

# Overdetermination Underdetermined

(Penultimate Version; final version forthcoming in *Erkenntnis*)

Sara Bernstein

Nonreductionists about minds and material objects think they have faced down the challenges posed by widespread causal overdetermination. Famously, Kim<sup>1</sup> attacked nonreductive physicalism about the mental on the grounds that it would involve ubiquitous overdetermination of physical effects by mental and physical properties. More recently, Merricks<sup>2</sup> objected to views that countenance material composition and constitution on the grounds that effects of ordinary objects would be systematically overdetermined by the things that make them up.

Two major lines of defense have emerged in response to the accusation of widespread overdetermination: denial and acceptance. Deniers argue that cases like mental causation and causation by particles and their sums don't really count as overdetermination, since these cases don't share all the features of "traditional" cases involving, for example, multiple rock-throwers.<sup>3</sup> On the other hand, overdetermination accepters hold that it is okay to admit widespread overdetermination, since it is not "the problematic kind."<sup>4</sup> Accepters argue that the type of systematic overdetermination generated by nonreductive theories of minds and objects is different and less worrisome than the variety involving multiple rock-throwers, and poses little threat to nonreductive metaphysics. Both lines of defense employ differences between types of causal overdetermination to argue that the type generated by nonreductive metaphysics is not to be feared.

This paper throws down the gauntlet anew. I argue that attention to the differences between types of overdetermination discharges very few burdens for the nonreductive metaphysician.

Here's the plan. First, I give a taxonomy of the various types of causal overdetermination, so that we can talk more carefully about the differences between the types. Second, I articulate some heretofore ignored problems with overdetermination, and argue that attention to differences between the types doesn't address them. Third, I show

that a prominent recent strategy for avoiding overdetermination—*compatibilism*—changes the subject rather than solves the problem.

### 1.1 What is Causal Overdetermination?: A Surface Answer

Here is a typical example of causal overdetermination: Billy and Suzy each throw a rock through a window, and either rock alone could have shattered the window. It's a case of causal overdetermination (hereafter, just "overdetermination") because there are multiple distinct causes, each of which are sufficient to bring about the effect.

Overdetermination is to be distinguished from *preemption*, in which one cause prevents the other cause from bringing about an effect, either through one causal chain hindering the other (early preemption) or one causal chain interrupting the other by bringing about the effect (late preemption). Overdetermination is different from *joint causation*, in which each cause is necessary to bring about an effect in the circumstances.<sup>5</sup>

From these ideas, we can formulate a working notion of overdetermination:

(Surface Answer) Causes  $c_1$  and  $c_2$  *causally overdetermine* an effect  $e$  if  $c_1$  and  $c_2$  are distinct, they actually occur, and they are individually sufficient to cause  $e$  in the way that it occurs.

According to Surface Answer, there is overdetermination when there are two separate, actual causes, and each alone can bring about a particular effect in the way that it would have occurred had only one cause been present. The anatomy of an overdetermination case has three parts: multiple actual causes, multiple complete causal processes attached to the causes, and the effect.

We move from Suzy and Billy to central questions of ontology because nonreductive theories of minds and objects are committed to things that make up other things, and to the causal efficacy of both maker and makee. Nonreductive theories of objects, for example, hold that there is a rock in addition to the sum of particles that makes it up, and that both the rock and the particles are causally efficacious. Nonreductive physicalists about the mental hold that the physical makes up the mental, and that both the physical and the mental are causally efficacious.

(Surface Answer) is a good working notion of overdetermination because it

encompasses many sorts of cases:

(a) **Two Rocks Overdetermination.** Suzy and Billy each throw a rock at a window at the same time and at the same speed; both rocks hit the window at the same time; the window shatters. Suzy's rock ( $c1$ ) and Billy's rock ( $c2$ ) are each sufficient to cause the shattering of the window.

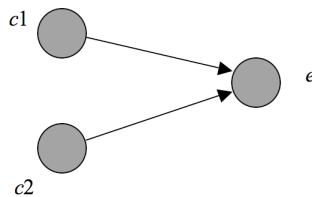
(b) **Mental/ Physical Overdetermination.** Suzy throws her rock. Suzy's mental property  $M$  (her intent to throw the rock) and the physical property  $P$  (the neural state that underlies the intent to throw the rock) are each sufficient to bring about the shattering of the window.

(c) **Mereological Overdetermination.** The single rock that Suzy throws is composed of parts. Both the rock, and the parts that compose the rock, are sufficient to bring about the shattering of the window.

(d) **Constitutive Overdetermination.** The rock is made of a sum of minerals. The sum of minerals materially constitutes, but is not identical to, the rock. Both the sum, and the rock, cause the shattering of the window.

(e) **Quantitative Overdetermination.** Only 1/2 pound of rock is required to shatter the window. Suzy's rock is 1 pound. Each 1/2 pound of the rock causes the shattering of the window.

(f) **Sanitized Overdetermination.** Neurons  $c1$  and  $c2$  each fire at precisely the same time and with precisely the same speed. Either neuron causes  $e$  to fire in exactly the same way: removing either  $c1$  alone or  $c2$  alone would not change whether or how  $e$ 's firing occurs.



Even before we can agree on a working notion of overdetermination, the controversy begins: Are all of these cases overdetermination? Or only some of them?

Those with a taste for parsimony are suspicious of calling (b), (c), and (d) overdetermination. Some hold that (b) is widespread, but unproblematic. Others deny that it is a case of overdetermination at all. Even using our working definition above as a guide, there is intense debate about whether each of the specific cases above qualifies as overdetermination.

First, let us begin by noting that there is a general technical sense of “overdetermination” that just means “causal redundancy,” and it applies to all of the cases listed above. The source of confusion is that the word “overdetermination” is sometimes used as a merely technical term that denotes causal redundancy, and other times used as a pejorative term for *worrisome* causal redundancy.<sup>6</sup> In the debate over the counterfactual account of causation, for example, “overdetermination” labels any situation in which there are multiple causes sufficient to bring about an effect in the way that it occurs. But in the debates on mental causation and ordinary objects, “overdetermination” is often used to denote a kind of *worry* or *discomfort* about causal redundancy—causal redundancy *plus something else troubling about it*. Dialectical moves that argue “This isn’t *really* overdetermination” or “This may be overdetermination, but it is not the *problematic* kind” often combine the technical and pejorative senses of “overdetermination”. For this reason, claiming that the label “overdetermination” doesn’t apply isn’t helpful if the claim doesn’t include analysis of the discomfort surrounding the label.

Commonly, nonreductive metaphysicians of minds and objects hold that they are not committed to “the bad kind” of overdetermination because their kind differs from Two Rocks in metaphysically important ways. But without an account of what the problems with overdetermination *are*, it is unclear what work these differences do in easing the discomfort with it.

Disentangling these issues will take several steps. First, we must clearly articulate the precise differences between types of overdetermination (in the strictly technical sense of the term) rather than simply gesturing at these differences. Next, we must clarify what overdetermination “in the pejorative sense” *means* by looking at the problems with overdetermination. Finally, we must examine what the appeal to the differences between types of overdetermination achieves with respect to solving these problems.

