THE UNIVERSITY *of* EDINBURGH

Physicalism and the Causal Exclusion Argument

Jonas Fogedgaard Christensen

PhD

The University of Edinburgh

2010

Declaration

Pursuant to The University of Edinburgh's Postgraduate Research Assessment Regulations, (section 2.5) I hereby declare that the thesis has been composed by me, that the work is my own, and that it has not been submitted for any other degree or professional qualification.

Jonas Fogedgaard Christensen, 30/8 2010.

Acknowledgements

Many thanks to my supervisors, Dr. Jakob Hohwy, Prof. Andy Clark, and (especially) Dr. Jesper Kallestrup, to the Danish Research Council for generous financial support, and to my wife Elodie. I have benefited greatly from discussions with many people over the years – too many to mention. Thanks to you all.

Table of Contents

Introduction 1

Chapter 1: Defining Physicalism

1.1 Introduction 7
1.2 Physics 8
1.3 Being Physical and Hempel's Dilemma 9
1.4 Do Fundamental Chemical or Biological Properties Falsify Physicalism? 13
1.5 Physicalism and Being 'Nothing over and Above' 15
1.6 From Supervenience Physicalism to Grounding Physicalism 17
1.7 What if there are no Fundamental Properties? 21
1.8 Grounding and Supervenience 23
1.9 Grounding Physicalism and Standard Supervenience 24
1.10 Grounding Physicalism and Minimal Supervenience 26
1.11 Strong and Global Supervenience 28
1.12 The Modal Strength of the Relevant Supervenience 28
1.13 Grounding of the Non-Physical in the Physical: Which Supervenience is Entailed? 32
1.14 Concluding Remarks 35

Chapter 2: Emergentism and Reductive vs. Non-Reductive Physicalism

2.1 Emergentism 36
2.2 Using the Nomic/Metaphysical Distinction to Sort Emergentism from Physicalism 39
2.3 Physical and Nomic Supervenience 39
2.4 Physical Supervenience and Emergentism 41
2.5 What is a Non-Physical Law? 42
2.6 Emergentism and Dispositional Essentialism 45
2.7 Howell: Trans-Ordinal Laws Contaminate the Physicalist Grounding Base 46
2.8 Emergentism is Incoherent if Dispositional Essentialism is True 48
2.9 Concluding Remarks on Physicalism and Emergentism 49
2.10 Reductive and Non-Reductive Physicalism 50

2.11 Non-Physical Properties and Broadly Physical Properties 50
2.12 Generating Broadly Physical Properties from Fundamental Physical Properties 51
2.13 Applying the Distinction Between Non-Physical and Broadly Physical Properties 53
2.14 General Worries about Composite Properties 54
2.15 Concluding Remarks 55

Chapter 3: The Causal Exclusion Argument and the Drainage Problem

3.1 Causes 57
3.2 Causal Powers 58
3.3 Causation 59
3.4 The Causal Exclusion Problem 60
3.5 Mental Causation 60
3.6 The Causal Closure Principle 61
3.7 The Causal Exclusion Principle 64
3.8 Distinctness 66
3.9 The Exclusion Argument 67
3.10 Tidying Up the Argument 68
3.11 Solving the Exclusion Problem 70
3.12 The Generalization Problem 71
3.13 The Drainage Problem 73
3.14 Kim on the Drainage Problem: Levels versus Orders 73
3.15 Composite Overdetermination 74
3.16 Kim on the Drainage Problem: Inter-level Reduction 78
3.17 Causal Drainage and Inter-level Reduction of Properties 79
3.18 Against Inter-level Reduction as a Solution to the Drainage Problem 81
3.19 Solving the Drainage Problem I 83
3.20 Solving the Drainage Problem II 85
3.21 Inter-level Reduction of Causal Powers 86
3.22 Reduction of Powers and Causal Overdetermination 88
3.23 Concluding Remarks 91

Chapter 4: Motivating and Defending the Subset View

4.1 Mental Causation and Weak Overdetermination 93
4.2 Reduction of Powers without Reduction of Properties 93
4.3 In What Way are Mental and Physical Properties Distinct? 95
4.4 The Subset View and the Exclusion Problem 97
4.5 Shoemaker's Subset View 98
4.6 Shoemaker on the Exclusion Problem 100
4.7 The Proportionality Claim and the Closure Principle 104
4.8 The Proportionality Claim and the Subset View 104
4.9 Reconciling the Proportionality Claim with the Subset View 105
4.10 Yablo and the Subset View 107
4.11 Is the Subset View Committed to the Proportionality of Causes? 110
4.12 The Proportionality Constraints and the Exclusion Problem 112
4.13 The Argument from Parts and Wholes 112
4.14 The Subset View and Physicalism 114
4.15 The Problem of Backward Looking Powers 117
4.16 The Subset View and Functionalism about Mental Properties 120
4.17 Must a Proponent of the Subset View Accept Shoemaker's Functionalism? 123
4.18 Motivating the Subset View Independently of Functionalism 124
4.19 Kim on the Subset View 126
4.20 Motivating Reductionism over Non-Reductionism 128
4.21 Limits of the Subset View 130

Chapter 5: The Subset View and Overdetermination of Causal Powers

5.1 Introduction 133
5.2 Determinable and Determinate Properties 134
5.3 Gillett and Rives' Objection 135
5.4 A New Exclusion Problem 137
5.5 Rejecting NOP 138*
5.6 Varieties of Power – Simple, Compound, and Joint 139
5.7 Grounding Joint Powers in Conjunctive Properties 142
5.8 A Potential Pessimist Rejoinder 144
5.9 Grounding Joint Powers in Properties Bestowing Conditional Powers 145

5.10 Further Potential Pessimist Rejoinders 147
5.11 Concluding Remarks on the Grounding of Joint Powers 149
5.12 Weakening NOP 149*
5.13 Concluding Remarks 152

