

COMPATIBILISM, INTRALEVELISM, AND BI-DIRECTIONAL DETERMINATION

by

JAY GEYER

B.A., University of Illinois, 2010

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Master of Arts
Department of Philosophy

2012

Signature Page

This thesis entitled:
'Compatibilism, Intralevelism, and Bi-Directional Determination'
Written by Jay R. Geyer
Has been approved for the Department of Philosophy

Michael Tooley
(committee chair)

Date

Robert Rupert
(committee member)

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Geyer, Jay R. (M.A., Philosophy)

Compatibilism, Intralevelism, and Bi-Directional Determination

Thesis Directed by Professor Michael Tooley

A leading view on mental ontology over the past three decades has been a non-reductive physicalist view (NRP), which affirms that the mental is irreducible, supervenient upon its physical correlate and causally efficacious, and also affirms a weakened articulation of the causal closure of physics. Jaegwon Kim's well known causal exclusion argument has challenged the plausibility of NRP on the grounds that the causal sufficiency of the mental's supervenience base either excludes the mental from causal efficacy or entails that any effect of the mental will be causally overdetermined. In this thesis I examine two response strategies to Kim's exclusion argument that deny the entailment of widespread overdetermination. In chapters two and three, I argue that both strategies fail and that NRP does indeed entail widespread overdetermination, making it an untenable theory. In the final section I motivate a novel mental ontology that avoids the problems that beset NRP.

ACKNOWLEDGMENTS

I am grateful to Michael Tooley, Rob Rupert, David Barnett, and Noel Saenz for all of their helpful comments and suggestions.

CONTENTS

CHAPTER

I. INTRODUCTION.....	1
Compatibilism.....	2
Intralevelism.....	4
Bi-Directional Determination Theory.....	6
Non-Reductive Physicalism and the Causal Exclusion Argument.....	7
II. BENNETT’S COMPATIBILIST SOLUTION.....	9
Removing a Horn from Bennett’s Dilemma.....	13
Overdetermination and Redundancy.....	16
Conclusion.....	20
III. INTRALEVELIST VIEWS.....	21
Hybrid Overdetermination and Exclusion.....	26
The Generalization Problem and the Drainage Problem.....	28
Kim’s Response.....	29
Mental-to-Physical Causation.....	32
Conclusion.....	36
IV. BI-DIRECTIONAL DETERMINATION.....	38
Concerning the Lack of Better Alternatives.....	39
The Central Challenge of BDT.....	40
BDT1.....	41
BDT2.....	45
BDT3.....	47

Another Virtue of BDT.....	49
Conclusion.....	50
BIBLIOGRAPHY.....	52

FIGURES

Figure

1. Simple Causal Chain.....	9
2. Non-Simultaneous Overdetermination.....	10
3. Schlosser's Intralevelism.....	22
4. Classical NRP and The Layered View.....	24
5. The Layered View.....	27
6. BDT1.....	43
7. Co-variation, Dependence.....	44
8. Co-variation, Common Cause.....	45
9. BDT2.....	46
10. BDT3.....	47

CHAPTER 1

INTRODUCTION

Jaegwon Kim's causal exclusion argument has long been a thorn in the side of Non-Reductive Physicalism (NRP) about the mind (1998, 2005). Kim argues that for the mental to be meaningfully causally efficacious in NRP it must be able to causally interact with the physical. However, given the physicalist commitment to the causal closure of physics, every physical effect already has a sufficient physical cause. In NRP the mental cause cannot be reduced to that physical cause and so it would seem that the mental is either excluded from causal efficacy by its physical supervenience base that is sufficient for any putative mental effect, or else every effect of the mental is causally overdetermined. Widespread causal overdetermination is generally taken to be highly problematic, while a complete epiphenomenalism concerning the mind is an unacceptable commitment for any mental ontology.

In this thesis I shall examine two response strategies to Kim's exclusion argument. Both strategies deny that NRP entails widespread overdetermination. In sections two and three, I shall argue that both strategies fail and that NRP does indeed entail widespread overdetermination. I shall assume that this entailment makes NRP an untenable theory, and I will have nothing to say here in response to some philosophers' claim that there is nothing problematic about widespread overdetermination.¹

If NRP is shown false by the exclusion argument, then the question of what the mental is and how it causally interacts is left open. In the final section of the thesis I shall argue that extant mental ontologies do not provide a satisfactory answer to this question. In light of their failings, I shall tentatively offer an alternative account of the mental that I call 'Bi-Directional Determination Theory', or 'BDT'. Despite some clever arguments to the contrary, NRP entails

¹ See Ted Sider (2003) for a defense of widespread causal overdetermination as an acceptable commitment.

widespread overdetermination and thus ought to be rejected; in its place we ought to endorse BDT.

Compatibilism (Introduction)

Several philosophers have challenged whether NRP entails widespread overdetermination, and in this thesis I shall examine two strategies for denying that NRP entails widespread overdetermination. The first is offered by Karen Bennett (2003)² and takes as its starting point a certain understanding of ‘overdetermination’. A necessary condition for an effect to be overdetermined is that there are at least two sufficient causes for that effect. This condition is accepted by all parties, but it is too permissive to be an analysis of ‘overdetermination’. At the very least, another necessary condition is needed. Bennett argues that this extra condition should be understood as a pair of counterfactuals where if one sufficient cause did not obtain the other would still cause the effect and vice versa.³

Armed with this necessary condition for overdetermination, Bennett argues that the mental and physical in NRP cannot be said to overdetermine any effect due to the mental’s supervenience on the physical.⁴ Because of this supervenience relation, Bennett’s counterfactuals either come out false or vacuously true in NRP, depending on how localized or comprehensive we take the physical supervenience base of the mental to be. Thus, despite having two simultaneous sufficient causes, the effects of the mental in NRP are not overdetermined as the counterfactual necessary conditions for overdetermination are not met.

² Other philosophers who endorse similar strategies include Steven Yablo (1992) and Terence Horgan (1997).

³ I have put this in terms of the counterfactuals being necessary conditions for overdetermination, but we could just as easily say they are necessary conditions for *problematic* overdetermination. In that case, Bennett would argue that although the effects of the mental in NRP are overdetermined they are not problematically overdetermined (2003, p. 474).

⁴ Unless otherwise specified, when I use ‘supervenience’ I mean supervenience and ontological dependence (what Kim calls ‘strong supervenience’ (2005, p. 33)). Weak supervenience would just be a co-variation relation.

The view that an effect can have two sufficient causes without being overdetermined can be called ‘compatibilism’,⁵ and in this thesis I shall refer both to this account of overdetermination and to the response to the exclusion argument that is based on this account as ‘compatibilism’.

I shall offer a two-part response to Bennett. First, I shall argue that on a localized understanding of the physical supervenience base, Bennett’s counterfactuals actually come out true and there is widespread overdetermination in the NRP-ist picture. Bennett does not offer a reason to prefer the comprehensive understanding of the supervenience base over the localized understanding, instead presenting her argument as a dilemma, where on either understanding the mental does not overdetermine its effects. By denying Bennett’s localized/falsity strategy, her dilemma argument loses a horn, and so fails. The second part of my response is to argue that even if Bennett can give reasons to prefer the comprehensive understanding of the supervenience base, on which the counterfactuals come out vacuously true, she has discharged none of the initial worries about overdetermination in NRP as these worries do not track her counterfactual necessary conditions. Bennett’s compatibilist project admits that there are two distinct sufficient causes for any effect of the mental, but denies that these are cases of overdetermination (2003, p. 473). If I am right, then the mere existence of two distinct sufficient causes for some effect, what I shall call ‘redundant causal sufficiency’ is enough to generate all of the worries associated with overdetermination. In that case, Bennett’s compatibilist project has not really helped NRP against the exclusion argument, even if we ought to understand ‘overdetermination’ differently than we used to.

Intralevelism (Introduction)

⁵ The term in this context was coined by Terrence Horgan (1997).

The second sort of attempt to deny that NRP entails widespread overdetermination is the sort offered by Amie Thomasson (1997), John Gibbons (2006), and Markus Schlosser (2009). While each offers a different story about how the mental could be causally efficacious without violating closure or overdetermining its effects, all deny:

Causing Supervenient Properties (CSP): To cause a supervenient property to be possessed, the property's base property must be caused to be possessed (Kim, 1998, p.42).⁶

This shared feature makes all of these theories what Andrei Buckareff calls 'intralevelist theories' (2011, p. 403). In intralevelist theories, the mental only causally interacts with relata at the mental level and never with physical relata. This sort of picture neatly avoids the charge of overdetermination as the mental never causes anything physical and thus does not compete with any physical cause.

I will focus primarily on Thomasson's intralevelist account, what she calls the 'Layered View', but my critiques of her view will be general enough to apply to any intralevelist theory. There are two general worries for any intralevelist account: First there is the issue of whether CSP can be denied; whether supervenient property instantiations can be caused without causing their supervenience base. If supervenient properties are excluded from being direct effects, then intralevelist views are a non-starter, at least as physicalist accounts of the mind where the mental supervenes on the physical.

I shall argue that there are good reasons to doubt CSP as it appears to have two implausible entailments: First, if supervenient properties are excluded from being direct effects, then, coupled with overdetermination worries, they also seem to be excluded from being direct

⁶ This phrasing is from Buckareff (2010, p. 404-405)

causes.⁷ As there seems to be no principled way of limiting such exclusion to mental properties, Kim is faced with the specter of epiphenomenalism about all supervenient properties. Second, if supervenient properties are excluded from causal efficacy, then only property instantiations at the bottom-most level of supervenience are causally efficacious. However, as Ned Block points out (2003), it is an open question of physics whether there is a bottom-most level of supervenience. This makes the existence of causation itself, on Kim's view, contingent on an open question of physics. I shall argue that it is far from clear that Kim has a satisfactory response to either of these concerns, what I shall call the 'generalization problem' and what is known as the 'drainage problem'. As CSP is highly questionable, intralevelist views are off the hook with regard to the first issue; supervenient properties need not be epiphenomenal, and mental properties can cause other mental properties without causing the dependence base of those mental effects.