## 1.2 Causal Overdetermination: A Taxonomic Framework

To articulate the differences between types of overdetermination-in-the-technical-sense, the first step is to look at Surface Answer as a *schema*, rather than a fully specific

definition. The schema is:

(OD Schema) Causes  $c_1$  and  $c_2$  overdetermine an effect  $e$  if (i) they are **distinct**, (ii) they **occur**, and (iii) they are each **sufficient to cause  $e$  in the way it occurs**.<sup>7</sup>

The schema has four open definienda: what it is for causes to be distinct, what it is for causes to occur, what it is to be sufficient to cause something else, and what it is for the effect to be the same as if only one cause had been present.

By filling in the open definienda of the overdetermination schema, we can sort the different species of overdetermination based on various axes of difference:

**Distinctness.** Causes are not distinct *simpliciter*; they are distinct in specific ways. Causes can be mereologically distinct, modally distinct, numerically distinct, or analytically distinct (to name only a few).<sup>8</sup> Two things are mereologically distinct if they share no spatiotemporal parts; modally distinct if there can be one without the other; numerically distinct if they have different properties; and analytically distinct if the existence of one cause doesn't analytically entail the existence of the other.<sup>9</sup> Numerical distinctness is the most general kind of distinctness: it simply amounts to nonidentity. Numerical distinctness is thus implied by all other types of distinctness.

Different attitudes about what sort of distinctness is relevant to overdetermination can help us understand the inclination to treat the foregoing cases differently. The intuitive difference between Two Rocks and particles/ sum sorts of cases is that Billy and Suzy's rocks are distinct in a way that the rock and the sum of particles are not. We can imagine using a pair of event scissors to "snip out" Billy's throwing of the rock, leaving Suzy's rock to shatter the window, but we can't easily do so in the case of the particles and the rock: if we snip out the particles, it seems, we also snip out the rock. This difference reflects different sorts of distinctness between the causes. Suzy's and Billy's rocks are numerically, mereologically, modally, and analytically distinct, whereas the particles and the sum are only numerically distinct. The mental property is numerically

distinct from the physical property, but (arguably) not mereologically or modally distinct.

Cases of overdetermination by numerically distinct causes are extremely widespread. Since there is numerical distinctness between objects wherever there is property difference, the outcome caused by a single ordinary object is overdetermined by its quarks, atoms, and every “lower level” sum of entities that make it up. Numerical distinctness overdetermination is everywhere: in every effect caused by objects that differ modally from the sums that make them up, every effect caused by a mental property that is different from its physical property realizer, and every effect caused by a sum of parts that differs from what it composes.<sup>10</sup>

Given that a commitment to widespread overdetermination is sometimes taken as a *reductio* of nonreductive theories, it is easy to see why proponents of those theories would respond by trying to restrict the relevant notion of overdetermination to exclude cases where the causes are only numerically distinct. However, some strategies mean to deny overdetermination-in-the-technical-sense<sup>11</sup>, and other strategies mean to deny overdetermination-in-the-pejorative-sense.<sup>12</sup> I will examine these strategies at length in the next section.

**Occurrence.** Billy’s and Suzy’s rocks exist independently: they’re both located in spacetime and neither is ontologically derivative of the other. But how about the rock over and above its sum of particles, or over and above its parts? Does it “occur”?<sup>13</sup> Constitutive overdetermination and mereological overdetermination arise from the acceptance of an extra object in addition to constituting or composing parts. Mental/physical overdetermination usually arises due to the acceptance of an extra set of causal powers responsible for efficacy of the mental *qua* mental.

Extra loci of controversy arise with these three kinds of putative overdetermination. For constitutive and mereological overdetermination, the key question is whether the composite object has an extra set of causal powers. This issue is intertwined with the question of whether existence is wedded to causal efficacy. If what it is for an object to exist is for it to have certain causal powers associated with it, then

accepting the particles in addition to the sum generates overdetermination. With mental/physical property overdetermination, the locus of controversy is whether the mental can have powers at all. Nonreductive metaphysicians who already accept objects in addition to sums, or mental properties in addition to physical properties, admit that these dependent entities “occur”.

**Causation.** Different types of causation generate different types of overdetermination. If causation is a transfer of conserved quantity (or “oomph”) from one thing to another, then two causes overdetermine an effect if each quantity transferred is sufficient to bring about an effect. (Ned Hall (2004) calls this and related types of causation “production”.) If causation is a matter of counterfactual dependence between cause and effect, then multiple counterfactual dependencies create overdetermination (or joint causation).<sup>14</sup> If causation is a matter of lawful regularity, then multiple coincident regularities generate overdetermination.

I will not discuss the metaphysics of causation in detail here, but I will note that discussions of mereological overdetermination often assume that the type of causation in question is “oomph” causation. The puzzle of overdetermination is framed as the following: if there is an extra object on top of the parts, why aren’t there double forces transferred from, say, the rock to the window? Some hold that mereological overdetermination and mental/physical causation are rendered less strange by accounts of causation in which energy transfer does not play a central role. I am not convinced, but it is worth noting that some such as Loewer (1997) hold that using a counterfactual account of causation rather than an “oomph” theory of causation vindicates the efficacy of the mental. In any case, I will limit the present discussion to views of mental causation that involve “oomph”-style causation.

**Nature of the effect.** When Billy and Suzy throw their rocks at the window, the shattering of the window occurs differently than if only one rock had been thrown. Perhaps the window shattering via Suzy and Billy has a property—say, shattering with

extra force—sufficient to render the effect a case of joint causation. A view that permits one to avoid counting Two Rocks as a case of overdetermination is *modal fragility*. According to Lewis (1986b), we can call an event fragile if “it could not have occurred at a different time, or in a different manner.” A similar view takes events to be *fine-grained*. This view involves a very fine level of granularity at which we count events. Kim (1976) holds that events are individuals that instantiate a property at a time. There is an event for every individual, every property, and every time: my delicately buttering a piece of toast is a different event than my buttering the toast. On this view, it takes both causes to bring about the exact group of properties that occur. These views are unified by the idea that events are very precise: the extra cause makes a difference to the exact way that the effect occurs, and so it is a different event altogether.

On the other end of the fragility spectrum is *modal robustness*, according to which the numerically same event might have different features. A light shattering in the actual world is itself a forceful shattering in some nearby world. Each cause need only bring about the effect in roughly the same way that it would have occurred without the presence of the other in order for overdetermination to obtain. This is a sort of “impure”<sup>15</sup> overdetermination in which a singly-caused effect is the same as a doubly-caused effect in virtue of loose identity criteria for events. Two Rocks is joint causation under modal fragility but overdetermination under modal robustness.

The nature of the effect is clearly salient for Two Rocks type of overdetermination, in which removing one cause changes the way the effect occurs. But many ignore this feature with respect to constitutive, mereological, or mental/ physical overdetermination, in which removing one cause doesn’t change the way the effect occurs. The locus of controversy for this kind of overdetermination is the relationship between the causes, rather than the way that the effect occurs. As I shall discuss shortly, this omission is more important for the debate about mental/physical overdetermination than commonly assumed. Making a difference to the way the effect occurs is important for accounting for the precise causal contribution of each cause.



## 2.0 The Trouble with Overdetermination

### 2.1 Nonreductive Physicalism and Overdetermination

Now that we have a precise account of the differences between types of overdetermination-in-the-technical-sense, we are in a better position to evaluate several claims. First, we are in a better position to evaluate whether nonreductive physicalism can deny overdetermination-in-the-technical sense. Second, we are in a position to evaluate which of these types of overdetermination, if any, deserve the pejorative label.