Chapter 6: Alternative Non-Reductive Solutions to the Exclusion Problem

6.1 Introduction 153
6.2 Amending the Closure Principle – Causal Sufficiency versus Causal Efficacy 154
6.3 The Difference-making Account of Causation and the Exclusion Problem 156
6.4 Replacing Closure with Closure 159*
6.5 The Empirical Evidence for Closure 161
6.6 Difference-making and the Subset View 163
6.7 Against Menzies' View 165
6.8 Concluding Remarks on Menzies' View 166
6.9 Disambiguating 'Property' 167
6.10 The Trope Theory 167
6.11 The Trope Solution to the Exclusion Problem 168
6.12 The Trope Solution and the Subset View 170
6.13 Gibb: Trope-Identity Entails Type-Identity 172
6.14 From Exact to Inexact Resemblance between Tropes of the same Type 172
6.15 Determination Dimensions and Resemblance between Tropes .. 175
6.16 Determination Dimensions and Psychophysical Trope-Reduction 177
6.17 Concluding Remarks on the Trope Solution 178

Conclusion 179

Bibliography 181

Introduction

Natural science tells us that the world is fundamentally physical - everything is ultimately constituted by physical properties and governed by physical laws. How do we square this picture of the world with the apparent fact that there are genuine causal relations at levels that aren't described by physics? The problem of mental causation is at the heart of this issue. There are probably two reasons for this. Firstly, if there are any non-physical properties at all, surely mental properties are among them. And secondly, the reality of mental causation is arguably more important to us than the reality of any other kind of causation. Without it, it would be hard for us to make sense of ourselves as agents with free will and moral responsibility. The main purpose of this thesis is to defend a view that accepts a scientific worldview and still allows for mental properties to exist, be non-physical, and be genuine causes of actions and behaviour. Some philosophers are pessimistic that all these goals can be achieved. They think that the only way for mental properties to fit into the causal structure of the world is if these mental properties are really physical properties. I do not find the argument for this view compelling. As I will show, it relies on an implausibly strong constraint on causes that must be amended. Once amended, a new position emerges, the so-called Subset view, which is actually motivated by the very premises that initially pushed us towards a reductive view of mental properties. Below is a brief summary of the main points of the thesis.

In chapter 1 I discuss how best to understand the doctrine of physicalism – roughly the view that physics provides the fundamental inventory list of the universe. There are two central issues associated with fleshing out this metaphor. First, we need a non-trivial account of what the view amounts to. And second, we need an account that allows for the existence of non-physical properties without giving up the claim that the physical has priority in some sense. As for the former, I argue that physicalism makes a substantial and non-trivial claim in at least two ways – one, it denies the existence of fundamental properties that are such that they are incapable of being incorporated into natural scientific theory, and two, it denies the existence of fundamental mental properties. As for the latter, I argue that the relationship between the physical and the non-physical is one of metaphysical grounding. This

account is in contrast to accounts that attempt to specify the relationship in modal terms using the notion of supervenience. This is not to deny the importance of supervenience when defining physicalism. While the truth of any such supervenience-based account is insufficient for the truth of physicalism, it *is* necessary. The reason for this is that grounding has modal consequences – it entails supervenience. *Which* kind of supervenience is determined by what goes in the grounding-base for a given set of entities. If laws of nature are included, then the supervenience entailed is nomic supervenience. I discuss which kind of supervenience is entailed in the case of the non-physical being grounded in the physical and argue that, contrary to what is often expressed in the literature, the right answer is that it should be no stronger than a form of nomic supervenience.

In chapter 2 I discuss a potential problem that arises from this conclusion. Emergentism is the view that, while the world is fundamentally physical, there are some non-physical properties, so-called emergent properties, which are causally independent, in a strong sense, from physical properties. Historically, emergentism has been considered an anti-physicalist position. However, it is not obvious why we should think that is the case given the account of physicalism I endorse in chapter 1, since the emergentist can accept that everything non-physical is grounded in physical properties and the laws of nature. The solution is to further restrict what goes in the grounding-base for the physical. The physicalist should not allow *all* laws of nature in that grounding base, only physical laws. This raises the problem of distinguishing physical laws from non-physical laws. In the literature, the main focus tends to be on the non-causal laws that connect physical properties to emergent properties (what the emergentists called trans-ordinal laws). However, I argue that it is no straightforward matter to give a criterion for physical laws that exclude trans-ordinal laws from being physical. Instead, the problematic laws from a physicalist perspective are rather the causal laws in virtue of which emergent properties gain their causal independence from physical properties. These laws should not be considered physical and so, should be excluded from the physicalist grounding-base for non-physical properties. When this is done, physicalists and emergentists can be seen to disagree on what the world is like. Emergentists claim that there causal powers at the level of emergent properties which are not grounded in the physical. Physicalists, on the other hand, deny this.

I end chapter 2 by considering how to draw the distinction between reductive and non-reductive physicalism. Reductive physicalists claim the non-physical properties, in particular mental properties, are identical to physical properties. The non-reductive physicalist denies this. However, given that I stipulate physics to be the science of fundamental properties, it seems that reductive physicalism becomes the paradoxical view that physicalism is true and that mental properties are reducible to fundamental properties. But this cannot be right since the existence of fundamental mental properties would make physicalism false. I argue that what is needed is some less restrictive notion of physical properties. These properties can function as a reduction-base for mental and other non-physical properties on a reductive physicalist account. Such physical properties need not be fundamental but are nevertheless broadly physical in the sense that they can be generated directly from fundamental physical properties. The relations suitable to generate them are conjunction and disjunction, and in particular, mereological fusion. It is the latter that allows the reductive physicalist to locate mental properties at mereologically complex levels, in particular at the level of neural states. The physical properties of such states are composite properties, ultimately decomposable into fundamental physical properties. The reductive physicalist's claim is thus that every non-physical property is identical to a conjunction, disjunction, or mereological fusion of fundamental physical properties. Correspondingly, the non-reductive physicalist denies this.

