically necessitates M ; and if features have non-causal quiddities in addition to powers, satisfaction of *SCCP* won’t ensure that M is ‘wholly constituted’ by P . But as Wilson (2011) argues, for purposes of establishing that broadly scientific feature M is in fact “nothing over and above” physical feature P (that is: is physically acceptable), neither non-causal quiddities (which play no role, either experimentally or theoretically, in scientific theorizing) nor goings-on in worlds with different laws of nature, are relevant. As such, satisfaction of *SCCP* is sufficient for purposes of ensuring physical acceptability; and *pace* Melnyk, commitment to a controversial theory of properties (i.e., along lines of that endorsed in Shoemaker 1980) is not required in order to implement the strategy.

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 non-reductive 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 the having of 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 discussion in, 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 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) have still 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 E ’s production. That M ’s distinctive power profile contains just those powers relevant or “proportional” to E ’s production provides a principled reason for taking M ’s efficacy *vis-à-vis* E to be distinctively different from P ’s.

Another case for taking the having of a distinctive power profile to be a basis for M ’s causal autonomy, which may be in place even if a given entity is only singly-realized, reflects that the distinctive power profiles of entities are associated with distinctive systems of laws governing the entities. Plausibly, systems of laws track causal joints in nature—in particular, more or less abstract causal joints, with fundamental physical laws tracking a causal joint which is highly specific and sensitive to micro-physical details, and special science laws, ranging from chemistry to biology and psychology and beyond, tracking causal joints which 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 which tracks a causal level of grain which is broadly

independent of micro-physical details, independent of whether 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 relevant *qua* the (qualitative, intentional) mental event it is, in virtue of M 's distinctive power profile, taken as a whole, tracking the comparatively abstract causal level of grain associated with psychology.

Again, the suggestion here is that there are two ways for a mental event M to be causally relevant—two ways for a mental event to be efficacious *qua* the type of mental event it is—as compared to the physical event upon which it 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. Such a posit would violate *Physical Causal Closure*, and so is incompatible with physicalism. Another way—that at issue in the powers-based strategy for accommodating non-reductive 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 effect's production.

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

3.1.1 Functional accounts

First, consider functionalist accounts, according to which realized 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 so 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 those 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) 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” (81); see also Clapp 2001. A mereological account of realization is promising for non-reductive physicalist purposes, insofar as proper parts of more fundamental aggregates are distinct from and yet in a sense nothing over and above these aggregates, and may be efficacious without inducing overdetermination, as when both I and my eye cause a wink.¹⁸ Does a mereological account satisfy *SCCP*? It must, if a given mental event *M* is to be guaranteed to be casually efficacious (at all); for on some accounts of features, these have non-causal aspects (‘quiddities’), in which 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 non-reductive 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 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. (78–9)

When any such feature is multiply realized, its realizing types will share all the powers of the realized type, but will differ with respect to further powers. This type-level 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 non-reductive realization in terms of the determinable/determinate relation, the relation of increased specificity paradigmatically holding between colors and their

¹⁸The example is inspired by remarks in Paul 2002.

shades. Yablo (1992) 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 objects 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 (259).

Yablo anticipated that many might find his suggestion puzzling; but as suggested in 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 2009).

To start, consider a patch that is red, and more specifically scarlet. Sophie the pigeon, trained to peck at any red patch, 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 to a token power of its determining scarlet instance. Moreover, given Sophie's training, she would have pecked even if the patch had been a different shade of red (burgundy, say); but not so for Sophie's cousin Alice, 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 those of their determinate types. This relation will hold, in turn, between token powers of determinable and determinate instances; hence a determinable/determinate account of realization satisfies *SCCP*.

Summing up: a wide variety of apparently diverse accounts of non-reductive realization plausibly 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 non-reductive physicalist response to the problem of mental causation, this is primarily thanks 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 or 'positive compatibilist'¹⁹ responses to the problem implement a strategy that is structurally similar to that implemented by non-reductive 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 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 those that in fact determine the outcome of the choosing.

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

¹⁹See note 10: again, a positive compatibilist aims to give a positive account of the compatibility at issue; and here we include among the compatibilist accounts of interest those which are neutral on the status of *Causal Determinism* as actually true.

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 non-empty 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, 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 $\{C\}$ which are a non-empty proper subset of the total causal antecedents of M , and (ii) it is possible that a choosing M' of the same type as M occur, having relevant antecedents $\{C'\}$ of the same type as $\{C\}$, but where the *total* antecedents of M' are such as to completely determine the outcome of M' ²⁰ 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 *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. (c.f. Ayer 1954). From broader perspectives, the relevant antecedents might also include events tracking cultural influences,

²⁰As per *CCA*; the possibility of indeterministic scenarios is not to the point here.

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. (193–4)

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 undetermined. Moreover, the assumption that the subset of antecedents leaves the choice undetermined plausibly entails (indeed, has as its content) that it is possible that a choosing M' of the same type as M occur, 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 φ s 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 so? To start, note that if, for example, I am brainwashed with beliefs and desires leading to my 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—e.g., 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 they had; hence underdetermined that the agent would identify with the first-order beliefs and desires leading to the agents 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, here we again see the 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 relevant *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 point out, this approach too 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 (197)

The relevant antecedents in this case would then include those relevant to determining whether the agent was deliberating under conditions where they would, by the lights of the prevailing system of law, morality or interaction, be held responsible for the outcomes of their 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 account implements a proper subset strategy, encoded in satisfaction of *SCCA*.