The second major obstacle for intralevelist views is the issue of physical-to-mental and mental-to-physical causation. Direct causal relations between levels are explicitly ruled out by intralevelist accounts. At a glance, this should strike us as totally unacceptable. That tissue damage causes pain, or that intentions cause actions must be accounted for in any acceptable mental ontology. The success of the intralevelist project will depend on how well they can cash out our intuitions about physical/mental interaction without appealing to causation. While supervenience seems to offer intralevelists an upward explanatory mechanism by which to cash out physical-to-mental causation, they cannot perform this operation in reverse as supervenience is asymmetric. A mental ontology that cannot offer any story about mental-to-physical causation

⁷ If only the supervenience base can be a direct effect, and if, plausibly, the supervenience base has a sufficient cause at its level, then all supervenient causes could only overdetermine the supervenience base. On pain of extremely widespread overdetermination being a feature of the universe, such putative supervenient causes should plausibly be eliminated, excluded by their supervenience bases.

is an untenable mental ontology and I shall show that intralevelism should be rejected on these grounds.

Bi-Directional Determination Theory (Introduction)

In the third section of this thesis, I shall tentatively offer a novel mental ontology. BDT has the following virtues: 1. It does not entail causal overdetermination; 2. It does not reduce the mental to the physical; 3. It offers a way to understand mental-to-physical causation; and 4. It allows for a robust form of the causal closure of the physical principle. 1 and 4 give BDT a theoretical advantage over classical NRP (the sort of theory Bennett's compatibilism attempts to preserve), 2 gives BDT an advantage over reductive physicalism and compromise views like Kim's,⁸ and 3 gives BDT an advantage over intralevelist views.

BDT achieves these advantages by positing bi-directional determination relations between the mental and the physical.⁹ Causal processes happening at the mental and physical level never directly causally interact with each other. In other words, there are no interlevel causal interactions. In this respect, BDT is an intralevelist view. Unlike other intralevelist views however, BDT allows us to cash out mental-to-physical causation in the form of the mental's determination of the physical. The mental's ability to determine the physical means that the mental is not ontologically dependent on the physical, and so BDT is not a physicalist theory. While BDT faces some objections, its unique theoretical advantages over its competitors at least earn it a place at the table of mental ontologies.

⁸ Kim's view is that qualitative mental states cannot be reduced, but because they cannot they are epiphenomenal. Behavioral mental states can be reduced and thus get to be causally efficacious (2005, p. 162)

⁹ There will be much more detail in section four concerning just what this determination relation could be.

Non-Reductive Physicalism and the Causal Exclusion Argument

Considering what it promises to deliver, it is not surprising that NRP has established a prominent place among extant mental ontologies. By positing an irreducible mental, NRP sidesteps the worries that beset reductive physicalism. At the same time, NRP counts itself among physicalist mental ontologies by endorsing a type of causal closure of the physical and insisting that mental entities, though irreducible, still supervene on physical entities. To pull this off, NRP is committed to the following four claims.¹⁰ The first two make NRP at least minimally a physicalist theory:

Weak Closure: All physical effects have a sufficient physical cause.¹¹

Supervenience: Mental properties supervene (and ontologically depend on)
physical/biological properties

In order to accommodate the worries that beset reductive physicalism,¹² the mental is irreducible in NRP:

Irreducibility: Some mental properties are not reducible to physical properties.

Finally, NRP is committed to:

Efficacy: Irreducible mental properties can be both causes and effects distinct from their physical supervenience base.

¹⁰ I take these four commitments from Kim (2005). One might call a view with less commitments than these (say only supervenience and irreducibility) a non-reductive physicalist theory, though it is not clear that a theory could rightly be called physicalist without endorsing some form of causal closure, while a theory in which nothing mental is causally efficacious is probably not worth considering too seriously (even someone like Paul Churchland is unwilling to eliminate all mental states, choosing instead to reduce some, thereby rendering them causally efficacious (2007)). Of course how to understand the four commitments will be subject to nuanced disagreements about reduction and causation, but I take it that the exclusion argument will be generated by an endorsement of any reasonable interpretation of all four commitments.

¹¹ A more robust form of causal closure might stipulate that physical entities can only causally interact with other physical entities. This will be relevant in the fourth section of the thesis. Also, it should be noted that Weak Closure is at odds with the view, supported by evidence from quantum physics, that the universe is indeterministic. The principle could be reformulated to meet such concerns as ‘all physical effects have their objective probability determined by only physical entities.’

¹² These worries include the well-known explanatory gap concerns (Levine, 2001) and multiple realizability concerns (Putnam, 1967).

That some mental states are causally efficacious is a minimum requirement of any acceptable mental ontology, without which we would seem to surrender learning, agency, and such intuitively causal relations as tissue damage causing pain. The sort of view that I'm calling 'NRP' and that is targeted by the exclusion argument is that *irreducible* mental states are causally efficacious. Kim holds a hybrid reductivist view, where behavioral mental states like beliefs and intentions are reducible and causally efficacious, but qualitative states are irreducible and thus epiphenomenal (2005). Such a view is not what I'm calling 'NRP' and would not be vulnerable to the exclusion argument.

What Kim argues with the exclusion argument is that the four commitments of NRP are in conflict. If the mental is to cause anything it must cause something physical. Even if the mental is to cause a mental effect, that mental effect supervenes and ontologically depends on a physical realizer. Here, CSP comes into play with the claim that to cause a supervenient property instantiation, the base realizer must be caused. So, any effect of the mental is a physical effect. However, all physical effects already have sufficient physical causes given Weak Closure. So either the mental is excluded from causal efficacy by its supervenience base or else any mental effect is overdetermined. This widespread overdetermination is highly problematic, and so, argues Kim, NRP should be rejected (1998, 2005).

CHAPTER 2

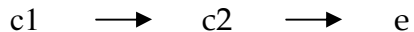
BENNETT'S COMPATIBILIST SOLUTION

The first response strategy to the exclusion argument can be called 'causal compatibilism' or simply 'compatibilism'. Compatibilists argue that some effects can have more than one sufficient cause without being causally overdetermined and argue that the effects of the mental in NRP are examples of such cases. Compatibilism's success will depend on how we ought to understand overdetermination. Here is a necessary condition of overdetermination that all sides, compatibilists included, accept:

1. Effect **e** is overdetermined only if there are two sufficient causes for **e**

While certainly a necessary condition for overdetermination, 1 falls far short of a complete analysis. A simple causal chain serves as a counter-example:

Figure 1: Simple Causal Chain

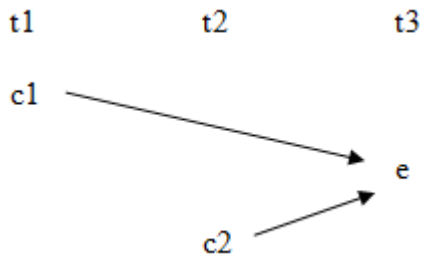


Here, **c1** is causally sufficient for **c2** which is in turn causally sufficient for **e**, but obviously **c1** and **c2** do not overdetermine **e**. Something must be added to 1. We might be inclined to think that a simultaneity condition will do the trick:

2. Effect **e** is overdetermined iff there are two causes for **e** at time **t** that are both sufficient for **e** at **t**.

2 does get the simple causal chain right, but stumbles in the following sort of case:

Figure 2: Non-Simultaneous Overdetermination



In the above figure, *c1* and *c2* are both directly causally sufficient for *e*, but are not simultaneous.

Intuitively, this is still a case of causal overdetermination, but 2 gets it wrong.

Presumably there are other conditions we could add to 1 to get a better analysis of overdetermination. A directness condition might be a good start:

3. Effect *e* is overdetermined iff there are two causes for *e* that are both sufficient for *e* and have no causal intermediaries between them and *e*.

Of course, 3 discounts almost all intuitive cases of overdetermination (the overdetermining causes of the firing squad victim's death turn out not to be the two rifles firing, but the two bullets penetrating his vital organs, or whatever the most proximate causes are to the victim's actual death). The following might do a bit better:

4. Effect *e* is overdetermined iff there are two causes for *e* that are both sufficient for *e* and each cause is part of a distinct, complete causal chain which intersects with the other causal chain at *e*.

We would need to add some extra conditions to 4 concerning just where in the causal chain a cause must reside to compete with a cause in another causal chain. Two baseballs striking the window overdetermine the shattering, but presumably one baseball striking the window does not overdetermine the shattering with some much more distant cause in the chain involving the second baseball striking the window: The boy who threw the second baseball being born is not

an overdetermining cause, for example. I am fairly confident that there is some set of conditions along these lines that can be added to 1 that would get us a satisfactory account of overdetermination and this would be an interesting project, but it is not my project. Let us just say that some sort of strategy along these lines works and call it 1+.

Regardless of the success of the 1+ project, Bennett already has a compatibilist necessary condition to add to 1 that would make sense of most problem cases. In addition to 1, Bennett argues that the following two counterfactuals must be satisfied for a case to count as a case of overdetermination:

O1: If **c1** (or mental event **m**) had happened without **c2** (or physical event **p**), **e** would still have happened

O2: If **c2** (or physical event **p**) had happened without **c1** (or mental event **m**), **e** would still have happened (Bennett, 2003, p. 480)

Besides tracking paradigmatic overdetermination cases like windows being shattered by simultaneous baseballs, Bennett argues that O1 and O2 correctly diagnose trickier cases like the ones in the figures above (2003, p. 477).¹³

If we accept O1 and O2 as necessary conditions of overdetermination, then Bennett's compatibilist project may have a chance. At least it is not ruled out from the outset as it might be in 1+. Bennett points to two general strategies for showing that effects of the mental are not overdetermined despite having both a sufficient physical and a sufficient mental cause. First, we could say that in these cases O1 or O2 is false, that if one cause occurred without the other, the effect would not obtain. Second, we could show that in mental causation cases either O1 or O2

¹³ However, like 4, O1 and O2 fail to screen off cases like the birth of the boy who would throw the second baseball overdetermining the window shattering with the first baseball striking the window. This may be fine for Bennett, as she only claims O1 and O2 to be necessary conditions of overdetermination. Both 4 and O1 and O2 require some other conditions to screen off cases of causes temporally distant from each other overdetermining an effect.

are true, but only vacuously true; that is, we could show the antecedent of O1 or O2 to be necessarily false by showing that either the physical or the mental cause could not occur without the other.

Bennett advances each of these strategies depending on whether we understand the physical realizer of the mental as localized or comprehensive (2003). The localized realizer is what Sydney Shoemaker calls the ‘core realizer’ and the comprehensive realizer is what Shoemaker calls the ‘total realizer’ (1981):

A total realizer of a property will be a property whose instantiation is sufficient for the instantiation of that property. A core realizer will be a property whose instantiation is a salient part of a total instantiation of it (Shoemaker, 2007, p. 21).