In chapter 3 I begin settling some terminological issues and sketching the metaphysical framework assumed by the thesis. Roughly, the role of properties is to bestow causal powers on the particulars that instantiate them. When these powers manifest, the property that bestows them is causally efficacious, or simply a cause. This leads to a presentation of the Exclusion problem for mental properties. The Exclusion problem is a paradox consisting of four individually motivated, yet jointly incompatible statements. First, the claim that there is mental causation. Few facts seem as obvious as this one. Surely, what we do is frequently a causal result of what we want, believe, hope, fear, etc. Second, the claim the world is causally closed at the physical level. What this means is that every physical effect has a sufficient physical cause. This is backed up by the success of science, which

indicates that there are complete physical causal explanations of all physical effects. Third, the claim that, as a rule, no effect has more than a single sufficient cause. In other words, if a property is a sufficient cause of an effect, then all other properties are *excluded* from being causes of this effect. This is known as the Exclusion principle. The Exclusion principle is typically motivated by principles of ontological simplicity. And lastly, the claim that mental and physical properties are distinct, which is established by the argument from multiple realization. Any three of these claims can be used to establish the denial of the fourth. In particular, if we accept that there is mental causation, that every physical effect has a sufficient physical cause, and that these effects have no other causes, then it follows that mental and physical properties cannot be distinct. This is known as the Exclusion argument for reductive physicalism.

The remainder of chapter 3 is largely devoted to a discussion of the so-called Drainage problem. If the Exclusion argument is sound, it seems that causal powers drain from higher to lower levels of mereological complexity – for instance, the effects caused by some higher-level broadly physical composite properties are also caused by the sum of the lower-level fundamental physical properties of which they are composed. This line of thought has devastating consequences if we accept that it is a contingent matter whether there are any fundamental properties at all. Perhaps the world decomposes endlessly with each level being more fundamental relative to the level above it. If so, causal powers would drain away into a bottomless hole and there would be no causation anywhere. I discuss various ways in which the proponent of the Exclusion argument can respond to this problem and ultimately argue that the only solution is to allow that higher-level composite properties have the very same causal powers as the sum of lower-level properties that compose them. This leads to an amendment of the claim that, as a rule, effects do not have more than a single sufficient cause. In many cases they do and what is common to these cases is that the various sufficient causes of a certain effect are causes in virtue of manifesting the very same causal power. I call this weak overdetermination.

In chapter 4 I consider how this conclusion can be applied to mental properties. If the causal powers of any mental property are also powers had by some physical

property, then they could both be allowed as sufficient causes of physical behaviour, even if these properties are distinct. This potential solution to the Exclusion problem is endorsed by the so-called Subset view for mental properties – the view that the causal powers of any mental property forms a proper non-empty subset of the powers of some physical property. I present Sydney Shoemaker's version of the Subset view, consider to what extent it is a physicalist view, and argue that it solves the Exclusion problem (though not exactly in the way Shoemaker himself suggests). While Shoemaker endorses certain controversial claims regarding the nature of causation and mental properties, I argue that these are not necessary commitments of the proponent of the Subset view. Rather, once the Exclusion principle is amended to allow for weak overdetermination, the Subset view is motivated by the very statements that supported reductionism about mental properties. At the same time, this removes the main motivation for popular reductive views. I end the chapter by considering some potential limits of the Subset view.

In chapter 5 I defend the Subset view against an objection according to which the Subset view solves the original Exclusion problem only on pain of creating another. The objection is the following: Even if we allow that *effects* can be overdetermined in the way suggested by the Subset view, we should not allow that the causal *powers* of objects are overdetermined by distinct properties as this would be ontologically unparsimonious. Since it is a central claim of the Subset view that the mental powers of an individual are determined by both mental and physical properties, the Subset view should be rejected as a consequence. I argue that this objection is based on new version of the Exclusion principle which we have little reason to accept. The reason is that some causal powers cannot help but be overdetermined by distinct properties. Hence, it cannot be a problem that the Subset view incorporates such overdetermination of powers. I end by expressing some general worries about the value of arguments from ontological parsimony.

In chapter 6 I argue that two current alternative non-reductive solutions to the Exclusion are either inferior to the Subset view or unlikely to work at all. The first is a solution according to which we should give up the claim that every physical effect has sufficient physical cause. Some effects are exclusively *caused* by mental

properties, though for any such effect there is a physical property which is causally *sufficient* for it. I discuss how the empirical evidence is arguably insufficient to rule out this possibility, though I ultimately reject it on the grounds that it, unlike the Subset view, leaves psychophysical supervenience a mystery. The second proposed solution is based on a trope theory of properties and claims that while reductionism is true for mental tropes, it is false for mental types. This makes it unmysterious how mental tropes can be causes in a world that is causally closed at the physical level, while still maintaining that mental and physical types are distinct. I discuss whether a trope can be of several distinct types and answer that it can only be so to the extent that the types in questions are individuated along the same dimensions (the so-called determination dimensions for properties). Unfortunately, we have good reasons for thinking that mental and physical types are individuated along *different* dimensions. This means that a trope cannot be of both a mental and a physical type, and this blocks the kind of psychophysical trope-reduction endorsed by the trope solution. Hence, the Exclusion problem is left unsolved.

In the end, I conclude that the Subset view presents an attractive theory of the relationship between mental and physical properties and an attractive solution to the problem of mental causation.

Chapter 1

Defining Physicalism

Abstract: In this chapter I discuss how best to define physicalism. I argue that a definition in terms of grounding is superior to a definition in terms of supervenience, but that some form of supervenience is entailed by the grounding relation. In the end, I argue that, for purposes of defining physicalism, the supervenience entailed is a form of nomic supervenience.

1.1 Introduction

Physicalism is often said to be view that there is nothing over and above the physical with one of the debates over physicalism being over how to flesh out this slogan.[1] Typically, this is framed in terms of properties. Hence, the physicalist slogan:

(Slogan): There are no properties over and above physical properties.