Views on nonreductive physicalism can be sorted along these two dimensions. Deniers deny, at the outset, that their views are committed to overdetermination-in-the-technical sense. Deniers generally hold that overdetermination-in-the-technical sense is avoided because the causes are too intimately connected to be redundant. The Denier's strategy is incomplete if she does not spell out how she avoids overdetermination-in-the-technical-sense while retaining her central theoretical posits. In contrast, Accepters admit overdetermination in the technical sense of the term, but repudiate a commitment to overdetermination in the pejorative sense of the term. The Acceptor's strategy is incomplete if she does not spell out what the problems with overdetermination are, and then argue that she avoids these problems.

In this section and the next, I will argue against both Denial and Acceptance. I will argue that nonreductive metaphysicians *are* committed to overdetermination in the technical sense, contra Denial. Then I will articulate and examine the problems with this particular kind of overdetermination, and argue that it deserves the pejorative label attached to it, contra Acceptance.

First it will be helpful to clarify how I use the terms “nonreductive” and “reductive”. These terms are controversial, and there is little widespread agreement on their appropriate usage. I will not take a stand on that issue here, but rather I will use the terms in the following way. I will take the “nonreductive” in nonreductive physicalism to

apply only to views according to which the mental *qua* mental is causally efficacious. This usage of “nonreductive” rules out several views that have traditionally fallen under the label. Token-token identity views such as Davidson’s anomalous monism,<sup>16</sup> according to which mental properties are distinct from physical properties even though mental events are identical with mental events, count as reductive rather than nonreductive according to my usage.<sup>17</sup> (So, too, does Shoemaker’s “subset” account, which I discuss at greater length below.) Views such as these avoid overdetermination-in-the-technical-sense, because they only stipulate one set of causally efficacious powers to begin with.

Further, since the major controversy about overdetermination centers around mental causation, I will restrict my focus to this topic.<sup>18</sup> Nonreductive physicalists about the mental standardly accept the following claims:

(Closure) The physical world is causally closed.

(Distinctness) The mental is distinct from the physical.

(Efficacy) The mental *qua* mental is causally efficacious.

Nonreductionist accounts differ on the precise nature of the relationship between the mental and the physical: some hold that it is constitution,<sup>19</sup> others assert that it is a subset relation of causal powers,<sup>20</sup> and others argue that it is a determinate/ determinable relationship.<sup>21</sup> What unifies nonreductive physicalist theories of the mental is that they accept some sort of distinctness of mental and physical properties, and take as a desideratum the retention of the causal efficacy of the mental *qua* mental. But given these commitments, it seems that each mentally caused outcome will be caused by a mental property and its physical realizer, leading to widespread mental/physical overdetermination in the technical sense. There is an additional question about whether it deserves to be counted as “the bad kind” of overdetermination. I will treat these two issues in turn.

## 2.2 The Nonreductive Physicalist is Committed to the Technical Sense of Overdetermination

Here I will give a brief argument against Denial, or the idea that the nonreductive

physicalist can deny widespread overdetermination in the technical sense. I will argue that her commitments, together with the anatomy of overdetermination, require the acceptance of widespread redundancy.

In order to deny redundancy, one must deny that there *are* multiple causes, or deny the actual sufficiency of one of the causes. The nonreductive physicalist is committed to the existence of the mental. That leaves only a denial that one of the causes is sufficient. I see two routes for denial. Neither route is consistent with the nonreductive physicalist's commitment to efficacy of the mental.

One route is to deny the causal potential of one of the causes. By *causal potential*, I mean the potential of a property to cause anything at all. Baseballs have causal potential, while ghost baseballs do not. Roughly, for something to have causal potential is for it to be the right sort of thing to be a cause.

The other route is to deny that a cause "follows through" and actually causes anything. By *follow-through*, I mean the completion of the causal process leading from the cause to the effect, as opposed to a process that "stops short" of bringing about an effect. If Billy throws a rock so weakly that it fails to reach the window, his rock-throw lacks "follow-through."<sup>22</sup>

Consider the distinction at work in a preemption case. Billy and Suzy each throw rocks at a window. Suppose that Suzy throws her rock slightly faster than Billy, and Suzy's rock shatters the window, whereas Billy's rock flies through the space where the window used to be. Both Billy's and Suzy's rocks have *causal potential* in the sense that they each have the ability to shatter the window. But Billy's rock lacks *follow-through* in the sense that the causal process originating in his rock-throw does not reach the effect. A ghost rock, in contrast, lacks causal potential altogether.

In order to deny overdetermination-in-the-technical-sense, one can deny the causal potential of one of the properties, or deny the follow-through of one of the properties. I'll run through all of the options in logical space, and conclude that none of them are compatible with the nonreductive physicalist's commitments.

Consider the diagrams of mental causation below:

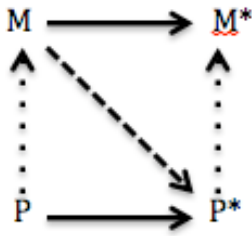


Figure 1

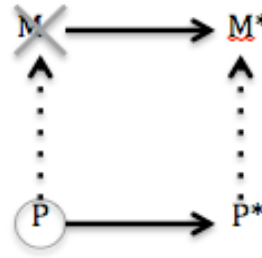


Figure 2

The thick arrows represent actual causation. The dotted arrows represent the realization relation between the physical and the mental. The diagonal dashed arrow represents downward causation, or causation from the mental to the physical; it also represents the putative overdetermining cause of every physical effect. In Figure 2 I have added circles and strikethroughs to the traditional diagram. Circles indicate causal potential, and strikethroughs represent the lack thereof. Note that, since M lacks causal potential in Figure 2, there is no downward causal arrow from M to P\*.

First, causal potential. Denying the causal potential of the mental cause is a nonstarter for the nonreductive physicalist, for a central tenet of her view is efficacy of the mental. Nor can she deny causal potential of the physical cause (P), for then the mental cause is the only cause of P\*. This commitment violates causal closure, as represented by the following diagram:

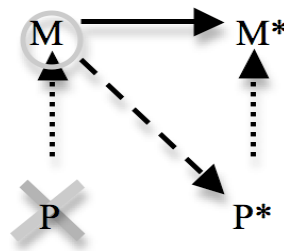


Figure 3

Both the physical cause and the mental cause, then, must have causal potential.

Perhaps she could hold that both causes have causal potential, but deny that one

cause “follows through,” as in an early preemption case in which Suzy’s rock hits the window and Billy’s rock stops short of the window. In this case, both P or M have the potential to cause P\*, but one “lazy” cause fails to follow through, like so:

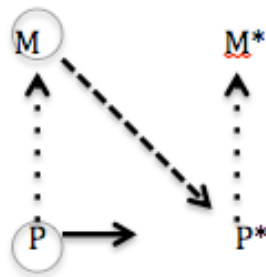


Figure 4

(Note that the shorter arrow indicates that the process running from P does not follow through to P\*.)