A paradigmatic core realizer would be c-fiber firing as the realizer for *being in pain*. c-fiber firing (or whatever the actual physical state of affairs is) is the salient aspect of the total realizer, but it by itself is not sufficient for grounding the mental state. We (or some future scientists) could recreate c-fiber firing in a petri dish, but this would not ground any mental state, let alone the mental state *being in pain*. For this mental state to obtain, not only must c-fibers be firing, but a whole host of other background conditions must obtain. Whatever these background conditions are, the set of them sufficient for the mental state (plus the core realizer) is the total realizer.

By definition then, total realizers cannot obtain without necessitating the mental state that they ground. It is on such an understanding of physical realizer that Bennett advances the vacuity claim. On the total realizer understanding, **p** could not obtain without **m** and so the antecedent of O2 is false and the conditional is vacuously true. Thus, if we understand the physical realizer as the total realizer, the conditions for compatibilist overdetermination are not

met and the effects of the mental are not overdetermined, despite having two simultaneous sufficient causes.

If, on the other hand, we understand the physical realizer as the core realizer, then **p** can certainly occur without **m** by being part of a different total realizer. Shoemaker intuitively refers to those aspects of the total realizer besides the core realizer as the ‘surround’ (2007). Paired with a different surround, the core realizer c-fiber firing could still obtain, but the total realizer it would be a part of might ground a wildly different mental state, such as *desiring to fly* (Shoemaker, 1981, p. 97). Bennett argues that on the core realizer understanding, O2 is false. That is, **p** cannot ground a different **m** without also causing a different **e**. If the surround is different, argues Bennett, then we have little reason to think that the effect would be the same as the one caused by the original **p** plus surround (2003, p. 488).

Bennett presents her argument as a dilemma for NRP’s accusers: Either we understand the physical realizer of the mental as the total realizer or the core realizer. If we understand it as the total realizer, then O2 is vacuously true. If we understand it as merely the core realizer, then O2 comes out false. Either way NRP does not entail overdetermination as both O1 and O2 must obtain for an effect to be overdetermined.

Removing a Horn from Bennett’s Dilemma

I shall offer two objections against Bennett: First I shall attempt to show that if we understand the physical supervenience base as the core realizer, then in many cases it is true that if **p** happened without **m** (or with a different **m**), **e** would still happen. If I am right, then one horn of Bennett’s dilemma argument is broken and the burden is on her to argue for the total realizer understanding of the supervenience base on independent grounds. Second, I shall argue

that even if she manages to do this, there is a more fundamental flaw with her project. The worries we might have about widespread overdetermination, epistemic or metaphysical worries, remain problems for NRP even if Bennett is successful. This is because these worries are grounded in there being two sufficient causes for some effect where one will do, an entailment of NRP by Bennett's own admission.

Let us begin by showing that under a core realizer understanding of the physical supervenience base, if **p** obtained without **m**, **e** would still have happened. For Bennett's falsity strategy to work, it must be the case that, at least in most cases of mental causation, **p** could not obtain without **m** without also causing a different **e**. I shall argue that the falsity strategy fails in light of the following counterexample: Take as our **p**, the core realizer, c-fiber firing. The mental state necessitated by the c-fiber firing plus surround is the mental state *experiencing intense pain*. The input stimulus is touching a hot stove and the output is avoidance behavior, say, jerking one's hand back. Now let us imagine that our subject has taken an anesthetic that dulls the pain by manipulating the surround while leaving the core realizer intact. When our subject, under the influence of the anesthetic, touches the hot stove his c-fibers fire, but instead of entering the mental state *experiencing intense pain*, he enters the mental state *experiencing moderate pain*. *Experiencing moderate pain* still supervenes on the core realizer, c-fibers firing, and both cause the output avoidance behavior of the subject jerking his hand back from the stove. This would seem to be a case of the very same **p** causing the very same **e**, while realizing a different **m**.

This sort of case works because, plausibly, a slight difference in the level of pain experienced will produce no discernible difference in avoidance behavior. The subject jerks back his hand just as vigorously. Generally speaking, **e**'s are pretty coarse grained events, at

least when they are actions like avoidance behavior. Mental states and their physical realizers, on the other hand, are typically characterized in a very fine grained way. We can easily specify small differences for **m**: The original **m** might be *experiencing excruciating pain*, because of an appendix on the verge of rupture, while the modified **m** may be *experiencing slightly less excruciating pain*, because the subject has taken a single aspirin. **es** are just not responsive to these subtle changes in mental states (and corresponding changes to the total realizers of those mental states). This is a general feature of **es**, at least when they are not mental states themselves, and so counterexamples to Bennett's falsity claim are ubiquitous. Any mental state whose effects are not other mental states can be manipulated by changes to the surround of the total realizer that grounds it, and these manipulations, so long as they are subtle, will not carry over to the effects of such mental states. So, in many cases, O2 comes out true on the core realizer understanding of **p** and the mental in NRP overdetermines its effects even on Bennett's compatibilist analysis: If **p** had happened without **m**, **e** would still have happened.

Now, it may be that on the core realizer understanding O2 does come out false for cases where the effect of the mental is another mental state as the effect mental state can be calibrated just as finely as the altered cause mental state. Even so, at least every *action* would be overdetermined in NRP.¹⁴ This would be widespread enough overdetermination to cast serious doubt on NRP. Every bodily movement I make that has a causal origin in a mental state would be overdetermined.¹⁵

¹⁴ This may give us reason to favor a weakened version of the exclusion argument, where only mental-to-physical causation generates either overdetermination or exclusion. An additional reason for favoring a weakened exclusion argument will become clear in section 3; a weakened exclusion argument would still run even if CSP turns out to be false.

¹⁵ There are other effects of the mental besides bodily movements that could also be overdetermined in this way due to their coarse-grained nature. For example, excruciating pain may cause me to lose consciousness, but so will slightly less excruciating pain. Losing consciousness does not admit of subtle levels of degree like experiencing pain does, so once again the same local **p** realizes a different **m**, but causes the same **e**.

So it seems that Bennett's falsity strategy fails and we are left with only the vacuity strategy. In that case Bennett no longer has a dilemma and there is nothing stopping us from understanding the supervenience base of the mental as the core realizer, in which case the NRP account of mental causation entails widespread overdetermination even on Bennett's compatibilist account. Bennett presents no argument for favoring the total realizer understanding over the core realizer understanding, but even if such an argument could be mustered, I shall argue in the next section that her project falls victim to a more general worry; that the problems associated with overdetermination are problems entailed by having redundant sufficient causes for an effect, an entailment of NRP even by the compatibilists' lights.

Overdetermination and Redundancy

If the NRP account of mental causation does not entail widespread causal overdetermination, it entails something close enough. We can see this by reminding ourselves why widespread causal overdetermination is taken by most to be problematic. There are several reasons to worry about widespread overdetermination, but here, for the sake of brevity, I will focus on only two¹⁶:

1. Metaphysical worries: Causation is such that only one cause can be sufficient for an effect at a time (at least in worlds like ours).

¹⁶ Ted Sider considers both of these reasons and a third, that the existence of widespread overdetermination would require some explanation about why these extra sufficient causes obtain, in his paper 'What's So Bad about Overdetermination' (2003). Sider's answer to his titular question is that there is nothing wrong with widespread causal overdetermination. Whether he is right lies outside of the scope of this thesis, although some of what I have to say in this section would suggest otherwise.

A fourth worry about widespread overdetermination is that it is highly improbable. The *a priori* probability of there being a kind of event that has two distinct other kinds of events linked by causal laws such that they *both* always cause the first kind of event seems to be rather low. In the NRP case, there seem to be a large number of kinds of events that are overdetermined and so the *a priori* probability of such widespread overdetermination seems to be even lower. Thanks to Michael Tooley for pointing out this worry.

2. Epistemic worries: Widespread overdetermination is unparsimonious and additional sufficient causes should be eliminated.

Whether these worries are insurmountable is not our present concern. Here we are assuming with the compatibilist that widespread overdetermination is a real problem and that the NRP account of mental causation is untenable if it entails widespread overdetermination. The question at hand is whether Bennett's project has allayed these underlying worries at all. If it has not, then even if O2 is only vacuously true and the effects of the mental in NRP are not overdetermined, these cases are still problematic for the reasons we thought overdetermination was problematic.

Let us take the two worries above in order to see if they are allayed at all in light of Bennett's compatibilist project, beginning with the metaphysical worry. The claim is that if an effect has one sufficient cause, there is nothing else for any other cause to do. Sider gives a slightly caricatured depiction of this sort of view:

Causation is a kind of fluid divided among the causes of an effect. If one potential cause acts to produce an effect, the fluid is *used up*, and no other potential cause can act (Sider, 2003, p. 3).

Sider claims that this is 'a view of causation that no one accepts' (2003, p. 3) but that this is the only sort of view on which overdetermination is implausible on metaphysical grounds. As a matter of fact some philosophers do endorse a view of causation not so different than this, including Kim himself (2007). Kim endorses some sort of productive or generative account of causation first suggested by Elizabeth Anscombe (1975) with later, more detailed, variants developed by philosophers like David Fair (1979) and Wesley Salmon (1980) where causal processes are necessarily physical processes. According to Kim:

Agency requires the productive/generative conception of causation. . . These causal processes all involve *real connectedness* between cause and effect, and the connection is constituted by phenomena such as energy flow and momentum transfer, an actual movement of some (conserved) physical quantity (2007, p. 236).

On the productive account of causation, it is far from clear that overdetermination ever happens, at least in worlds with laws like ours. Whatever the cause is that is sufficient for some effect, that cause is necessarily physical, but physical causation is something of a zero sum game, involving conservation of energy, momentum etc. If the victim in front of the firing squad has two bullets penetrate his heart simultaneously, his death is not overdetermined on the reductivist physicalist picture. Rather we would need to look at the exact details of the event: Did one bullet penetrate a more critical part of the heart, leading to his death before the other bullet could do its work? Did both bullets penetrate the exact same part of the heart at the exact same time? In the second case, both bullets contributed to the victim's death; both played some role in the physical events leading to the victim's heart stopping. There is no overdetermination in such reductivist physicalist accounts of causation, only, at most, cooperative causation. There is some physical fact of the matter describable in terms of energy and momentum transfer about which bullet was causally responsible for the death and, in the event that they cooperated to bring about the effect, to which extent each bullet was causally responsible.