Considering the physicalist slogan, two questions are particularly salient. First, what is it for a property to be physical? And second, what is it for a property to be nothing over and above physical properties? Starting with the former question, a plausible answer is that physics, the natural science, tells us which properties are physical.

(Physical): A property is physical iff it is a property posited by physics.[2]

Coupled with the physicalist slogan this gives us a starting point for a definition of physicalism:

(Physicalism1): There are no properties over and above the properties posited by physics.

[1] See for instance Dowell [2006], p. 26, Chalmers [1996], p. 41, Howell [2009], p. 83, Jackson [1998], p. 9, Wilson [2005], p. 426.
[2] Roughly, a property is a posit of physics if the predicate that refers to it is a part of physical theory.

Physicalism about other entities such as objects or events I take to be derivative of this definition. For instance, physicalism about objects is the view that every object in the actual world has only physical properties or properties that are nothing over and above physical properties, and physicalism about events is the view that every event in the actual world has only physical properties or properties that are nothing over and above physical properties.

1.2 Physics

Roughly, and without getting too bogged down in issues from the philosophy of science, physics we may take to be defined by its methodology (which includes empirical experiments, inductive reasoning, reproducibility, and so on) and its subject matter, which are the fundamental entities in the world. The former it shares with other natural sciences, the latter is unique to physics. Thus, physics here is to be understood as fundamental physics. Correspondingly, the properties posited by physics might be called fundamental physical properties. For now, however, I will simply refer to them as physical properties. Later on it will be necessary to make more fine-grained distinctions. Focusing on properties, physics is thus the science that is in the business of uncovering the fundamental properties in the world and their interactions, using a certain method distinctive of natural science. Thus, other special scientific disciplines use the same method, but treat only non-fundamental properties. Properties that are incorporable into scientific theories I will call *scientific properties*. Properties that are not I will call *unscientific properties*. As for fundamentality, we may take that to mean something along the lines of: not dependent for its existence on any other property (or properties), where the relevant form of dependence is metaphysical rather than causal. I will say more about this later.

One thing to note is that on this understanding of physics, many scientific disciplines typically thought of as being part of physics, will count as special scientific. For instance, assuming that atoms and their properties are non-fundamental, atomic physics will count as a special science. This, however, poses no problem for my account. For the purpose of defining physicalism, the important point is that we have a distinction between science that treats fundamental properties and those that do not. Whether one prefers to cash that out in terms of

physics versus other natural sciences, or in terms of fundamental physics versus other natural sciences, some of which will be disciplines of non-fundamental physics, is a terminological dispute. I prefer the former since it avoids the question of what delineates non-fundamental physics from other special sciences (on my account there simply is no such thing as non-fundamental physics), but my account can easily be translated into an account that opts for the latter. Simply replace talk of physics and its posits, with talk of fundamental physics and its posits.

1.3 Being Physical and Hempel's Dilemma

The definition of physical properties above gives rise to Hempel's Dilemma. The first horn of the dilemma is this: If physicalists base their physicalism on the properties postulated by current physics, physicalism is almost certainly false, as current physics and the list of properties it posits is bound to be revised in the light of new scientific discoveries. To most, this is taken as a decisive blow against basing physicalism on the posits of current physics.[3] The obvious alternative is to look ahead and base physicalism on the properties postulated by future completed physics. Such a list would not be open to revision (given that the physics providing it is complete). Hence, we might define physical properties as those being posited by completed physics:

(Physical2): A property is physical iff it is a property posited by completed physics.

This definition, however, runs afoul of the second horn of the dilemma. As Alyssa Ney puts it, if physicalism is based on the properties postulated by completed physics, "the view is trivial since, in order to become complete, physics will of course expand as much as it has to in order to cover all phenomena, even if this means in the end positing irreducible mental entities".[4] This, however, cannot be quite right. If the subject matter of physics is the domain of fundamental properties, and physics goes about discovering these properties using a certain scientific method, there are two kinds of properties left out of the story told by completed

[3] Andrew Melnyk is one of the few that has defended such a strategy. See Melnyk [1997] and Wilson [2006] for some problems with this defense.
[4] Ney [2008], p. 2.

physics. First, completed physics will not expand beyond the domain of fundamental properties, and second, completed physics will be silent about fundamental properties that are undetectable using the methods of science or detectable but incapable of being incorporated into scientific theory.[5] Hence, it is wrong to suggest, as Ney does, that complete physics must cover *all* phenomena. This response, however, does not yet effectively deal with the second horn of Hempel's Dilemma. As Ney also points out, completed physics may end up positing mental properties, not because it must do so to cover all phenomena, but because mental properties may turn out to be fundamental. Now, if physicalism is unconditionally based on the properties postulated by completed physics this means that physicalism is compatible with mentality being fundamental. This suggests a different way in which physicalism based on completed physics is threatened by triviality - physicalism would be trivialized in the sense that it fails to contrast with paradigm examples of anti-physicalist positions. For instance, take Cartesian or strong property dualism to be the view that the domain of fundamental properties is bifurcated between mental and non-mental properties that causally interact.[6] It certainly seems possible (though at present unlikely) that completed physics will end up vindicating such a view. And if so, physicalism as defined above and based on the properties postulated by completed physics would be compatible with Cartesian dualism. Of course, accepting this does not make physicalism wholly trivial. Physicalism still contrasts with positions that involve fundamental properties which, for some reason, cannot be incorporated into physical theory as such properties naturally wouldn't be part of the posits of completed physics. It just fails to contrast with all the positions we thought it did.