According to (Closure), P\* must have a physical cause. If M is the only cause of P\*, there is a violate of (Closure):

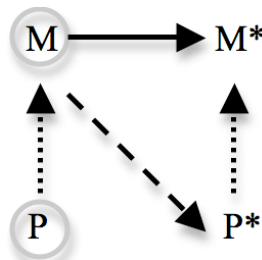


Figure 5

Thus P must follow through to P\*.<sup>23</sup>

Nor can the nonreductionist admit that the mental cause fails to follow through. On that view, the mental cause *could* cause P\*, but for some reason, it does not. The idea is that M is the right sort of thing to cause P\*, but the causal process running from M doesn’t follow through to P\*, as follows:

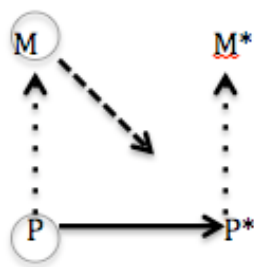


Figure 6

If the mental cause is always preempted, its causal potential is moot. Rejecting follow-through of the mental amounts to an alternative form of epiphenomenalism in which the mental has the potential to cause  $P^*$ , but is causally inert due to its lack of follow-through. This view violates (Efficacy). To retain the efficacy of the mental qua mental, the mental cause must be efficacious in practice as well as principle. Thus both causes must “follow through” in order for the nonreductionist to maintain the full causal picture that she wants.

I conclude that the nonreductive physicalist must accept overdetermination-in-the-technical-sense. Denying that the mental and physical are causally redundant involves denying the efficacy of one of the causes, either through denying its potential or denying its follow-through. Neither view is acceptable to the nonreductive physicalist, who is antecedently committed to causal closure and to efficacy of the mental qua mental. Acceptance of two causes with full causal potential and full follow-through makes overdetermination in the technical sense a *fait accompli*.<sup>24</sup> Now onto the problems.

### 2.3 Easy Trouble

When nonreductive physicalists hold that their type of overdetermination isn't problematic, what they mean is that the explanatory and metaphysical burdens that it incurs can be easily discharged. Well-known burdens of widespread overdetermination include *ubiquity* and *coincidence*.<sup>25</sup>

With respect to ubiquity, Accepters can hold that widespread redundancy is not inherently troublesome. For example, humans have two kidneys, each of which is sufficient to filter out the toxins in the body. But there doesn't seem to be a problem of

why there are two kidneys (even among evolutionary biologists). A causally crowded world alone is not an intrinsic reason to deny theories committed to widespread overdetermination.

Nor is an ontologically excessive metaphysics necessarily an incorrect metaphysics, according to the Acceptor. Parsimony is not the only criterion for evaluating metaphysical theories, or else nihilism would reign uncontroversially.<sup>26</sup> One desideratum of nonreductive physicalism is to model the intuition that the mental *qua* mental is causally efficacious while remaining naturalistically friendly. Presumably, the nonreductionist holds that successfully meeting these desiderata counterbalances the problem of metaphysical excess: nonreductive theories trade parsimony for explanatory and intuitive power. Nonreductive physicalists are better off openly accepting this tradeoff rather than denying that such an exchange is necessary.

The coincidence concern is similarly dissolved upon close examination. Mental/physical overdetermination is not coincidental: psychophysical laws explain co-occurrence. Different sorts of overdetermination yield different built-in explanations. The coincidence objection only properly applies to cases of overdetermination with modally distinct causes, like Two Rocks. Mental/physical overdetermination is not coincidental because there are rich reasons for the co-occurrence of mental and physical properties.

Now, I agree that careful attention to differences between mental/physical overdetermination and Two Rocks shows that the nonreductionist does not accept a world of unexplained coincidence. Nor is ubiquity a problem. These are “easy” troubles that are dissolved upon close examination. If the pejorative sense of overdetermination just means “overdetermination that is ubiquitous” or “overdetermination with unexplained coincidence,” then the Acceptor’s strategy is successful: she has articulated the problems with overdetermination, and shown that her type of overdetermination does not suffer from these problems.

But these easy problems are not the end of the story for the nonreductive physicalist. For it seems unlikely that the pejorative sense of “overdetermination” just means “overdetermination that is not coincidental” or “ubiquitous overdetermination”.

Suppose that Billy is an uncertain rock-thrower and will only throw his rock at the window if Suzy throws hers. Suzy throws her rock and Billy throws his. In this case, the occurrence of multiple sufficient causes is not coincidental, since Billy's rock-throw depends on Suzy's rock-throw. But, intuitively, this case is less suspicious than mental/physical overdetermination.

It also seems unlikely that the pejorative sense of "overdetermination" just means "redundancy that is widespread". Suppose that every mischievous sibling-pair on earth conspires to throw rocks at their respective windows at precisely the same time. Intuitively, this is a case of widespread overdetermination, and it is *weird*, but it does not invoke the same metaphysical discomfort as mental/physical overdetermination.

For these reasons, I do not think that coincidence and ubiquity are the real sources of discomfort. Rather, I think that the nonreductive physicalist faces a deeper, unarticulated set of troubles that cannot be treated by attention to differences between Two Rocks and mental/physical causation. As I will now argue, the discomfort stems from unexplained differences between the different sorts of overdetermination.

### 2.3 Hard Trouble

Return to Two Rocks. A distinctive feature of this case is that removing one rock makes a difference to the way the shattering occurs. There is a precise explanation for this difference between the one rock and two rock scenarios: each rock contributes a specific force to the shattering of the window, such that removing one rock removes a force from the shattering of the window. But mental/physical overdetermination has no such difference, since, holding the physical cause fixed, removing the mental cause does not change the way the mentally caused outcome occurs. This difference, I suggest, is the reason overdetermination is problematic for the nonreductive physicalist: it reinforces the central tension of the nonreductive physicalist's view rather than resolving it.

Let's examine this idea in closer detail. Drawing on the two major open definienda of the overdetermination schema—types of distinctness between the causes, and how the effect occurs—we can identify two major questions about overdetermination. First, there is the question of where the extra causation in a case of overdetermination



“comes from”—that in virtue of which there is an extra causal contributor to the outcome. Second, there is the question of where the extra causation “goes”—the precise way that the extra cause contributes to the outcome. In Two Rocks, neither of these explanatory tasks is challenging: it is not mysterious where the extra causation “comes from”, since there is an extra rock. Case closed! Nor is there a mystery about where the extra causation “goes”, since the extra rock causes the window to shatter with double the force with which it would have shattered had there been only one rock. To put it simply, Two Rocks is *causally satisfying*: there is a precise account of the causal powers of both rocks, and of the individual contribution of each rock to the shattering of the window. Removing one rock-throw has an easily definable result: the window shatters with less force.

To alleviate the discomfort with overdetermination, the nonreductive physicalist must give a causally satisfying picture of the exact causal contribution of the mental cause minus the causal contribution of the physical cause, just like Two Rocks. But the nonreductive physicalist has no such picture.

More importantly, she can't provide a causally satisfying picture that is friendly to her view. The more similar mental/ physical overdetermination is to Two Rocks, the further the nonreductive physicalist strays from her titular commitments. The less similar mental/physical overdetermination is to Two Rocks, the more mysterious, derivative, and un-causation-like the mental cause seems.<sup>27</sup> With respect to where the extra causal power of the mental “comes from”, mental causal powers too independent from physical causal powers threaten physicalism, while mental causal powers too dependent on physical powers threaten nonreductionism. And with respect to where the extra causation “goes”, an easily demarcated extra force of the mental à la Two Rocks evokes substance dualism. But a mental cause whose removal makes no difference to the effect suggests that it is not a real causal contributor in the first place.