I am not at all concerned here with the merits of a reductivist physicalist view of causation. The point is that the metaphysical worry is a real worry for Bennett's chief opponent. If Kim thought that overdetermination was just not possible in worlds with laws like ours, would he be swayed by Bennett's conclusion that the NRP account of mental causation does not involve overdetermination? I think it is clear that he would not be swayed. Bennett's solution

acknowledges that any effect of the mental has both a sufficient mental cause and a sufficient physical cause, only she denies that these are cases of overdetermination. ‘But that is irrelevant,’ Kim might say, ‘the problem is just that there are *two sufficient causes* for the same effect.’ If ‘overdetermination’ is not the right term for this sort of occurrence, then let us call it ‘redundant causal sufficiency’ instead. Redundant causal sufficiency is what generates the metaphysical coherence worry and the entailment of redundant causal sufficiency by the NRP account of mental causation is admitted by Bennett.

Is it redundant causal sufficiency or compatibilist overdetermination that generates the second worry about parsimony? The worry is that if all putative effects of the mental already have a sufficient physical cause (entailed by causal closure and supervenience) then any additional mental cause for some mental or brain event seems unnecessary and eliminable. Again, whether mental causes are eliminable is not our present concern. Our concern is whether it is compatibilist overdetermination or redundant causal sufficiency that generates these elimination worries. Once again, it seems clear that it is redundant causal sufficiency. If we think that mental causes are eliminable it is because they are redundant and the compatibilist solution admits that there is a sufficient physical cause for every effect of the mental. Two sufficient causes where one will do is unparsimonious, regardless of whether this counts as a case of overdetermination or not.

The compatibilist might push back here and point out that I am stacking the deck a bit by calling the mental cause ‘redundant’, or saying that there are two sufficient causes ‘where one will do’. Maybe the compatibilist could deny this, instead arguing that unlike firing squads and baseballs, it is not the case that in NRP the mental cause is redundant. It is not clear what this response amounts though, without saying either that the mental cause is not sufficient for its

effects, or that the mental is not distinct from the physical. Surely the NRP-ist picture is not one in which the mental and physical instantiate different links in a chain of sufficient causes. If the mental cause is sufficient for its effect and cannot be reduced to the physical cause of that effect, then it sure sounds like there are two distinct causes sufficient for a single effect. It is pretty much analytic that one sufficient cause *will do*, and being a physicalist theory NRP must give primacy to the physical cause, so the mental cause is redundant.¹⁷

Conclusion

Bennett's compatibilist re-definition of 'overdetermination' fails to save NRP from the exclusion argument. First, because even on her definition of 'overdetermination' it is not clear that widespread overdetermination is not entailed by NRP. If the physical supervenience base is understood as the core realizer, then many effects of the mental (at least all actions) are overdetermined even on Bennett's definition. Even if the supervenience base should be understood as the total realizer, and NRP does not entail compatibilist overdetermination, it entails all of the same problems that were grouped under the label 'overdetermination worries' as NRP entails redundant causal sufficiency.

¹⁷ Steven Yablo has argued that while the mental and physical can both be *causally relevant* to (and even *causally sufficient* for) the same effect, we can preference one over the other as the *cause* and, he argues, sometimes the mental cause is to be preferred over the physical cause as the mental cause is more 'commensurate' with the effect (1992, p. 274). If Yablo is right, then the NRP-ist need not give primacy to the physical cause, but nevertheless it seems that *some* cause is redundant if both are sufficient. It is difficult for me to understand Yablo's contention that something can be causally sufficient for an effect without causing that effect. Causation for him seems to be a context relative relation (being commensurate with an effect the hallmark of the true cause), but this weak sort of relation strikes me as an inadequate notion of causation.

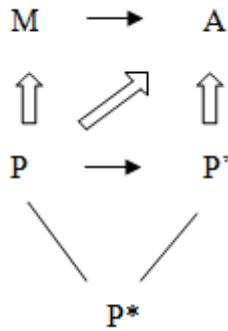
CHAPTER 3

INTRALEVELIST VIEWS

The second response strategy denying that NRP entails widespread overdetermination (on pain of exclusion) is to deny CSP, the principle stating that supervenient properties cannot be caused except by causing their supervenience base. It is this principle that ensures that not even mental-to-mental causation can occur without downward causation to the physical. As any physical effect of the mental already has a sufficient physical cause, *all* effects of the mental (being physical effects) would be overdetermined if CSP is true. Intralevelist views, by denying CSP, claim that mental causes can produce effects at the same level without downward causation to the physical, and thus need not overdetermine their effects.

All views that I shall call ‘intralevelist’ views deny CSP, though they do so in different ways. The intralevelist approach favored by Marcus Schlosser (2009) appeals to the Causal Theory of Action as Schlosser understands it. Schlosser argues that the mental causes actions and that actions themselves take place at the mental level. Actions, like mental properties, supervene on physical realizers, but Schlosser understands supervenience to be a diachronic relation (or at least it *can* be a diachronic relation). This allows the action (A in the figure below) caused by the mental property instantiation (M) to partly share the mental property’s supervenience base (P). The thought is that because the action can supervene at least partly on P, the mental does not need to cause the later supervenience base (P’), and so P’ is not causally overdetermined. P causes P’, M causes A, and A supervenes on both P and P’.

Figure 3: Schlosser's Intralevelism



P and P' can thus be thought as temporal parts of a temporally extended supervenience base P*. It is P* that is the ontological grounds for A and M and so M does not need to cause P' in order to cause A.

John Gibbons' intralevelist account appeals to an analogy between mental causation and reasons-explanations of actions to explain how one mental property can cause another without causing its physical supervenience base (2006). Say that at t1 I want to quench my thirst and I want a beer.¹⁸ At t2 I reach into a cooler of beverages and I grab a beer. So, I have two desires followed by two actions. According to Gibbons, my wanting a beer is reason enough to grab a beer, but not reason enough to reach into the cooler. The reason for reaching into the cooler is that I want to quench my thirst. My wanting a beer is explanatorily relevant to my reaching into the cooler only to the extent that reaching into the cooler provides a means to getting a beer. So, each action is paired with a sufficient reason, but there is explanatory competition; after all, wanting to quench my thirst and wanting a beer are both explanatorily relevant to my reaching into the cooler. Gibbons thinks that it is perfectly reasonable to understand this as a case of my wanting a beer causing my grabbing a beer, despite this grabbing action being dependent on my reaching into the cooler, which in turn is already caused by my wanting to quench my thirst.

¹⁸ This example is from Andrei Buckareff (2011) and seems to capture the sense of what Gibbons is after a little better than Gibbons' own example involving a light switch.

What goes for wants and actions goes for mental causation, claims Gibbons: M can cause M' without causing P' even though M' depends on P'.

Both of these intralevelist accounts are unsatisfying for the following reasons: In Gibbons' case it seems obvious that the analogy he attempts to draw is a poor one.¹⁹ Mental properties supervene (and ontologically depend) on their physical realizers, but this is clearly not the same relation that desiring a beer instantiates with desiring to quench one's thirst, or that grabbing a beer instantiates with reaching into the cooler; I can desire a beer without being thirsty and I can grab a beer without reaching into the cooler (I could grab the beer from the fridge instead). The plausibility of CSP rests on the supervenience (and ontological dependence) of the mental on the physical. A case where there is no supervenience is just obviously not a case where CSP applies, and so appealing to such a case and then making a simple analogy to the mental is unjustified. Schlosser's account, on the other hand, depends on the very strange notion of diachronic supervenience, but there seems to be no reason at all to accept such a view. A's partial dependence on earlier physical supervenience base P seems an awful lot like causation smuggled in under another name.

I will not spend any more time critiquing either of these two views however, as I find Amie Thomasson's view to be a much more promising, though ultimately flawed, intralevelist account. What I have to say against Thomasson's view counts against intralevelist views in general, and so the extent to which I succeed against her view will count against Schlosser and Gibbons' views as well.

Thomasson's view is simply that causation happens at the mental level and at the level of the physical realizer, but never between these levels. Mental relata cause other mental relata and physical relata cause other physical relata, but the physical never causes the mental and the

¹⁹ Buckareff makes this same point (2011).

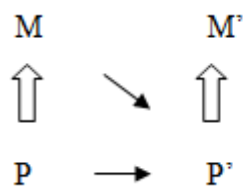
mental never causes the physical. The relation shared by the mental and the physical is what Thomasson simply calls ‘dependence’. ‘Dependence’ here (and throughout this thesis unless otherwise noted) broadly ‘refers to a synchronic relation between higher-level entities and lower-level entities that contribute to their being the way that they are’ (Thomasson, 1998, p. 184). Thomasson’s main claim is that these two types of explanation, causation and dependence, do not compete:

. . . Causation occurs only within a level; there is no upward or downward causation. Layers are not connected by causal relations but by relations of determination, dependence and material constitution (Thomasson, 1998, p. 187).

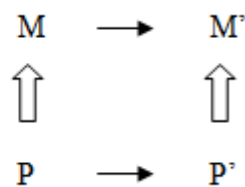
This view Thomasson calls the ‘Layered View’. According to the Layered View, causal relations depend upon, but are not reducible to the most fundamental causal relations instantiated at the level of particle physics. Below are two diagrams. The first is of the classical NRP picture that Kim attacks where CSP is assumed. The second is Thomasson’s view, where CSP is denied:

Figure 4: Classical NRP and The Layered View

Classical NRP (with CSP)



The Layered View (no CSP)



The Layered View is certainly the simplest of the intralevelist views and its elegance leaves it vulnerable to relatively few objections. I shall explore two broad criticisms of the Layered View, each of which counts against intralevelist views in general. First, there is the issue of whether a strongly supervenient property can be caused without causing that property’s supervenience/ontological dependence base; in other words, whether CSP can be denied. The

question of how exactly causation is inter-related with dependence is a thorny question over which Kim and Thomasson are in stark disagreement. According to Kim, the dependence relation excludes the causal relation, with the result that the Layered View only trades one type of overdetermination (causal overdetermination of P') for another (the hybrid overdetermination of M'). Here I use the expression 'hybrid overdetermination' to refer to cases where two different determination relations each sufficiently determine the same entity. In the Layered View, vertical dependence and horizontal causation both determine M', comprising a case of hybrid overdetermination. According to Thomasson, these two determination relations do not compete and we can speak of P' grounding M' and of M causing M' without any problems.