Some philosophers have accepted that, in principle at least, physicalism is compatible with such strong forms of dualism. J. L. Dowell, for instance, bases her physicalism on the posits of future completed physics and requires only that the fundamental properties that this physics posit must be capable of being incorporated into theories in a way which respects, as she calls it, "the hallmarks of scientific

[5] Detectability can here be both direct and indirect. The most obvious example of an undetectable property would a property with no causal powers.
[6] Put aside the fact that Descartes himself was a substance dualist, not a property dualist.

theories".[7] In other words, a fundamental property, mental or otherwise, which fails to be incorporable into scientific theories, will not count as physical on Dowell's account. Her example of such a property is one which bestows miraculous powers, for instance the property *being an angel*. No further constraints are needed on what is to count as physical, according to Dowell. In particular, "if no actual mental property is among the basic physical ones, as it seems overwhelmingly likely, that's a matter to be settled *a posteriori*".[8]

Others disagree. Jessica Wilson, for instance, makes the point that:

> Given that physicalism is an anti-dualist doctrine, then while […] the physics-based boundaries of the physical may stretch, they cannot stretch so far as to encompass fundamental mentality. Hence physicalists (and their rivals) have good reason to impose the NFM (no fundamental mentality) constraint on their operative account of the physical.[9]

Wilson has two reasons for imposing this additional constraint. First, a version of physicalism imposing a *no fundamental mentality* constraint (henceforth a NFM constraint) on the physical guarantees, contra a form of physicalism based on Dowell's account of the physical, that physicalism will be distinct from its traditional rivals, in particular certain forms of dualism or idealism. This is important in so far as, historically, physicalism, and its ancestor materialism, have been partly defined as being in opposition to dualism and idealism. Secondly, the philosophical interest in defining the physical is in large part driven by a desire to formulate and motivate the mind-body problem, understood as the problem of seeing how mentality can arise or have its place in a world that is fundamentally non-mental. But if mentality is, in principle, allowed in the domain of fundamental physical properties, then a formulation of the mind-body problem in terms of the physical versus the mental is no good and would dissolve the mind-body problem, rather than motivate it. Hence, an account of the physical which guarantees that the physical is non-mental is preferable.

[7] Dowell [2006], p. 26.
[8] Dowell [2006], p. 28.
[9] Wilson [2006], p. 70.

Of these two reasons, the former seems to me to carry the greater weight. After all, it might just be that the mind-body problem should *not* be framed in terms of the physical versus the mental. Furthermore, should it turn out that there are fundamental mental properties, the mind-body problem as it has been traditionally understood would presumably be rejected as illegitimate seeing that, contrary to what is at present assumed, the mind does *not* arise from something fundamentally non-mental. In other words, assuming that an account of the physical *must* motivate the mind-body problem is begging the question against Dowell's account. What seems unlikely however, is that physicalists in general would accept that their position is compatible with dualism, or would continue to label themselves as physicalists should dualism be vindicated by future physics.[10] This suggests that Wilson is ultimately right in thinking that a NFM constraint is needed in some form or another. Now, this does not yet decisively settle the issue in favour of Wilson, as one can accept Dowell's account of the physical and impose the NFM constraint on physicalism itself, rather than the definition of physical properties. In this case, physicalism would be the view that there are no properties over and above the properties posited by completed physics, and that no mental property is a physical property. And following Wilson's reasons for having the NFM constraint in the first place, this seems to be the more natural option. After all, the claim is that it is a defining feature of *physicalism* and its history that it is opposed to fundamental mentality, not a defining feature of the posits of physics. This also avoids the peculiar result that physics could end up positing properties that are not physical. I therefore endorse an account of physicalism which, following Wilson, imposes a NFM constrain, however, I suggest that it is best imposed on the definition of physicalism itself rather than the definition of the physical. Our starting point for a definition of physicalism therefore needs to be slightly amended:

(Physicalism2):

> There are no properties over and above the properties posited by completed physics, and no mental property is a physical property.[11]

[10] Wilson mentions some of these in Wilson [2006], p. 70. See also Joseph Levine [2001], p. 70 and David Papineau [2001], p. 12.

[11] A less discussed version of Hempel's Dilemma would be one that targets the methodology of physics. The first horn of the dilemma is this: The methodology of current science may turn out to be ineffective or insufficient, and will eventually be replaced by different methods that allow us to gain

1.4 Do Fundamental Chemical or Biological Properties Falsify Physicalism?

One might worry that other posits of completed physics would falsify physicalism in the same way the discovery of fundamental mental properties would. For instance, would it not be equally problematic for physicalism if completed physics posits fundamental chemical or biological properties? And so, should we not then amend the definition of physicalism to exclude fundamental chemical or biological properties as well? Wilson argues that since, in fact, chemical and biological properties are ontologically dependent on physical properties this question can be put aside.[12] But this is too quick since a definition of physicalism should be independent of how the actual world is. Even if we assume that it is beyond dispute that chemical and biological properties are non-fundamental, there is still a question of whether or not this must necessarily be so for physicalism to be true. Wilson may think that what it would take for physicalism to be true is, among other things, for chemical and biological properties to be non-fundamental, but if so she should endorse a 'no fundamental chemical or biological properties' clause in addition to NFM. The worry is that this makes physicalism an unsystematic or *ad hoc* position.

My own view is that there is no need to impose a clause that rules out fundamental chemical or biological properties. There are three reasons for this. Before giving them, however, it is useful to first introduce the so-called 'levels view' according to which reality can be stratified following the axioms of mereology. At the bottom we have the level of mereological atoms, what I will call micro-objects, i.e. the smallest objects in the world. As we move up through the levels, the compositional complexity increases. We find elementary particles, then atoms, then molecules, then cells, then tables and chairs, and so on. With this picture in mind we can

---

complete and accurate knowledge of the world and its properties. Perhaps these methods of future science will even be radically different from the methods of current science. So if physicalism is committed to the claim that the fundamental properties of the world can be exhaustively described using the methods of *current* science, physicalism is almost certainly wrong. This, however, leads to the second horn. If the physicalist gives up the claim that the world is fundamentally such that it can be completely and accurately described using the methods of current science, all she is left with is the negative claim that there are no mental properties among the fundamental properties. While some philosophers embrace such a purely negative definition of physical properties, at least for some purposes (see for instance Spurrett & Papineau [1999]), I do not find out of the question to simply deny the first horn of the dilemma. After all, current scientific methodology has proven much more stable than current scientific theories. Science textbooks can become outdated in a few years, while most of the central features of current scientific methodology have been in place for centuries.