As an illustration of the first dilemma, consider the “subset strategist” who accepts distinctness of mental and physical properties while holding that the causal powers of the former are a proper subset of the latter.<sup>28</sup> This view provides an answer to the question of where causal powers come from, but at the expense of nonreductionism:

if mental powers just are physical powers, this is a closet reduction of the mental to the physical. If there is only one set of efficacious powers shared between the mental and physical properties, redundancy is avoided at the expense of causal independence of the mental. The subset strategist avoids the troublesome sort of overdetermination because she avoids the *technical* sort of overdetermination: there is only one set of causally efficacious properties to begin with.<sup>29</sup>

Similarly with the constitution theorist.<sup>30</sup> According to this view, the relationship between the physical property and the mental property is like a lump of clay and the statue that it constitutes: there is overlap with respect to a set of physical properties, but a difference with respect to modal and relational properties.<sup>31</sup> Here again, there isn't overdetermination-in-the-technical-sense because there's only one set of causally efficacious properties: that with respect to which the mental and the physical overlap. The mental *qua* mental doesn't do independent causal work.

Views that avoid causal redundancy to begin with are not nonreductive in the way that matters. For these views only satisfy (Distinctness) by holding that mental and physical properties are nonidentical, *not* by holding that the causal powers of the mental are distinct from the causal powers of the physical. According to these views, there is only one set of causal powers that the physical and mental share. But shared causal powers do not vindicate efficacy of the mental *qua* mental, because the mental itself is not the source of the efficacy. Either the nonreductive physicalist accepts that mental causal powers are distinct from physical causal powers, and thus accepts the technical sense of overdetermination; or she denies the efficacy of the mental *qua* mental and strays from her nonreductionism. Strategies that only posit one set of causal powers avoid overdetermination, but at the cost of the independent efficacy of the mental.

The nonreductive physicalist faces a different but equally intractable dilemma with respect to the question of where the extra causation *goes*. To get a handle on this problem, consider that Two Rocks features *additivity* of causes: both rocks contribute individual forces to the shattering of the window, so that it shatters differently than if only one rock had been thrown.<sup>32</sup> Additivity of this sort can be summed up by the simple slogan: double the causation, double the force of the effect. There is no mystery about

where the extra causation “goes”, because the extra cause is straightforwardly additive.

But mental causation for the nonreductive physicalist isn't additive like Two Rocks: the individual causal contribution of the mental cause doesn't contribute an extra force or “oomph”. On the one hand, a view according to which the causal contribution of the mental cause *is* easily demarcated à la Two Rocks is suspiciously dualistic, and thereby threatens the physicalist's commitments. If the causal power of the mental is sufficiently independent that it creates an extra *oomph*, it is hard to see how mental causation remains friendly to physicalism. On the other hand, a view according to which the mental cause isn't additive in the same way as the physical cause endangers the independent efficacy of the mental *qua* mental: if the causal contribution of the mental cause to the outcome isn't easily demarcated or identifiable, what makes it real causation? If mental causation is real causation, where does the extra causation go? Why aren't physical outcomes caused with double the *oomph* with which they would have been caused had only the physical cause been present?

The discomfort lies in the fact that mental/ physical overdetermination isn't enough like Two Rocks. Two Rocks is straightforward, satisfying, and unmysterious: two rocks, two forces, double the force of the shattering. But mental/physical overdetermination is causally unsatisfying. The differences between Two Rocks and mental/ physical overdetermination only reinforce this tension.

Now, I do not think that overdetermination in the pejorative sense just means “redundancy without a causally satisfying explanation for the origin and contribution of each cause.” Rather, I contend that the argument over the pejorative sense of overdetermination reflects the deeper unresolved discomfort with the dissimilarity between Two Rocks and mental/physical overdetermination. The idea is that Accepters try to show that their sort of overdetermination is unproblematic because it is unlike Two Rocks. But they would be well-served if mental causation were rather *more* like Two Rocks, at least along the dimension of efficacy of the mental cause.

A word of clarification regarding the dialectical role of the Two Rocks example. Accepters hold that its dialectical role is to model the dissimilarities between the two-

rock-throw type of overdetermination and mental/physical overdetermination. I maintain that the dissimilarities—ubiquity and coincidence—are not as important as the nonreductive physicalist thinks. In addition, I assign Two Rocks the second dialectical role of showing that the nonreductive physicalist would be better off with *more* similarities between her type of overdetermination and Two Rocks, particularly with respect to how causally satisfying it is.

To see these points more clearly, consider the following thought experiment. Imagine a world in which a ubiquitous Overdetermination Demon causally overdetermines every brain state. And suppose, further, that the Overdetermination Demon's individual causal contribution to each mentally caused outcome is additive, easily observable, and cleanly demarcated. For example, suppose that when my desire for coffee causes me to fetch the coffee, the Overdetermination Demon also causes me to fetch the coffee. And suppose that, by carefully examining my mental processes, scientists are able to ascertain the exact causal contribution of the Overdetermination Demon to my coffee-fetching in exactly the way that we are able to ascertain the causal contribution of a single rock in Two Rocks. If the demon stops doing his work, then each mentally-caused outcome occurs differently.

This scenario does not generate the worry that the causal powers of the Overdetermination Demon crowd out the causal powers of brain states.<sup>33</sup> Demon World just shows that *if* we are already committed to the efficacy of brain states and the overdetermination demon, then Demon World is full of overdetermination in the technical sense. Demon World is causally satisfying because the extra causal power arising from Overdetermination Demon is pre-established, as is the causal contribution to each mentally caused outcome.

But no such comfort exists with mental/physical overdetermination. Contrasting Demon World with mental/physical overdetermination shows that the residual discomfort surrounding the latter is not with ubiquity, coincidence or causal excess (all of which Demon World have), but rather with the lack of clarity surrounding the capacity and contribution of the mental cause. If nonreductive physicalism can vindicate the mental *qua* mental while staying true to physicalism, then overdetermination will not be the

problem. Arguing about overdetermination only changes the subject from the greatest challenge of nonreductive physicalism—the inherent instability of the view—to the lesser problem of ontological excess. But excessive metaphysics was not the problem in the first place.

Several objections might arise. First, one might hold that I have performed a bait-and-switch: I have promised to give a detailed account of overdetermination-in-the-pejorative-sense, but instead I have given an independent argument about the instability of nonreductive physicalism.

Here it is worth pausing to recapture the overall project: to show that differences between types of overdetermination do not help the nonreductive physicalist in the way she imagines. By digging deeper than the commonly considered problems about overdetermination-- ubiquity and coincidence—it is clear that there is still a discomfort with overdetermination *that has nothing to do with those two problems*. The source of discomfort around mental/ physical overdetermination is the metaphysical unclarity surrounding the contribution of the mental cause. This discomfort reflects the underlying instability of nonreductive physicalism.

Second, one might hold that I have placed too high a demand on the nonreductive physicalist. Perhaps, one might hold, it is enough that mental and physical causes are numerically distinct but overlap with respect to their causal powers.

I do not wish here to argue about what it takes to be a nonreductive physicalist. Rather, I think I have done enough to establish that the nonreductive physicalist should be uncomfortable with the difference between her sort of overdetermination and Two Rocks, for reasons having to do with her ontology of the mental. The burden is on the nonreductive physicalist to provide a causally satisfying picture of causation by the mental *qua* mental, and she does not do so by holding that there is one set of causal powers with respect to which the mental and the physical overlap. The idea is that accepting overdetermination as unproblematic doesn't at all vindicate the efficacy of the mental *qua* mental.