I shall tentatively side with Thomasson regarding this first worry. Hybrid overdetermination must not entail hybrid exclusion (exclusion of the causal relation by the dependence relation) as this would seem to entail two major problems: First is the issue of generalization. If M is excluded as a cause of M' because M' depends on P', and this exclusion falls directly out of hybrid overdetermination, what is to keep us from generalizing this exclusion to all supervenient property instantiations? Epiphenomenalism about supervenient properties such as acidity, mass, color and shape would be an unacceptable entailment of hybrid exclusion and so if hybrid exclusion generalizes, it ought to be rejected. I shall argue that Kim does not have a satisfactory response to the generalization worry and so hybrid exclusion and CSP are left in serious doubt. Because of this, hybrid exclusion should not count strongly against the Layered View or intralevelist views in general.

The second problematic entailment of hybrid exclusion is Ned Block's Drainage Problem (2003). Block has argued that the bottom-only causal picture painted by Kim not only means that causal powers 'drain' to the lowest level in the supervenience structure, it also means that

the existence of causation at all in the universe is dependent on an open question of physics, whether the universe is gunky or not. If the universe is gunky and physics has no bottom-most level, then causal powers would disappear entirely. Here again, I find Kim's response to this worry to be unsatisfying. The generalization problem and the drainage problem cast serious doubt on the claim that hybrid overdetermination entails hybrid exclusion and so Kim's objection to the Layered View on the basis of hybrid exclusion fails.

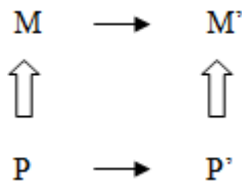
The second broad criticism of the Layered View is that it gives up too much in sacrificing mental-to-physical causation, and that just mental-to-mental causation is not enough to give us a causally efficacious mental that we should care about. Thomasson thinks that intuitive cases of interlevel mental causation can largely be cleared up as cases of 'loose talk' (1998, p. 191). I shall argue that no explanatory strategy along these lines will yield a satisfactory account of mental-to-physical causation, something any mental ontology must account for. Thomasson may be able to give an account of seeming physical-to-mental causation, cashed out in terms of dependence, but no alternative explanation can be found for mental-to-physical causation as the dependence relation is asymmetric. As the Layered View cannot countenance a satisfactory explanation in this area, it ought to be rejected.

Hybrid Overdetermination and Exclusion

CSP, the principle that states that to cause a supervenient property, that property's base must be caused, carries some strong initial appeal. Remember that supervenient properties are ontologically dependent on their base properties. So, if M' ontologically depends on P', then it is at least the case that P' necessitates M' and is sufficient for M'. Any explanation of the existence of M' that does not countenance P' would seem to be insufficient. The Layered View

does countenance P' as part of the explanation of the existence of M' , but Thomasson argues that the dependence relation instantiated by P' and M' does not compete with any mental cause of M' . Let us again set up the Layered View diagram:

Figure 5: The Layered View



M' is caused by M and also depends on P' . So, in LV there is no causal overdetermination of P' , but there is hybrid overdetermination of M' involving vertical dependence and horizontal causation. According to Thomasson, this hybrid overdetermination is completely innocuous and the two explanatory mechanisms do not compete to the exclusion of one of them.

Kim argues to the contrary, that dependence and causation compete and that dependence excludes causation. Considering a lump of bronze and the lump's yellowness which supervenes on some microproperty R , Kim makes the case for the dependence on R excluding any causal explanation of the lump's yellowness (such as being yellow at an earlier time):

As long as the lump has microproperty (R) at t , it's going to be yellow at t , *no matter what happened before t* . Moreover, unless the lump has (R), or another appropriate microproperty (with the right reflectance characteristic), at t , it cannot be yellow at t . Anything that happened before t seems irrelevant to the lump's being yellow at t ; its having (R) at t is fully sufficient in itself to make it yellow at t (Kim, 2005, p. 37).

Similarly, any antecedent mental properties are causally irrelevant, given that P' is fully sufficient for M' at t .

The Generalization Problem and the Drainage Problem

Does hybrid overdetermination entail hybrid exclusion? Epiphenomenalism and drainage give us good reasons to suspect that it does not. Hybrid exclusion as a general principle would entail that causation only happens at the bottom-most level of the supervenience structure, leaving all supervenient properties epiphenomenal: All supervenient properties are sufficiently accounted for by dependence relations and so are not the (direct) effects of any cause. These supervenient properties can cause no effects at any level of supervenience higher than the base as these effects would be excluded by dependence. So, the only level of the supervenience structure that can be an effect is the base, or bottom-most level. Only one level of the supervenience structure at a time can be a cause without causal overdetermination as only the base level can be an effect. It is implausible to preference any level with causal efficacy to the exclusion of the base level, and so only the base level can be either a cause or an effect, with every level above the base being epiphenomenal.

Generalization is a major worry for Kim's picture: To what extent does hybrid exclusion generalize and can Kim contain hybrid exclusion to only apply to mental properties? If all supervenient properties are excluded from causal efficacy by their supervenience base, then many properties that we thought efficacious turn out to be epiphenomenal. Such intuitively causally efficacious properties as acidity, fecundity, beauty, shape and mass would be excluded from causal efficacy. Kim must either deny hybrid exclusion and thus strip CSP of its underlying motivation, embrace epiphenomenalism, or tell a plausible story about how exclusion applies to mental properties in NRP, but not to the properties of the special sciences, aesthetic properties, ethical properties, or properties uniquely instantiated by composite objects.

In addition to the generalization problem, there is another problematic entailment of hybrid exclusion. To say that causal powers only occur at the bottom-most level of the supervenience structure pre-supposes that there is a bottom-most level, but as Ned Block points out, it is an open question, about which physicists disagree, whether there is an ultimate bottom level to physics (2003).²⁰ If there is no bottom-most level to physics then there is nothing stopping the drainage and it seems that there is no causation to be found anywhere in the physical universe.

It is important to note that the drainage objection is not supposed to depend on it in fact being the case that the universe is gunky. Block's point here is that if Kim's view is correct then whether there is causation in the universe rests on an open question in physics. That there is causation in the universe is not an open question, and so this putative feature of Kim's view, that causation is contingent upon an open question in physics, makes his view suspect (Block, 2003).

Kim's Response

The basic claim in hybrid exclusion is that if supervenient properties *cannot be reduced*, then they are epiphenomenal. Kim's response to the generalization and drainage problems keeps faith with this central claim, but stresses that many of the properties threatened by the generalization problem can in fact be reduced. If they can be reduced, then these properties get to share in the causal powers of their supervenience base. While these reducible properties lack *novel* causal powers, they are not epiphenomenal, argues Kim. Kim agrees that intuitively, the properties of the special sciences are causally efficacious and so they must be reducible to the properties they ultimately supervene on, properties of basic particle physics (2005, p. 54).

Reducing these properties is not the same as eliminating them. The whole point of Kim's project

²⁰ See Dehmelt (1989)

is ‘to clarify the options available to the physicalist: If you deem yourself a physicalist, you must choose between (reduction and epiphenomenalism).’ Kim is fine with that project being extended beyond mental properties to supervenient properties more generally.

When it comes to the drainage problem, Kim employs a similar tactic. Causal drainage only occurs when we have both supervenience *and* irreducibility. But, argues Kim, we have good reason to think that when it comes to purely physical supervening levels, higher levels can be reduced to lower levels. The property ‘being water’ is generally regarded as being identical to the property ‘being H₂O’. Such identities probably hold between H₂O at the molecular level and its constituents at the level of atoms, and between the atoms and their constituents at the sub-atomic level, and so on.

Unless we have reason to think that irreducibility will hold ‘all the way down,’ we have no reason to think that causal drainage will go on forever. Reduction is the stopper that will plug the cosmic hole through which causal powers might drain away (Kim, 2005, p. 68).

I am not convinced that Kim’s responses here vindicate hybrid exclusion. First, regarding his response to the drainage problem, it is not obvious that all physical levels reduce downwardly. Kim argues that the drainage objector must claim that irreducibility holds all the way down, but actually the burden of proof is on Kim to establish that reduction holds all the way down. Suppose the universe is gunky, and suppose that we have good reason to think that physical levels reduce downwardly from the level of ordinary objects down to ten levels below the standard model of particle physics. What is to make us think that reduction will hold between the tenth level below the standard model and the eleventh? For that matter, what reason would we have to posit reduction between 1,000th and 1,001st levels below the stand model? If

irreducibility obtains between any two levels, then all levels higher than that level are cut off, as it were, from the ultimate source of causation. The drainage objector does not need to posit irreducibility all the way down. She simply needs irreducibility to hold somewhere in the structure. Even if the irreducibility only obtains fairly high in the structure, say between the molecular and the atomic levels, then on Kim's picture ordinary objects are excluded from efficacy. If the irreducibility obtains between two deep levels, say between the 100th and the 101st levels below the standard model, then even if reduction holds from the 101st on down, we still have an unrecognizably epiphenomenal universe. What Kim needs is some sort of general principle ensuring that reduction holds among all physical levels. He does not offer such a principle and it is not clear how one could be justified.

If Kim could produce such a principle, then he would indeed have a strong response to the drainage problem. His response to the generalization problem is another matter. For supervenient properties in general to have causal powers on Kim's picture, they must be reduced. The success of his response will depend on how successfully we can reduce all the supervenient properties we take to be causally efficacious. Some supervenient properties seem to lend themselves to reduction more than others. Properties of the special sciences may have an especially good chance of being reducible. A property like fecundity can perhaps be functionalized and then reduced to properties at the chemical level and ultimately to properties at the most basic level of physics.

However, there are reasons to resist reduction even in these cases. One reason is multiple realizability. Just as mental properties can be realized by many different physical properties, so too can properties like fecundity, adaptability, etc., in which case, goes the objection, they cannot be reduced. While multiple realizability is a pretty conclusive objection to *type* reduction, I take

Kim to be correct in resisting this sort of objection in these cases. It is property *instances* that are being considered for reducibility. If the general kind to which those instances belong is multiply realizable, then it just means that there are different reduction bases for those various instances (Kim, 2005, p. 56).