[12] Wilson [2006], p. 75.

distinguish physicalism *simpliciter* from a specific form of physicalism, micro-physicalism. Micro-physicalism is physicalism with the added claim that the fundamental properties posited by completed physics will all be properties of micro-objects. Micro-physicalism is a popular form of physicalism, so much so that it is sometimes taken for granted as being physicalism *simpliciter*. As an example, consider Michael Tye's description of physicalism:

> Physicalism, in its most general form, is the thesis that no nonphysical ingredients are needed to account for anything in the actual world: the physical ingredients alone suffice […]. Indeed, the microphysical ingredients alone suffice. Once the microphysical facts about the world are fixed, everything else in the world is automatically determined.[13]

But chemical and biological properties, whether fundamental or not, are not properties of the smallest objects in the world so should they turn out to be fundamental, micro-physicalism would be false, as the micro-physical alone would not suffice to account for everything in the world. The first reason why physicalism need not exclude fundamental chemical or biological properties then is this: If some indeed have the intuition that fundamental chemical or biological properties would falsify physicalism, this can be partly accommodated by the fact that such properties would falsify micro-physicalism. And this, I suggest, is enough to explain the intuition (if indeed it is an intuition) that physicalism would be falsified by the discovery of fundamental chemical or biological properties. Second reason: The historical considerations which tell against the compatibility of fundamental mentality and physicalism are not present to the same extent in the case of chemical and biological properties, as few have ever wanted to posit the existence of such fundamental properties.[14] Hence, the historical development of physicalism has not been one in which physicalism has defined itself in contrast with such views, at least not to the extent that physicalism has defined itself in contrast to views that accept fundamental mentality. Third reason: it seems possible that the actual world is *gunky*, i.e. is such that any part of it has proper parts. Framing things in terms of levels, this is equivalent to there being no lowest or bottom level. But if any part of

[13] Tye [2009], p. 25

[14] It could be argued that the British Emergentists posited fundamental chemical and biological properties, and that physicalism is historically as opposed to emergentism as it is to mind-body dualism. I think it is questionable whether the emergentists did in fact posit fundamental chemical and biological properties. I will return to this question in the next chapter.

the world itself has proper parts, then there are no micro-objects and hence no objects to have micro-properties. And this means micro-physicalism would be false in a gunky world.[15] The truth of micro-physicalism then depends on the world being non-gunky. But does the truth of physicalism also depend on the world being non-gunky? This seems like the kind of question physicalism should be neutral about. If there are fundamental properties at *some* levels, if these properties are scientific, and if none are them are mental, that should be good enough for physicalism.[16] [17]

1.5 Physicalism and Being 'Nothing over and Above'

So far, our definition of physicalism tells us that physicalism is true if and only if all properties are nothing over and above the properties posited by completed physics, and no mental property is a physical property. What is it for a property (or a set of properties) to be nothing over and above another property (or set of properties)? Our definition of physical properties restricts the possible answers to this question. In particular, it rules out the most obvious interpretation, namely that all properties might be nothing over and above physical properties in the sense that all properties are *identical* to physical properties. If this is our interpretation of the 'nothing over and above' locution, physicalism would be a radically eliminativist position, in so far as it is unlikely that most properties around us, such as *being a table*, *being a person*, *being a mountain*, and *being a philosophy department*, will ever turn out to be among the fundamental properties posited by physics. And while some physicalists, perhaps, would be happy to embrace the conclusion that such properties do not exist, it would obviously be wrong to expect all physicalists to do

[15] It might be objected that micro-physicalism need not be committed to the claim that fundamental properties are properties of micro-objects. Instead, micro-physicalism may only entail the weaker view that fundamental properties are properties of objects below a certain level of compositional complexity, for instance below the level of atoms. In this case, micro-physicalism can be true in a gunky world. This may be Wilson's view since she endorses the claim that "the physical entities are those existing at relatively low orders of complexity" (Wilson [2006], fn. 1), though she may just take that as a contingent fact about physical entities. In any case, even if micro-physicalism is defined in this way, it gives us no compelling reason to think that micro-physicalism is identical to physicalism *simpliciter*. In fact, it seems ad hoc to take it as a conceptual truth about physics that it treats only properties below a, more or less arbitrary, level.

[16] The related possibility of there being no fundamental properties, and not just no bottom level, is more troublesome. It is helpful, however, to postpone a treatment of this issue until after the notion of 'nothing over and above'-ness has been fleshed out.

[17] For some further reasons why physicalism need not be micro-physicalism see Hüttemann [2004] and Papineau [2008].

so.[18] As an alternative to identity, physicalists have looked to supervenience as a notion that will allow them to cash out 'nothing over and above' in a non-eliminativist way. Supervenient properties, the claim is, need not be identical to their subvenient base-properties, yet in some important sense, they are nothing over and above those. If so, we can flesh out the physicalist slogan in the following way:

(Physicalism3):

> Every property is either a property posited by completed physics or a property supervenient on physical properties, and no mental property is a physical property.