### 3.0 Compatibilism About the Mental: A Case Study in Overdetermination Denial

Many nonreductive physicalists attempt to dodge these “hard troubles” by arguing that a physical cause without a mental cause is metaphysically impossible. The suggestion is that we cannot evaluate the difference-making capacity of one cause without the other, since there cannot be one without the other. But, as I will now argue, this strategy doesn’t give a causal account so much as change the subject. I will focus on the most prominent of these strategies, *compatibilism*, and argue that it fails. I will argue that the compatibilist doesn’t avoid overdetermination in the technical sense or in the pejorative sense. Since the compatibilist does not distinguish between these two senses of overdetermination, I will proceed first by testing for the technical sense, and then by testing for the pejorative sense.

Compatibilism is driven by the idea that mental properties and their physical realizers are too tightly connected to be overdeterminers. (Note that this sort of compatibilism is entirely different than the “compatibilism” of the free will debate.) The compatibilist accepts the standard commitments of the nonreductive physicalist—(Closure), (Distinctness), and (Efficacy)—but adds the stipulation that one can hold those commitments while remaining uncommitted to causal overdetermination. The compatibilist could summarize her position thusly: “It is true that I posit distinct, individually sufficient, efficacious causes for a single event. But this doesn’t count as *overdetermination*, because the two causes aren’t distinct in the kind of way that overdetermination needs.”

Bennett’s (2003) compatibilism focuses on the modal relationship between causes as a strategy for denying mental/physical overdetermination. Her idea is that overdetermination requires causes that are modally distinct. Two things are modally distinct when there can be one without the other.<sup>24</sup> The mental and the physical are not modally distinct, Bennett argues, and hence the mental does not overdetermine the physical.

Bennett proposes the following test for causal overdetermination: *e* is overdetermined by the mental and physical only if both of the following counterfactuals

are non- vacuously true:

(O1) if m had happened without p, e would still have happened.

(O2) if p had happened without m, e would still have happened.

Roughly speaking, a counterfactual is vacuous if its antecedent is metaphysically impossible. Since the antecedent of (O2) appeals to a metaphysically impossible world, the counterfactual is vacuous. Based on the necessitation of the mental by the physical, Bennett holds that there is no world where  $c_1$  occurs without  $c_2$ . Thus (O2) is vacuous. She writes:

“If one of the causes guarantees the existence of the other, there is no issue about skipping over some worlds to get to one where the antecedent of the relevant overdetermination counterfactual holds. There are no further worlds to skip to. To put the point more formally: if one of the causes necessitates the other, if it is at least metaphysically impossible for the one to occur without the other, then one of the overdetermination counterfactuals will come out vacuous. And there is something to be said for the idea that the vacuity of one of them means that the effect is not overdetermined.” (Bennett, 2003)

To get a better grip on this claim, imagine that a baseball is  $c_1$  and a plurality of particles that makes it up is  $c_2$ . If there is no possible world where there is a plurality of particles without a baseball, one of the counterfactuals is vacuous, and so the baseball and the plurality of particles do not overdetermine the effect. A mental property and the physical property that realizes it, Bennett maintains, are related in roughly the same way: since there is no possible world where there is a physical property realizer without a mental property, (O2) is vacuous. Thus there is not overdetermination.

Bennett’s strategy involves biconditionalizing on modal distinctness, i.e., holding that two things are eligible to be overdeterminers *only if* there can be one without the other. Using the schema framework, we thus have the following notion of overdetermination:

(Modal Distinctness OD) Causes  $c_1$  and  $c_2$  overdetermine an effect  $e$  **iff** (i)  $c_1$  and

$c_2$  are **modally distinct**, (ii) they occur, and (iii) they are each sufficient to cause  $e$  in the way that it occurs.<sup>34</sup>

The idea of (Modal Distinctness OD) is that modal independence between causes is a necessary condition on the occurrence of overdetermination. Applied to mental causation: suppose that my being in pain causes me to take an aspirin. My being in pain is realized by a neural state. According to the modal distinctness biconditionalization, if there is no possible world where there is a neural state without the pain state, then my pain and my neural state are not overdeterminers of my taking an aspirin.

### 3.1 Compatibilism does not Eliminate Overdetermination-in-the-Technical-Sense

Since the modal relationship between the causes does not bear on the efficacy of either cause, compatibilism does not avoid the technical sense of overdetermination.

Consider the following counterexample to (Modal Distinctness OD):

(Alarm) An alarm will sound if it detects the presence of the chemical dichromate or anything colored orange. A bit of dichromate (which is necessarily orange) appears. The alarm sounds.

A weak modal relationship holds between the chemical and the color: due to its chemical makeup, dichromate is always orange. But the modal indistinctness between the causes does not eliminate overdetermination in the technical sense: both the chemical and the color are individually sufficient to bring about the sounding of the alarm. Intuitively and metaphysically, each cause is separate. That two causes are modally bound doesn't itself have any bearing on the separate causal powers of the chemical and color. The relationship between causes is irrelevant to their status *as causes* given antecedent commitment to their efficacy. Similarly, the nonreductive physical's antecedent acceptance of two causes drives her to overdetermination in the technical sense.

Note that it is not the compatibilist's appeal to *modal* indistinctness per se that is unsatisfying. Rather, no appeal to the relationship between the causes eradicates either cause. Consider the following example:

(Assassin Twins) Twins in the assassination business go everywhere together. Each twin shoots Victim with a separate gun. Either shot would have been sufficient to kill Victim.



Here, the two twins always co-occur in the actual world.<sup>35</sup> But their co-occurrence doesn't mitigate their causal independence nor their individual causal responsibility for the death of Victim.

The lesson is that a modal relationship between the causes doesn't attenuate the individual causal powers of either one. Identifying the relationship between the causes—modal or non-modal-- does not eliminate overdetermination in the technical sense of the term.

### 3.2 Compatibilism Does Not Eliminate Overdetermination in the Pejorative Sense

Compatibilism doesn't avoid overdetermination in the pejorative sense, either, since it does not resolve the discomfort surrounding the causal contribution of the mental. Rather, it changes the subject from the contribution of the mental cause to the relationship between the physical and mental causes. But compatibilism remains silent on the causal capacity of the mental. The modal relationship between the causes doesn't itself explain such a difference, since causes that are modally indistinct can still be causally independent and additive, as in the case of the dichromate and orange flame overdetermining the sounding of the alarm.

The compatibilist's strategy is causally unsatisfying because it ignores informative worlds when using the counterfactual test to evaluate whether something is a case of overdetermination. Recall that Bennett holds that the evaluation of the overdetermination counterfactuals requires us to skip over worlds where  $c1$  occurs without  $c2$ . To assess whether  $c1$  and  $c2$  overdetermine  $e$ , the compatibilist looks directly to worlds where  $c1$  and  $c2$  co-occur, skipping the metaphysically impossible worlds where  $c1$  occurs without  $c2$ . Bennett's justification for skipping over these worlds is that a counterfactual is vacuous if it appeals to impossible worlds.