Maybe there are other reasons to resist reduction in the case of special science properties and maybe they succeed. Even if they do not, there are other classes of supervenient properties that seem less likely to be reducible. If they cannot be reduced, then by Kim's lights, they are epiphenomenal. Such classes of properties include ethical properties and aesthetic properties. The success of Kim's response to the generalization problem will depend on the extent of properties left unreduced. I do not have the space here to begin to tackle a survey of this scope, but I suspect that there will be a pretty significant remainder of unreduced properties that will have to be counted as epiphenomenal and that this number will be large enough for the generalization problem to carry considerable force. Does Kim's hybrid exclusion and CSP survive the generalization problem and the drainage problem? It is far from clear that they do. At the very least, these problems cast serious doubt on hybrid exclusion; enough doubt, I contend, that CSP can hardly be used to decisively put down intralevelist views.

Mental-to-Physical Causation

If Kim is right about hybrid exclusion, then the exclusion argument is much simpler and stronger than we thought. Simply by being irreducible and supervenient, the mental would be out of the running for being causally efficacious. Since hybrid exclusion is highly dubious, Kim cannot succeed quite so easily, and must instead draw upon the Weak Closure principle to generate exclusion. Assuming that the mental can cause things, and assuming it can directly

cause mental effects, we still have widespread causal overdetermination if the mental regularly causes physical effects, as per Weak Closure these physical effects already have sufficient physical causes. The Layered View purports to avoid this rendition of the exclusion argument by denying that the mental ever causes anything physical. This should immediately raise some worries. If the causal efficacy of the mental is something that must be vouchsafed by any mental ontology, is mere mental-to-mental causation enough? Intuitively some mental effects have physical causes (like a drug alleviating pain) and some physical effects have mental causes (like my intention to go to the store resulting in my bodily movement to the store). The Layered View succeeds in vouchsafing mental causation only if it can provide a satisfactory alternative explanation for these sorts of interlevel causal intuitions.

Regarding physical-to-mental causation, there may indeed be an alternative story to tell involving dependence. When I take aspirin for my headache, certain properties of the drug engage in causal relations with physiological properties at that level in my body. Eventually the causal sequence progresses until the supervenience base of the mental state ‘intense headache’ gives way to the supervenience base of the mental state ‘mild headache’. So, what might be ordinarily understood as physical-to-mental causation is instead understood as physical-to-physical causation plus dependence in the Layered View.

However, if this story succeeds in explaining putative physical-to-mental causation, it obviously cannot be reversed to explain putative mental-to-physical causation as the dependence relation appealed to is asymmetric. In fact, there exists no relation in the NRP-ist’s toolbox to explain seeming downward causation other than causation itself. Bottom-up physicalist metaphysics will not tolerate a dependence relation that extends downward. At best, a horizontal relation like causation can be bent downward as in the classic NRP picture, but this option is

ruled by the Layered View to avoid causal overdetermination. Thus, in the Layered View, antecedent mental states have no meaningful explanatory relationship with subsequent physical states.

This seems obviously wrong. Actions are at least explained, if not caused, by intentions. Agency itself requires intentionality, a paradigmatic mental state. Even if we locate intentionality at a physical level, there seem to be causal relations between intentions and other mental states such as beliefs, sensory experiences and memories. For example, I intended to go to the store because I saw that I was out of flour, remembered that the recipe I was making required flour, and formed the belief that I needed to get some flour. As the Layered View cannot offer an alternative explanation to downward causation (and, in fact, has no downward explanatory relation to draw upon at all) we must reject it for failing to provide an adequate account of mental causal efficacy.

Thomasson recognizes the Layered View's inability to provide an account of mental-to-physical causation as a serious problem for the Layered View:

It is somewhat more difficult to provide a satisfactory analysis of cases of purported mental/physical causation (more difficult than for physical-to-mental causation) on this model. Yet there are mitigating factors which can make this easier to accept

(Thomasson, 1998, p. 191 – my parenthetical note).

The first mitigating factor for Thomasson is to remember that on the Layered View we can at least have direct causation among mental events and indirect upward causation from mental events to higher level events (by directly causing mental events that those higher-level events depend on). So, the baseball shattering the window cannot be said to be caused by Julie's intention to shatter the window, but Julie's intention to shatter the window can be said to

indirectly cause higher-level events like ‘the school’s suffering yet another act of vandalism’ (Thomasson, 1998, p. 192).

Thomasson’s second mitigating factor is to remember that the alternatives to the Layered View are not attractive:

Once the alternatives are made evident, many nonreductivists might find that the costs of positing only higher-level causal relations fade in comparison to the costs of losing the distinctive status of the mental or accepting overdetermination and violating the causal closure of the physical (Thomasson, 1998, p. 192).²¹

I do not find either of those factors to mitigate much the unattractiveness of the Layered View’s inability to cash out mental-to-physical causation. Regarding the first factor, it is fine and well that the mental can cause *something* in the Layered View, but the objection is just that the mental cannot cause *physical* things in the Layered View. No doubt the Layered View is superior in this respect to a total epiphenomenalism about the mental, but surely no one seriously endorses that view. The second factor is worth considering and in the final section I will treat this issue more in depth; suffice it to say that all extant mental ontologies have their bullets to bite. However, it is worth wondering whether Thomasson is right in estimating widespread overdetermination or reduction to be less plausible commitments than a mental that cannot, directly or indirectly, cause anything physical. But, Thomasson’s point is well taken in that I do not find any of the three alternatives to be very attractive.

It is worth pointing out that the Layered View’s inability to countenance downward causation generalizes in a way not unlike the way the exclusion problem generalizes too far for Kim. According to the Layered View, some supervenient property instantiations are causally

²¹ Implied in the above statement is that Thomasson does not find Weak Closure to be a satisfying articulation of the causal closure of physics; more to follow on this point in the final section.

efficacious, but those properties do not cause downwardly, nor can any downward explanatory relation be appealed to. This would seem to imply that the loss of jobs cannot be explained by a bad economy, or that the death of some particular animal cannot be explained by a rival species outcompeting that animal's species for resources in an ecosystem. So it seems that the Layered View has its own generalization problem.

Conclusion

To conclude the critical part of this thesis, classical NRP does entail widespread overdetermination. Bennett's appeal to a compatibilist definition of 'overdetermination' does not succeed, even on her definition. Even if it did, redefining 'overdetermination' does not alleviate any of the original worries about overdetermination as these worries are generated by a feature of NRP that all NRP-ists must acknowledge, that there are redundant sufficient causes for any effect of the mental.

Intralevelist views such as Thomasson's Layered View would save NRP from widespread causal overdetermination by denying that causation can occur between levels. It is far from clear that hybrid overdetermination leads to hybrid exclusion, as Kim argues it does. If it did, then hybrid exclusion would generalize, leaving too many supervenient property instantiations epiphenomenal, and the drainage problem would suggest that causation is contingent on the universe not being gunky, which is an open question of physics. As it is far from clear that hybrid exclusion is entailed by hybrid overdetermination, CSP is insufficiently motivated and direct mental-to-mental causation is safe. There are, however, other reasons to reject intralevelist views like the Layered View: By denying interlevel causation, the physicalist is left with no possible downward explanatory mechanism. This is unacceptable, not only

because mental-to-physical explanation is ruled out, but also because the problem generalizes to implausibly rule out any downward explanation among supervenient entities.

So, NRP is in a quandary. On intralevelist interpretations of NRP, NRP is unable to provide an essential tenet of any mental ontology, downward explanation from the mental. The only possible downward explanatory mechanism available to the physicalist is causation, primarily a horizontal relation which perhaps may be bent. If causation can extend downward, as the classical formulation of NRP depicts, then we have widespread causal overdetermination as any physical effect of the mental already has a sufficient physical cause. I suppose if the NRP-ist must choose, the classical account is to be preferred to the intralevelist account as losing mental-to-physical explanation is completely unacceptable. Then, however, the NRP-ist must embrace widespread causal overdetermination. Some have certainly taken this route, but it seems, at the very least, to be a nasty pill to swallow.

CHAPTER 4

BI-DIRECTIONAL DETERMINATION

Up to this point, my thesis has been entirely negative: classical NRP, a leading mental ontology that purports to give us irreducible and causally efficacious mental states without surrendering basic physicalist commitments, entails widespread causal overdetermination. I have assumed that this entailment means we should reject classical NRP. Intralevelist views avoid overdetermination, but at perhaps the even greater cost of mental-to-physical causation. It would be reasonable at this point to wonder what the right account of mental ontology is. The general consensus about mental ontology seems to be that we must sacrifice at least one of our pre-theoretical commitments to have a coherent theory; everyone must choose a bullet to bite. The classical NRP-ist may very well choose to bite the overdetermination bullet. Intralevelists may choose to bite the mental-to-physical causation bullet. Kim, by endorsing a hybrid reductivist account bites the efficacy bullet, at least concerning qualitative mental states (or, you might say he bites the irreducibility bullet regarding behavioral mental states). Less popular, one could reject the physicalist commitment of causal closure.

In this final section of the thesis, I want to motivate a mental ontology that bites none of the above bullets. My theory, if correct, will allow for an irreducible mental that is causally efficacious in all the ways we care about, does not violate causal closure, and does not entail overdetermination. Of course, my theory does this only by biting a different bullet; it denies supervenience. In a nutshell, the theory I am tentatively suggesting, which I call ‘Bi-Directional Determination Theory’ or ‘BDT’, is an intralevelist view that posits determination relations²² not only from the physical to the mental, but also from the mental to the physical. If the mental

²² There will be more details in a later section on what this determination relation might be. At this point it is enough to say that it is synchronous and non-causal and performs roughly the same role as ‘dependence’ plays in the Layered View.

determines the physical, then the mental in BDT does not supervene on the physical, at least not in the strong sense involving ontological dependence. Hence, BDT is not a physicalist mental ontology as the strong supervenience of the mental on the physical is a basic commitment of physicalism.²³

Concerning the Lack of Better Alternatives

Before I attempt to spell out the nuts and bolts of BDT, I want to employ a similar tactic to Thomasson's second mitigating factor; that is, I want to convince the reader of the bleakness of the landscape of competing mental ontologies, at least for those holding a certain set of pre-theoretical intuitions. Suppose I am firmly committed to there being mental states that are both irreducible and causally efficacious. Suppose I find widespread overdetermination to be very problematic, and suppose I think mental-to-physical causation must be vouchsafed at any cost. Finally, suppose that I find Weak Closure to be an insufficient articulation of the closure commitment; that is, I find direct causal relations between physical and non-physical relata to be highly implausible. This seems like a reasonable set of pre-theoretical commitments to hold, and probably comes pretty close to our naïve view of the mind. None of the views considered thus far satisfy this set of commitments. BDT purports to.