Summing 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 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 to the extent that these accounts are promising so far as providing a response to the problem of free will compatible with *Causal Determinism*, this is primarily thanks to their operating to ensure satisfaction of the proper subset condition expressed in *SCCA*, and the associated basis for accommodating the efficacy of mental choosings *qua* free in the relevant (antecedent) sense. As such, a wide variety of apparently diverse compatibilist accounts are at a structural level more similar than they appear.

3.3 The parallel between compatibilism and non-reductive physicalism

Non-reductive 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, these are powers; in the other, these are causal antecedents.²¹ 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 provides a basis for showing that M is causally

²¹The compatibilist strategy can be massaged so as to also involve a proper subset relation between powers associated with two complex events corresponding to, first, the total antecedents of a mental choosing M , and second, the relevant antecedents in combination with M . But making out the structural parallel doesn't require that the strategies be exactly the same, so we won't pursue this further exercise.

relevant to the production of the effect E in question, *qua* the type of mental event M is.

This structural similarity may provide the basis for addressing a concern about the soft determinist's strategy—namely, that identification of a given subset of causal antecedents won't make sense of how the choosing could be free, since the mere presence of a subset of antecedents doesn't establish that M 'selects' or 'determines' the outcomes of the choosing.²² Here we are inclined to think that the parallel to the strategy implemented in the mental causation debate may be useful, as a case-in-point of how connecting the free will and mental causation debates may be fruitful.

To start, the soft determinist, like the non-reductive 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 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 non-reductive 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 may the soft determinist 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 distinctive form of causal relevance identified by non-reductive physicalists 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?

²²Thanks to a referee for raising this concern.

2. 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 non-reductive 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 non-reductive physicalism is useful in revealing that apparently diverse soft determinist/positive compatibilist accounts, like apparently diverse non-reductive physicalist accounts, are deeply structurally similar.
5. This parallel is also useful in revealing that soft determinist/positive compatibilist and non-reductive physicalists accounts are unified in implementing structurally similar ‘proper subset’ strategies for resolving their respective problems of mental causation.

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

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*, 84:654–674.
- Ayer, Alfred J., 1954. “Freedom and Necessity”. In *Philosophical Essays*, 271–284. London: Macmillan, London and Basingstoke. Reprinted in ?, p. 15–23.
- Bernstein, Sara, forthcoming. “Overdetermination Underdetermined”. *Erkenntnis*, 1–24.
- Bohm, David, 1952. “A Suggested Interpretation of Quantum Theory in Terms of “Hidden Variables”, Parts I and II”. *The Physical Review*, 85:166–179, 180–193.
- Churchland, Patricia S., 1986. *Neurophilosophy: Toward A Unified Science of the Mind-Brain*. 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*, 98:111–136.
- 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*, 40:295–6.
- Fodor, Jerry, 1974. “Special Sciences (Or, The Disunity of Science as a Working Hypothesis)”. *Synthese*, 28:77–115.
- Fodor, Jerry, 1987. *Psychosemantics*. Cambridge: 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*, 56.4:191–201. Hawthorne originally published this article under the name ‘John O’Leary-Hawthorne’.
- Holbach, Paul-Henri, 1770. *The System of Nature: Or, Laws of the Moral and Physical World*. New York: B. Franklin. Translated by H.D. Robinson.
- Horgan, Terence, 1989. “Mental Quausation”. *Philosophical Perspectives 3: Philosophy of Mind and Action Theory*, 47–76.
- Kane, Robert, 2002. “Introduction”. In Robert Kane, editor, *The Oxford Handbook of Free Will*. Oxford: Oxford.
- Kim, Jaegwon, 1989. “The Myth of Nonreductive Materialism”. *Proceedings and Addresses of the American Philosophical Association*, 63:31–47. Reprinted in ?.
- Kitcher, P. S., 1984. “In Defense of Intentional Psychology”. *Journal of Philosophy*, 81:89–106.
- Lewis, David, 1981. “Are We Free to Break the Laws?” *Theoria*, 47:113–21.
- McKenna, Michael and Justin Coates, 2008. “Compatibilism”. In Edward Zalta, editor, *Stanford Encyclopedia of Philosophy*. URL = <http://plato.stanford.edu/archives/sum2015/entries/compatibilism/>.
- Melnyk, Andrew, 2006. “Realization-based Formulations of Physicalism”. *Philosophical Studies*, 131:127–155.
- Melnyk, Andrew, 2008. “Can Physicalism be Non-reductive?” *Philosophy Compass*, 3:128–1296.
- Morris, Kevin, 2013. “On Two Arguments for Subset Inheritance”. *Philosophical Studies*, 163:197–211.
- Ney, Alyssa, 2008. “Defining Physicalism”. *Philosophy Compass*, 3:1033–1048.
- O’Connor, Timothy, 2002. “Free Will”. *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 *Art, Mind, and Religion*, 37–48. Pittsburgh: University of Pittsburgh Press.
- Shoemaker, Sydney, 1980. “Causality and Properties”. In Peter van Inwagen, editor, *Time and Cause*, 109–35. Dordrecht: D. Reidel.
- Shoemaker, Sydney, 2000/2001. “Realization and Mental Causation”. In *Proceedings of the 20th World Congress in Philosophy*, 23–33. Cambridge: Philosophy Documentation Center. Published in revised form in ?, 74–98.
- Steward, Helen, 2015. “What is Determinism”. *Flickers of Freedom blog*.
- 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*, 145:149–169.
- Wilson, Jessica M., 2011. “Non-reductive Realization and the Powers-based Subset Strategy”. *British Journal for the Philosophy of Science*, 94:121–154.
- Yablo, Stephen, 1992. “Mental Causation”. *The Philosophical Review*, 101:245–280.