But this rule exploits a mere technicality about the necessary connection between  $c1$  and  $c2$  to excuse skipping over the very worlds that are causally informative—the ones in which the causal capacity of the mental can be evaluated without the presence of the physical realizer. Even granting the metaphysical impossibility of worlds where  $c1$  occurs without  $c2$ , those impossible worlds are *most* relevant to examining the real contribution

of the mental cause.

To illustrate this idea, reconsider Two Rocks. Suppose that we want to know exactly how Suzy's rock contributes to the shattering of the window. In order to evaluate the causal capacity and contribution of Suzy's rock, we skip to the nearest possible world in which only Suzy throws her rock at the window. It is uninformative to look exclusively at worlds in which both Suzy and Billy both throw their rocks, if the explanandum is the individual causal contribution of Suzy's rock. Rather, the worlds that matter in assessing the causal contribution of Suzy's rock are the ones in which only she throws her rock. To ignore the individual-cause worlds is to change the subject away from the central explanandum.

Similarly, the informative metaphysical facts about mental causation stem from worlds where we can evaluate the causal capacity and contribution of the mental without the presence of the physical realizer. Skipping over these impossible worlds to co-occurrence worlds is equivalent to examining only the worlds in which Billy *and* Suzy throw their rocks. Co-occurrence worlds are metaphysically uninformative with respect to the causal capacity of individual overdetermining causes. Intuitively, we care about what happens when the mental cause occurs without the physical one. For the purpose of examining mental causation, the impossible worlds are the most relevant ones.<sup>36</sup>

One might be skeptical that counterpossibles—counterfactuals with antecedents that appeal to impossible worlds-- can be of use in shedding light on mental causation. I argue in Bernstein (forthcoming) that counterpossibles can be used for such a purpose, and space forbids me from mounting a full defense of this claim here. But a few remarks will help.

First, we can reason counterfactually about impossibilities in much the same way we can reason about possibilities. Consider:

If the mathematician hadn't failed to prove that  $2+2=5$ , she wouldn't be in dire financial straits.

Intuitively, this claim is non-vacuous: we can certainly make sense of it, and it also seems

true. The informative worlds are not the worlds in which the mathematician fails to prove that  $2+2=5$  (namely, all the possible ones); rather, the impossible world shows what would happen if she had proved it (namely, she would be in better financial straits). *Prima facie*, there is enough intuitive evidence to hold that counterfactuals involving impossibilities can sometimes be non-vacuous.

Second, Bennett's argument rests on the vacuity of counterpossibles. Her view is that proving the vacuity of a mental causation counterfactual shows that there is no physical/ mental overdetermination. Showing that counterpossibles can be non-vacuous in some instances (including in this instance) is a direct challenge to her counterfactual test for overdetermination.<sup>37</sup>

Third, if one accepts Bennett's counterfactual test on overdetermination, then one already accepts that mere possibilities can shed light on causation at the actual world. There is no reason not to accord impossibilities the same status, particularly when the relevant counterpossible ("If the physical property had occurred without the mental property, the effect still would occurred") is more informative than the relevant counterfactual ("If the physical property and mental property had not occurred, the effect would not have occurred.")

Now, it is a difficult and interesting question what is going on at these impossible worlds. But that's precisely the question that needs to be answered: what, precisely, the causal contribution of the mental cause absent the physical cause *is*. And note that the impossibility of the worlds does not mean "all rules are off" with respect to evaluating counterpossibles. As in normal counterfactual evaluation, we skip to the impossible worlds most similar to our own.<sup>38</sup> Thus, assuming nonreductive physicalism, the relevant impossible worlds are the same as our world in all the relevant respects, minus the presence of the physical realizer. We do not, for example, skip to dualist or eliminativist worlds, since those worlds are presumably less like our world than the nonreductive physicalist worlds not containing the physical realizer.

This technical difference about the relevance of impossible worlds reflects a deeper conceptual difference between the compatibilist and me. The compatibilist denies

mental/ physical overdetermination because she restricts the conceptual content of her view to worlds in which the mental supervenes on the physical. But I hold that the very source of discomfort around mental/physical overdetermination—the exact causal contribution of the mental cause—can be resolved only by examining the mental cause in isolation. Until we understand the exact causal capacity and contribution of the mental independent of its physical grounds, the mental/physical overdetermination that stems from nonreductive physicalism is causally unsatisfying.

#### 4. Conclusion

I have distinguished between technical and pejorative senses of overdetermination, and argued that nonreductive theories of minds and objects are committed to both. Commitment to the technical sense of overdetermination springs from the acceptance of multiple causally efficacious entities. The pejorative sense of overdetermination punishes nonreductive views that incorporate a lack of metaphysical clarity about the causal contribution of the dependent cause. Attention to differences between these sorts of cases and Two Rocks reinforces explanatory burdens rather than discharges them.

There is an independent question as to when these troubles have been relieved. The foregoing discussion makes clear that it is not when we point to mental causation and claim “that is not causal overdetermination” or “that is not *the problematic sort of overdetermination*”. The gauntlet remains.

---

<sup>1</sup> Most prominently in Kim (1993) and (2005).

<sup>2</sup> In Merricks (2001).

<sup>3</sup> Deniers include Yablo (1992) and Thomasson (2007) and arguably Bennett (2003, 2008).

<sup>4</sup> Overdetermination accepters include Loewer (2007), Pereboom (2002), Schaffer (2003), and Sider (2003).

<sup>5</sup> An astute referee points out that joint causes are not necessary to bring about their effects in cases involving indeterministic causation. For the purposes of this discussion, I will be assuming determinism.

<sup>6</sup> Sider (2003) also notes this confusion.

<sup>7</sup> A more precise formulation holds that there is overdetermination when  $c_1$  and  $c_2$  are distinct, they occur, are not part of the same causal process, and are each sufficient to cause  $e$  in the way that it occurs. Thus two causes in the same causal process at

---

different times, e.g., the big bang and a rock throw later in time, do not overdetermine an effect. For my purposes, the more precise formulation won't be necessary.

<sup>8</sup> Funkhouser (2002) and Stoljar (2008) sort overdetermination mainly on the distinctness axis. Here I sort overdetermination on four axes of difference.

<sup>9</sup> Thomasson (2007) uses analytic distinctness to argue against systematic overdetermination for ordinary objects.

<sup>10</sup> For the present purposes, I will set aside views that hold that composition is identity.

<sup>11</sup> For example, Yablo (1993), Shoemaker (2003), (2007) and Paul (2007).

<sup>12</sup> For example, Bennett (2003) and Sider (2003).

<sup>13</sup> The use of "occur" here is admittedly awkward, and stands in for deeper questions regarding the existence of objects versus the occurrence of events, and whether there is an object/ event distinction. Here I will set these issues aside.

<sup>14</sup> There is a well-known problem of overdetermined effects being left entirely without causes on the counterfactual view. The counterfactual "If  $c_1$  hadn't occurred,  $e$  wouldn't have occurred" is false because  $c_2$  would have caused  $e$ . "If  $c_2$  hadn't occurred,  $e$  wouldn't have occurred" is also false, since  $c_1$  would have brought about the effect. Here I am setting aside this problem so as not to take a stand on the truth or falsity of the counterfactual account of causation.

<sup>15</sup> So termed in Paul (2007).

<sup>16</sup> See Davidson (1980).

<sup>17</sup> Whether Davidson's anomalous monism is committed to epiphenomenalism is a matter of some controversy. Here I assume that the view does not imply efficacy of the mental qua mental.