At this point in the shared dialectic on mental ontology, there are some pretty entrenched positions and many philosophers have talked themselves into believing that the bullet their view has chosen to bite is not so unappetizing after all. Hence, many philosophers will not feel the motivation of BDT, and that is fine. My aim here is to simply point out that there is an area of logical space here that no extant mental ontology has filled and that is, as a matter of fact, no less

²³ Although the denial of supervenience certainly discounts BDT as a physicalist theory, I shall argue later on that BDT does much better with respect to the other basic physicalist commitment of causal closure than NRP, an ostensibly physicalist theory, does.

attractive an area to fill than any other (more attractive than the others, I think). I have no way of arbitrating between pre-theoretical commitments and, no doubt to some, a rejection of supervenience, and the rejection of ontological physicalism that such a rejection entails, will be a much less attractive bullet to bite than the one they already prefer. For such philosophers BDT is a non-starter in much the same way that the Layered View is a non-starter; the bullet to bite is just too big. My tactic here is to expose this entailment of BDT up front and to embrace it. Even if my reader finds that the rejection of supervenience and physicalism makes BDT a non-starter, hopefully she will at least appreciate that every alternative also violates a plausible pre-theoretical commitment for someone. Which pre-theoretical commitments are considered more sacrosanct than others will no doubt vary, and I have no intention here of arbitrating among them.

The Central Challenge of BDT

As mentioned earlier, BDT is a kind of intralevelist theory. Like the Layered View, BDT does not posit direct causal relations between different levels. Unlike the Layered View, BDT countenances a downward synchronous ('vertical') determination relation from the mental to the physical. What this relation is, and how to understand it, is the central challenge of BDT. The primary worry about positing a bi-directional determination relation is that some vertical relations, like ontological dependence or grounding, must be asymmetric. It is hard to know what it could even mean to say that A ontologically depends on B *and* B ontologically depends on A. If such a statement were to turn out true, we would probably conclude that A and B are in fact identical and not two distinct co-dependent things. For BDT to get past this initial worry we must either show that the vertical relation between the mental and physical is not of the

asymmetric variety, or we must individuate the relata on both the mental and physical side, so that it is not the case that the same relata are both grounded in and the ground of some other relata. In this section, I will articulate three different BDT models. The first model takes the first strategy of weakening the vertical relation, while the second and third models take the individuated relata strategy. I will ultimately endorse the second BDT model, but each model faces unique objections and by exploring each model we can arrive at a better grasp of the issues at stake.

BDT1

Suppose the vertical relation between mental and physical relata is non-asymmetric, but that the mental and physical are indeed distinct. What sort of relation could play such a role? It would need to be weak enough that we could sensibly say that M vertically determines P *and* P vertically determines M (where ‘determines’ is a sort of stand-in for whatever relation we ultimately posit), but also strong enough that it can be explanatory and serve to cash out indirect causation, in the way that ontological dependence does in the Layered View. So, for example, ontological dependence is a strong enough relation to be indirect causation bearing and explanatory. There was no difficulty in understanding Thomasson’s way of cashing out physical-to-mental causation in the Layered View: P causes P’ and P’ grounds M’. We can easily explain M’ in terms of its dependence on P’ and we can sensibly speak of P indirectly causing M’. To switch to a culinary example that Thomasson uses (1998, p. 183), where the cookies are put in the oven at t1 and are burnt at t2, we can say that at t2 the cookies are burnt *because of* the microstructure of the cookies at t2 (the microstructure that the macro-level description of the cookies ontologically depends on at t2). We can also say that changes in the

micro-structure of the cookies between t_1 and t_2 *caused* the cookies to be burnt at t_2 , though this causation is indirect in the Layered View. Ontological dependence, however, is too strong to be non-asymmetric. M and P singularly understood (not individuated into different entities at those levels) cannot ontologically depend on each other.

What kind of relation is weak enough to be non-asymmetric? Consider the following example: Imagine a sealed, gas-filled sphere whose surface is composed of a series of overlapping metal scales. At the push of a button, one can mechanically manipulate the surface area of the sphere by causing the scales to slide out of each other. The way the scales are overlaid guarantees that the only shape this closed object can take is a perfect sphere. All that can be manipulated is the size of the sphere. As the scales extract from one another, valves which only allow gas to flow into the sphere let in gas to relieve the negative pressure inside the sphere. There is a second button one can push to manipulate our sphere. This button activates a nozzle that forces gas into the sphere from a pressurized tank. As more gas is added to the sphere, the scales are allowed to expand to relieve the pressure inside, always perfectly maintaining the sphere's shape.

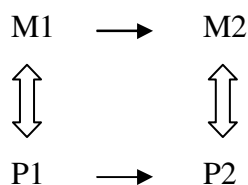
Pressing the first button seems to directly cause the sphere's surface area to increase, and pressing the second button seems to directly cause the sphere's volume to increase. Let changes to the sphere's surface area be S-changes and changes to the sphere's volume be V-changes. S-changes co-vary with V-changes in the above case as the closed object always remains a sphere. This co-variation is simultaneous. There may be a temporal lag between the build-up of pressure inside the sphere after pressing the second button and the expansion of the sphere's volume, but the instant the volume increases is the same instant the surface area increases. A similar temporal lag may occur between the mechanical expansion of the surface area and the inflow of

gas to relieve the negative pressure, but again, the instant the surface area increases is the instant the volume increases. This simultaneity rules out the possibility that the relation between the surface area and the volume is a causal relation as causal relations are generally understood to be diachronic.

In the above scenario, it seems plausible to posit a non-asymmetric determination relation between the surface area and the volume of the sphere. The laws of nature and the structure of the sphere are such that S-changes and V-changes are synced. S-changes determine V-changes and V-changes determine S-changes. One does not asymmetrically depend on the other the way changes to the aesthetic properties of a painting depend on the purely physical or natural properties of the painting.

The first strategy for dealing with the central challenge facing BDT must involve positing a relation between the mental and the physical similar to the relation between the surface area and volume of the sphere, a relation weak enough to be bi-directional. The two causal chains are synced up and any M-change necessitates a P-change and vice-versa. For the mental and physical something would need to play the constraining role of the metallic sphere, ensuring that P-changes really do necessitate M-changes. This role would most naturally be filled by psycho-physical laws. This articulation of BDT would look something like this:

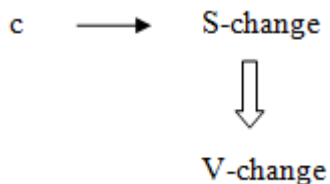
Figure 6: BDT1



The bi-directional arrows represent this weaker vertical relation guaranteeing the co-variation of Ms and Ps.

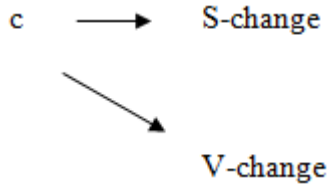
While BDT1 does avoid the central worry about the incoherence of bi-directional determination (I take it that there is nothing incoherent about the metallic sphere example), it has a fatal flaw: while this relation is weak enough to avoid the incoherence worry, it is not strong enough to be used to cash out indirect causation. The vertical relation between the mental and physical needs to be an explanatory or change-making relation, but this weaker relation, whatever it is, is not that kind of relation. In the sphere case, the S-changes do not sufficiently explain the V-changes. The V-changes cannot be explained without appeal to the causal story happening with the button being pushed and the scales sliding. The relation between the surface area and the volume is not causal, but neither can it be used to cash out indirect causation in the way that dependence can. When the first button is pressed a causal chain is initiated that has two direct effects, the increase of the surface area *and* the increase of the volume. The co-variation of the surface area and volume is just a contingent feature of the situation; the relation does not *make* the volume co-vary with the surface area in the way that ontological dependence *makes* the aesthetic properties co-vary with the natural properties of the painting. In other words, for BDT1 to work as intended, the metallic sphere case would have to work like this (where c is the pressing of the first button):

Figure 7: Co-variation, Dependence



Instead, it far more plausible to understand the case like this:

Figure 8: Co-variation, Common Cause



If S-changes and V-changes are best understood as two effects with a common cause, then the metallic sphere case cannot be appealed to as a model for mental-to-physical indirect causation.

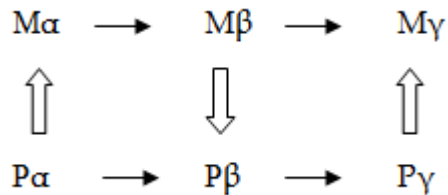
BDT1 is probably best understood as a sort of occasionalism. The relation between the mental and physical would not be directly causal, nor could the relation be used to cash out indirect causation. Instead mental and physical changes would both be effects of a common cause, just as the S-changes and V-changes are effects of common causes involving either the first or second button being pushed. The efficacy of the mental was something BDT was supposed to vouchsafe, and an occasionalist view fails in that respect, providing only a correlation between the mental and physical, but no explanatory or change-making relations capable of cashing out causation between the two levels. This, I take it, is the best we can do in formulating BDT along the lines of the first strategy of weakening the determination relation enough to be non-asymmetric. The failure of the metallic sphere case gives us reason to suspect that any relation between the mental and physical weak enough to be sensibly non-asymmetric will not be strong enough to bear indirect causal relations in the way that dependence does in the Layered View.

BDT2

So, the first attempt to articulate BDT, using the weaker notion of determination fails and we are left with the second strategy for avoiding the incoherence worry; individuating the rela

on both the mental and the physical side so that the same relata are not grounds for and grounded in other relata. The picture might look something like the following:

Figure 9: BDT2



Mental states of type α are always grounded in physical states of type α , while physical states of type β are always grounded in mental states of type β . Mental states can directly causally interact with other mental states, and physical states can directly causally interact with other physical states, but there is no direct mental-to-physical or physical-to-mental causation.²⁴ Unlike the Layered View, there is indirect causation in both directions; $M\alpha$ indirectly causes $P\beta$, just as $P\beta$ indirectly causes $M\gamma$.