How is supervenience defined? The standard definition of supervenience goes like this:

(Standard Supervenience):

> A standardly supervenes on B if and only if, for any two objects *x* and *y*, if *x* and *y* are indiscernible with respect to B, they are indiscernible with respect to A.[19]

A and B are usually taken to be sets of properties, though this is not an essential assumption. Supervenience is a relation that can hold between other kinds of entities as well. Which objects are quantified over? If the domain contains objects in all possible worlds, the definition gives us *strong supervenience*. And if the objects in the domain are entire possible worlds, the definition gives us *global supervenience*.[20] There is an abundance of philosophical literature on how to define supervenience and on whether or not supervenience of the non-physical on the physical (henceforth NP-P supervenience) can adequately support a physicalist

[18] To my mind, the notion of being 'nothing over and above' comes with such strong reductionist connotations, which many physicalists would reject, that the phrase is not particularly well chosen to describe the relationship between the physical and the non-physical. This is mainly a linguistic squabble though, and well chosen or not, the notion seems fairly ingrained in the literature by now.
[19] Indiscernibility is here with respect to all the features of A and B. In particular, if A and B are properties, the indiscernibility should be with respect to both instantiations and powers.
[20] If the domain contains objects in the same world, the definition gives us weak supervenience. It is generally agreed, however, that weak supervenience is too weak to be of use for the physicalist. For instance, a dualist can agree that any two objects in our world with the same physical properties will have the same mental properties.

metaphysics.[21] First and foremost, there is a worry that NP-P supervenience is not sufficient for physicalism. This is due to positions that are typically thought of as anti-physicalist positions, but which could nevertheless endorse NP-P supervenience. For instance, Moorean non-naturalism arguably includes the claim that moral properties supervene on natural properties, and British Emergentism arguably includes the claim that mental and chemical properties supervene on physical properties.[22] But Moorean non-naturalism and British Emergentism are typically seen as examples of positions that are in opposition to physicalism. The underlying general point here is to the effect that supervenience might just express necessary co-variation between properties, and not track the kind of substantial metaphysical dependence of the non-physical on the physical that the physicalist is looking to endorse. As Jaegwon Kim says:

> Supervenience itself is not an explanatory relation. It is not a "deep" metaphysical relation; rather, it is a surface relation that reports a pattern of property covariation, suggesting the presence of an interesting dependence relation which might explain it.[23]

Perhaps the supervenience relation can be modified in such a way that it, at least for most practical intents and purposes, will track metaphysical dependence. However, even if this is so, there is something backwards about approaching the issue in this manner. After all, supervenience is appropriate for formulating physicalism, only in so far as the relation is able to track an underlying metaphysical relation which ensures that the non-physical depends on the physical in the right way. As Kim says, supervenience is at best a *symptom* of a deeper metaphysical relation which explains why the supervenience holds. So as a starting point, it would be more natural to define physicalism directly in terms of such a metaphysical relation.

1.6 From Supervenience Physicalism to Grounding Physicalism

Some philosophers have done just that. For instance, Melnyk advocates what he calls *realization physicalism*.[24] His claim is that everything that exists is either

[21] See for instance Bailey [1998], Bennett [2004], Horgan [1993], Jack [1994], Moser & Trout [1995], and Wilson [2005].
[22] For a defense of this, see Horgan [1993].
[23] Kim [2003a], p. 167.
[24] See Melnyk [2003].

physical or realized by the physical and that the truth of this claim is at least sufficient for physicalism. Here the realization relation is supposed to be a relation of metaphysical dependence. Another example is Douglas Ehring's *part-whole physicalism*. Here the view is that mental properties are exhaustively *composed* of physical properties.[25] Ehring achieves this in a framework where property-instances are abstract particulars, or tropes, and property-types are sets or classes of tropes. According to Ehring, classes of tropes making out mental types will be exhaustively composed of (or have as subsets) sets of tropes making out physical types. In other words, mental properties are literally composed of physical properties and hence, the relation between physical and mental properties is a part-whole relation.[26]

But while we may agree that realization and composition are examples of the kind of metaphysical dependence relation the physicalist is after, when trying to express the relation between the physical and non-physical, it is problematic to define physicalism in terms of any of those specific relations. Where definitions in terms of supervenience have the problem that their truth is not sufficient for the truth of physicalism, definitions in terms of realization or composition have the problem that their truth is not necessary for the truth of physicalism. You can reject that the physical realize or compose the non-physical and still be a physicalist. Whereas a supervenience-based definition of physicalism has the problem that it is not specific enough and would allow positions that, intuitively, are non-physicalist to count as physicalist positions, introducing notions of realization or composition to describe the relationship between the physical and the non-physical makes the corresponding definition of physicalism too specific to be of use as a *general* definition of physicalism that all and only physicalists should accept. Is there a more general notion of dependence that can be of help to the physicalist? I think there is and it is the notion of *grounding* as it is used by Jonathan Schaffer. According to Schaffer, grounding is a primitive relation of metaphysical dependence, which can be used to define the notion of fundamentality and non-fundamentality:

(Fundamentality): A property is fundamental iff it is ungrounded.

[25] See Ehring [2003]. Ehring focuses his discussion on mental properties though presumably his account can be extended to apply to other non-physical properties as well.
[26] I will discuss a version of this view in more detail in chap. 6.

(Non-fundamentality): A property is non-fundamental iff it is grounded.[27]

Being primitive, the grounding relation itself cannot be analysed in terms of any deeper metaphysical relations. Nevertheless, according to Schaffer, there are "clear examples" of it. For instance, it is the relation that holds between: "(i) the entity and its singleton, (ii) the Swiss cheese and its holes, (iii) natural features and moral features, (iv) sparse properties and abundant properties, and (v) truthmakers and truths."[28] And intuitively, it is also the relation that holds between the physical and the non-physical if physicalism is true. An advantage of such a view is that it directly captures the spirit of physicalism that supervenience was introduced to indirectly capture, namely that the physical is fundamental and metaphysically primary, and that the non-physical is non-fundamental and metaphysically dependent on the physical. In other words, the grounding relation can replace the notion of being nothing over and above used in the definition of physicalism above, which gives us the following definition of what we can call *grounding physicalism*:

(Physicalism4): Every property is either a property posited by completed physics or a property grounded in those properties, and no fundamental property is a mental property.

Or, given the definition of grounding, alternatively (and simpler):

(Physicalism5): Completed physics provide a list of all fundamental properties and no fundamental property is a mental property.[29]

[27] Schaffer [2009], p. 373. Note that the terminology differs slightly from Schaffer's in that his definitions of fundamentality and non-fundamentality are not limited to properties but covers all entities.
[28] Schaffer [2009], p. 375.
[29] What if physicalism, thus defined, is true and there are unscientific but *non-fundamental* properties? Would this not contradict the spirit of physicalism, and does this not show the definition to be inadequate? Two responses: First, it might be argued that unscientific non-fundamental properties would not strictly speaking falsify physicalism, though they would falsify something like naturalism (which is plausibly the view that all properties are scientific). Second, and this is the response I am inclined to go for, we may assume that properties grounded in scientific properties are themselves necessarily scientific. Then the scenario described is impossible.