<sup>18</sup> However, I take my suggestions to generalize to nonreductive theories of ordinary objects as well.

<sup>19</sup> For example, Pereboom (2002) and Paul (2007).

<sup>20</sup> For example, Shoemaker (2007) and Wilson (2010).

<sup>21</sup> Yablo (1992), most prominently.

<sup>22</sup> There is a large literature on causal powers and causal efficacy that I will not address here. I take causal potential and causal follow-through to encompass causal powers.

<sup>23</sup> Some hold that there is a middle ground here that doesn't violate causal closure of the physical. Yablo's (1992) *proportionality* strategy holds that every event is "causally determined" by a physical event, but that only the mental cause is efficacious. I agree that this strategy fully respects causal closure of the physical, but do not think that it provides a causally satisfying picture of mental causation. (See section 2.3).

<sup>24</sup> For other sorts of ontologists, this lesson generalizes. One who holds that macro-level objects have causal efficacy in addition to their micro-level realizers has a de facto commitment to causal redundancy.

<sup>25</sup> See Sider (2003), Schaffer (2003), and Bennett (2003, 2008) for discussions of these burdens.

<sup>26</sup> Indeed, Merricks (2001) uses widespread overdetermination as an argument for mereological nihilism. Willard (2014) argues that there is no good reason to seek simplicity in metaphysical theories.

---

<sup>27</sup> The substance dualist does not have such problems. For she can just admit that causal properties of the mental are fully independent, and accordingly draw a picture upon which the mental and physical properties are additive. According to this picture, mental causation is like Two Rocks: double the causes, double the forces, double the impact. Dualism and nonreductive physicalism share the same commitment to redundancy, but dualism can give a causally satisfying picture of this redundancy whereas nonreductive physicalism cannot.

<sup>28</sup> Subset strategists include Shoemaker (2003), (2007) and arguably Yablo (1992).

<sup>29</sup> This objection depends on a particular reading of the subset strategy. In contrast to my reading, Hoffman (2007) holds that the subset strategy does bestow causal powers twice over. He calls this the “piling problem”.

<sup>30</sup> See Ney (2007) for a detailed argument that an appeal to the constitution relation does not solve the exclusion problem.

<sup>31</sup> Constitution strategists include Pereboom (2002) and Paul (2007).

<sup>32</sup> For a longer discussion of additivity, see Paul (2007).

<sup>33</sup> Of a similar thought experiment based on a meddling, Malebranchian god, Bennett (2003) also argues that the meddling god case is unlike the mental/ physical case.

<sup>34</sup> (Modal Distinctness OD) isn’t exactly equivalent to the conjunction of (O1) and (O2), for the former is expressed in terms of causal sufficiency while the latter are formulated in terms of counterfactuals, and there can be counterfactual dependence without sufficiency.

<sup>35</sup> Assume that the twins are modally separable, which is to say, one can occur without the other.

<sup>36</sup> Note that one need not be able to make sense of impossible worlds per se to agree with the idea that those are the informative worlds with respect to the contribution of the mental.

<sup>37</sup> See Brogaard and Salerno (2013) for another argument defending the non-vacuity of counterpossibles.

<sup>38</sup> See Nolan (1997) for an argument that impossible worlds can have a similarity ordering like that of possible worlds.

**Acknowledgements.** I owe many people thanks for feedback on this paper at its various stages. I am particularly grateful to Terence Horgan, Lewis Powell, L.A. Paul, Alex Rosenberg, and Daniel Silvermint for extensive feedback on this paper. I am also grateful to audiences at Australian National University, University of Missouri-St. Louis, University of Western Ontario, Duke University, Southern Methodist University, and the Eastern division meeting of the American Philosophical Association for helpful comments. Finally, I thank several anonymous referees.

## References

- Bernstein, S. "Omission Impossible", *Philosophical Studies*, forthcoming.
- Bennett, K. (2003) "Why the Exclusion Problem is Intractable, and How, Just Maybe, to Tract It", *Nous*, 37:3 (471-497)
- Bennett, K. (2008) "Exclusion Again" in Jesper Kallestrup & Jakob Hohwy (eds.), *Being Reduced: New Essays on Causation and Explanation in the Special Sciences*. Oxford University Press.
- Brogaard, B. and Salerno, J. (2013). "Remarks on Counterpossibles", *Synthese* 190 (4):639-660.
- Bunzl, M. (1979) "Causal Overdetermination", *The Journal of Philosophy* 76, 134-150.
- Davidson, D. (1980) "The Individuation of Events" in *Essays on Actions and Events*. New York: Oxford University Press.
- , 1970, "Mental Events", in Davidson 1980.
- Funkhouser, E. (2002) "Three Varieties of Causal Overdetermination" *Pacific Philosophical Quarterly*.
- Hoffman, Frank. (2007) "Causal Powers, Realization, Mental Causation" *Erkenntnis* September 2007, Volume 67, Issue 2, pp 173-182
- Kim, J. (1976) "Events as Property Exemplifications" Reprinted (1993) in *Supervenience and Mind*. Cambridge: Cambridge University Press, 33-52.
- Kim, J. (1993) *Supervenience and Mind*. Cambridge: Cambridge University Press. pp. 33-35.
- Kim, J. (2005) *Physicalism, or Something Near Enough*. Princeton: Princeton University Press.
- Loewer, B. (2007) "Mental Causation or Something Near Enough". In *Contemporary Debates in Philosophy of Mind*, Brian McLaughlin and Jonathan Cohen (ed.), Malden, MA: Wiley-Blackwell, pp. 243-264.

- Merricks, T. (2001) *Objects and Persons*. Oxford: Oxford University Press.
- Ney, A. (2007) "Can an Appeal to Constitution Solve the Exclusion Problem?" *Pacific Philosophical Quarterly*, Vol. 88.
- Nolan, D. (1997) "Impossible Worlds: A Modest Approach" *Notre Dame J. Formal Logic* Volume 38, Number 4, 535-572.
- Paul, L. (2007) "Constitutive Overdetermination" in *Topics in Contemporary Philosophy*, Vol IV: Causation and Explanation.
- Pereboom, D. (2002) "Robust Nonreductive Materialism," *Journal of Philosophy* 99, pp. 499-531.
- Shoemaker, Sydney. 2007. *Physical Realization*. Oxford: Oxford University Press.
- 2003. Realization, Micro-realization, and Coincidence. *Philosophy and Phenomenological Research*. Vol. 67, 1-23.
- Schaffer, J. (2003) "Overdetermining Causes", *Philosophical Studies* 114: 23-45.
- Sider, T. (2003) "What's so Bad about Overdetermination?" *Philosophy and Phenomenological Research* 67: 719-726.
- Stoljar, D. (2007) "Distinctions in Distinction" in Jesper Kallestrup & Jakob Hohwy (eds.), *Being Reduced: New Essays on Causation and Explanation in the Special Sciences*. Oxford University Press.
- Thomasson, A. (2007) *Ordinary Objects*. Oxford: Oxford University Press.
- Willard, M. (2014) "Against Simplicity," *Philosophical Studies*. 167: 165-181.
- Wilson, J. (2010) "Nonreductive Physicalism and Degrees of Freedom" *Br J Philos Sci* 61 (2): 279-311.
- Yablo, S. (1992) "Mental Causation," *The Philosophical Review* 101, 245-280.