The vertical relation in this articulation of BDT is a strong relation of ontological dependence or grounding, a relation strong enough to be explanatory and indirect causation bearing. Just as the aesthetic properties of a painting are dependent on or grounded in the purely natural properties of the painting, so $M\alpha$ is grounded in $P\alpha$ and $P\beta$ is grounded in $M\beta$. We can sensibly speak of changes to the purely natural properties of the painting causing changes to the aesthetic properties, though strictly speaking the relation between the natural and aesthetic properties is non-causal. When I spill ink on the painting, I cause a change to the natural properties of the painting and this in turn grounds the changes to the aesthetic properties, but natural properties do not directly cause aesthetic properties; they ground them.

²⁴ Note that Kim's hybrid exclusion worry once again appears as $P\beta$ is hybridly overdetermined. I will address the issue of hybrid exclusion in BDT in the next section.

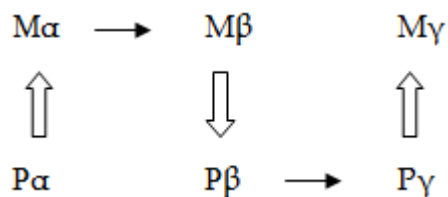
By individuating the relata on the mental and physical sides, BDT2 avoids the central worry for BDT models. In BDT2, it is never the case that the same relata ground and are grounded in some other relata. There is nothing incoherent about the figure above, even if it flies in the face of a bottom-up physicalist picture where the physical has ontological primacy.

BDT3

BDT2 entails hybrid overdetermination. As discussed in section three, it is far from clear that hybrid overdetermination entails hybrid exclusion. If it does, then BDT2 is no better off than the Layered View in this respect. In fact, one might think that BDT2 is even worse off. In the Layered View, it would just be the mental that is excluded, whereas in BDT2, exclusion would poke holes all over the mental/physical interface as both mental and physical entities are dependent (supervenient) on and irreducible to each other. Suffice it to say that if hybrid overdetermination entails hybrid exclusion, then both the Layered View and BDT2 are in serious trouble.

However, unlike the Layered View, BDT can be re-articulated yet again to accommodate hybrid exclusion. The new picture would look something like this:

Figure 10: BDT3



In BDT3, no entity is hybridly overdetermined and dependence never has the chance to exclude causation. Something like the above picture would seem to require some fairly complicated causal laws such that physical relata of type β can cause physical relata of type γ , but cannot be

caused by physical relata of type α . Then again, if hybrid exclusion is some sort of general constraint on causation, perhaps these causal laws simply fall out of the ontological dependence relations holding among the different types of Ms and Ps.

While BDT3 allows for hybrid exclusion, it has a rather startling entailment: There are all kinds of Ps that do not have direct sufficient physical causes. For that matter, these Ps do not have *any* direct sufficient cause, only indirect causes through their dependence base Ms. This is startling for two reasons. First, if there are physical entities whose existence has no physical cause, we might think that they are at least very rare; certain quantum phenomena and the pre-Big Bang totality of matter and energy are the only candidates that come to mind. It would be at least surprising if some fairly commonplace neural structures were routinely brought into existence without a physical cause.

The second reason this should give us pause is that whether $P\beta$ could come into existence without any physical cause seems to be an empirically verifiable claim. If physical structures were coming into existence without any sufficient physical cause, it seems that neuroscience would have detected such mysterious structures by now. At the very least, a perfected neuroscience could give a full account of the physical structures of the nervous system and their physical history. It should strike us as unlikely that in the perfected neuroscience we would find neural structures with short physical causal histories.

I contend that these startling features of BDT3 render this model of BDT untenable. Instead we ought to endorse BDT2. While BDT2 is vulnerable to hybrid exclusion worries, these worries are themselves suspect for their apparent entailment of widespread epiphenomenalism and causal drainage worries, as discussed in section three of the thesis. BDT2 avoids the uncaused physical structures objection as BDT2 is consistent with there being

complete sufficient physical causal chains extending from neural structures back to the first physical cause.

BDT2 is still susceptible to a more general physicalist worry; that it seems unlikely that physical entities would be ontologically dependent on non-physical entities. This, however, is simply the physicalist intuition that BDT admits of not being able to satisfy. While the ontological primacy of the physical would be a nice commitment to keep, remember that one cannot have everything one wants when it comes to mental ontology. By sacrificing the ontological primacy of the physical BDT gains an irreducible and causally efficacious mental while avoiding the overdetermination entailment of NRP.

Another Virtue of BDT

While the denial of supervenience discounts BDT as a physicalist mental ontology, it is worth pointing out that in one respect BDT does much better regarding physicalist intuitions than NRP does. BDT allows for robust causal closure:

Robust Causal Closure: Physical entities cannot directly causally interact with non-physical entities.

Robust Closure does much better than Weak Closure at getting at the physicalist intuitions undergirding closure principles in general. I take it that there are two reasons for thinking the physical system to be causally closed. One reason is because of *a priori* considerations: It just seems almost impossible to conceive of how a thing lacking all physical properties could cause changes in a physical thing. To cause a physical change is to cause a change describable in physical terms, such as a loss or gain of mass or energy. How could such a change be produced

by an immaterial thing? This worry dates back at least to Descartes and Princess Elisabeth of Bohemia's famous objection to Cartesian Dualism.

The second reason to think that physics must be closed is a more recent concern, and that is the lack of empirical evidence in support of non-physical causes with physical effects. Effects lacking physical causes would be great evidence against physicalism, but such mysterious events have yet to be empirically verified.

While Cartesian Dualism violates both kinds of causal closure intuitions, it is important to note that NRP still violates the *a priori* closure intuitions. Weak Closure, by merely claiming that all caused physical events have a sufficient physical cause, only gets NRP around the empirical worry. It is still highly mysterious how the mental in NRP, being irreducible to the physical, causes physical effects. Princess Elisabeth would still be perplexed. Robust Closure, while respecting both kinds of closure intuitions, would rule out NRP from the outset.

BDT, on the other hand, by positing no direct causal relations, allows for Robust Closure. In this respect it not only does better than Cartesian Dualism, but also better than NRP, an ostensibly physicalist theory. While BDT is not a physicalist theory, and while many physicalists will no doubt find the rejection of supervenience to be too high a price to pay for irreducible mental causation without overdetermination, perhaps BDT's ability to embrace Robust Closure will mitigate the physicalist intuitional fallout.

Conclusion

NRP entails widespread causal overdetermination despite the best efforts of its defenders. If we are committed to a causally efficacious mental that is also irreducible, then we must either embrace widespread overdetermination and reject Robust Closure, or endorse BDT. By

endorsing BDT, one does depart from physicalism as BDT denies supervenience, but perhaps in light of the sacrifices other mental ontologies must also make, denying supervenience will be seen as a reasonable cost. Much more work needs to be done, but at the very least, BDT is a theory deserving of some further consideration as it offers unique theoretical advantages over its competitors.

BIBLIOGRAPHY

- Anscombe, E. (1975). Causality and Determination. In E. Sosa (ed.) *Causation and Conditionals*. Glasgow: Oxford University Press.
- Bealer, G. (2007). Mental Causation. *Philosophical Perspectives*, 21 23-53.
- Block, N. (1980). Are Absent Qualia Possible? *The Philosophical Review*, 89 (2) 257-274.
- Block, N. (2003). Do Causal Powers Drain Away? *Philosophy and Phenomenological Research*, 67 (1) 133-150.
- Bennett, K. (2003). Why the Exclusion Problem Seems Intractable, and How, Just Maybe, to Tract It. *Nous*, 37 (3) 471-497.
- Buckareff, A. (2011). Intralevel Mental Causation. *Frontiers of Philosophy in China*, 6 (3) 402-425.
- Chalmers, D. (1996). *The Conscious Mind*. New York: Oxford University Press.
- Churchland, P. (2007). The Evolving Fortunes of the Eliminative Materialist. In B. McLaughlin and J. Cohen (eds.) *Contemporary Debates in Philosophy of Mind*. Malden, MA: Blackwell Publishing.
- Dehmelt, H. (1989). Triton, Electron, Cosman. . . : An Infinite Regression? *Proceedings of the National Academy of Sciences*, 86, 8618-8619.
- Fair, D. (1979). Causation and the Flow of Energy. *Erkenntnis*, 14 219-250.
- Gibbons, J. (2006). Mental Causation without Downward Causation. *The Philosophical Review*, 115 (1) 79-103.
- Hall, N. (2004). Two Concepts of Causation. In J. Collins, N. Hall and L.A. Paul (eds.), *Causation and Counterfactuals*. Cambridge, MA: MIT Press.
- Horgan, T. (1997). Kim on Mental Causation and Causal Exclusion. *Philosophical Perspectives*,

11 165-184.

Kim, J. (1998). *Mind in a Physical World*. Cambridge: MIT Press.

Kim, J. (2001). Lonely Souls: Causality and Substance Dualism. In K. Corcoran (Ed.), *Soul, Body and Survival: Essays in the Metaphysics of Human Persons*. Ithaca: Cornell University Press.

Kim, J. (2005). *Physicalism or Something Near Enough*. Princeton: Princeton University Press.

Kim, J. (2007). Causation and Mental Causation. In B. McLaughlin and J. Cohen (eds.) *Contemporary Debates in Philosophy of Mind*. Malden, MA: Blackwell Publishing.

Levine, J. (2001). *Purple Haze*. Oxford: Oxford University Press.

Lewis, D. (1980). Mad Pain and Martian Pain. In N. Block (ed.) *Readings in the Philosophy of Psychology*. Cambridge: Harvard University Press.

Putnam H. (1967). Psychological Predicates. In W.H. Capitan and D.D. Merrill (eds.) *Art, Mind, and Religion*. Pittsburgh: University of Pittsburgh Press.

Salmon, W. (1980). Causality: Production and Propagation. *PSA: Proceedings of the Biennial Meeting of the Philosophy and Science Association*, 2 49-69.

Schlosser, M. (2009). Non-Reductive Physicalism, Mental Causation and the Nature of Actions. In A. Heike and H. Leitgeb (eds.) *Reduction: Between the Mind and the Brain*. Frankfurt: Ontos Verlag.

Shoemaker, S. (1981). Some Varieties of Functionalism. *Philosophical Topics*, 12 (1) 93-119.

Shoemaker, S. (2007). *Physical Realization*. New York: Oxford University Press.

Sider, T. (2003). What's so Bad about Overdetermination? *Philosophy and Phenomenological Research*. 67 719-726.

Thomasson, A. (1998) A Nonreductivist Solution to Mental Causation. *Philosophical Studies*,

89 181-195.

Yablo, S. (1992) Mental Causation. *The Philosophical Review*, 101 (2) 245-280.