Given that physics is the science of fundamental properties and fundamentality is defined in terms of grounding, (Physicalism5) is equivalent to (Physicalism4). From this, we can see that physicalism can fail for two reasons. First, physicalism fails if there are fundamental mental properties. And second, physicalism fails if there are fundamental properties not posited by completed physics (whether these are mental or not). This would be the case if there are fundamental unscientific properties, i.e. fundamental properties which are undetectable using scientific methodology, or detectable but otherwise not incorporable into scientific theory. For now, I will call such properties whose existence falsifies physicalism (fundamental mental properties and unscientific properties), *anti-physicalist properties.*[30]

While I believe a grounding definition of physicalism captures the physicalist's commitments better than a supervenience-based definition, I also think it must be admitted that no definition will be completely univocal and rule out all grey areas. 'Physicalism' is a term of art with vague boundaries, so for some possibilities, we should accept there is no universally agreed upon fact of the matter as to whether or not they are acceptable from a physicalist perspective. For instance, say that completed physics reveals that properties of elementary particles are indeed the fundamental properties of the world. Would physicalism then be falsified if some properties of atoms turned out to be mental? According to the definition of grounding physicalism above, it would not (assuming atomic properties are grounded in the properties of subatomic particles). And perhaps some physicalists would happily accept this if it could be shown how the elementary particles which compose atoms ground, say, intentionality. Then again, others might think that the idea of mentality at the atomic level is an unacceptably anti-physicalist idea. Note that while the concepts employed in our definition of physicalism may themselves be vague, this does not account for the disagreement over atomic mentality. For

[30] A distinction can be made between what we can call restricted and unrestricted physicalism. Unrestricted physicalism makes a claim about all properties, both physical and non-physical. It says that there are no mental properties among the physical properties and that all non-physical properties are scientific. This is the kind of physicalism I have been concerned with so far in this chapter. However, physicalism might hold for a restricted class of properties. For instance, one may be a physicalist about chemical properties, while still holding that unrestricted physicalism is false, perhaps because one believes there are fundamental mental properties. Restricted physicalism, I believe, is best interpreted as the view that the class of non-physical properties about which one is a physicalist, is ultimately grounded in nothing but physical non-mental properties. If so, restricted physicalism about class of non-physical properties C allows that there are anti-physicalist properties, it just denies that any of them are in the grounding base for the properties in C.

instance, two philosophers can decide on a non-vague definition of 'physics' and 'fundamental', and still disagree about whether or not atomic mentality is compatible with physicalism. Of course, if our grounding definition of physicalism is compatible with atomic mentality, anyone thinking atomic mentality is ruled out by physicalism will think this definition is not sufficient for physicalism. Perhaps such a philosopher would want to add on to the definition that mentality can only be a property of objects above a certain level of mereological complexity, and that atoms fail to reach this level.[31] So be it. Adding on this further condition on physicalism won't improve our definition, as it will merely lead to other philosophers rejecting it as too strong. We simply have to accept that for certain scenarios, different philosophers will have different opinions over whether these are compatible with physicalism.

1.7 What if there are no Fundamental Properties?

Having settled the question of how best to understand the notion of a set of properties being nothing over and above another, as it is used in the definition of physicalism, we are in a better position to deal with the possibility of there being no fundamental properties. Say that a world is such that every level is dependent on a level immediately below it. It follows from this that there is no bottom level in such a world and that there is no level of fundamental properties. In such a world, there would not only be no completed physics (since every level is grounded in a relatively more fundamental level, the project of physics would be never-ending), physics would simply have no subject matter. (Physicalism4) would then be false (if there are no fundamental properties, non-fundamental properties are not grounded in them).[32] Is this unacceptable? I think not. First, recall that the notion of physics utilised in the definition of physical properties and physicalism is that of fundamental physics. It's a trivial point that there is no such thing as fundamental physics if there are no fundamental properties, so if the physical and physicalism is best defined in terms of fundamental physics, it should be a trivial point that there is

[31] I argued before that fundamental chemical and biological properties would not falsify physicalism. I stand by this argument though I won't put up much of a fuss if someone were to insist that this issue belongs in the grey area between physicalism and non-physicalism.

[32] Whether (Physicalism5) would be false depends on what one takes the truthvalue to be of sentences that involve reference-failure (completed physics have no reference in a world with no fundamental properties). For this reason, it is probably best to see (Physicalism5) as equivalent to (Physicalism4) only in worlds with fundamental properties.

nothing physical and that physicalism is false if there are no fundamental properties. Of course, this does not mean that dualism or idealism is true in such a world. Such views require fundamental mentality and if there are no fundamental properties there are no fundamental mental properties either. However, certain counterpart views of physicalism, dualism, idealism, and other positions that are defined in terms of fundamental properties, may hold true in a world with no fundamental properties. As a first stab at defining a counterpart version of physicalism, consider again a world with no bottom level and no fundamental properties. Assume that at a certain level, say the chemical level, all properties are non-mental and scientific, and that the same is the case *for all properties at every level below it*. In other words, when moving down levels we will eventually reach a point (in the case described, the chemical level) after which all properties are non-mental and scientific.[33] While physicalism is false in such a world (on my account of physicalism), the scenario described is still revealing as it undoubtedly captures the scientistic spirit of physicalism. After all, in the world described, all of reality would bottom out in (an infinitely long line of) purely scientific, non-mental properties. 'Naturalism' would perhaps be the best term to describe what holds true in this world, since everything is grounded in some scientific, or natural, non-mental properties, which themselves are grounded in relatively more fundamental scientific, non-mental properties, and so on, *ad infinitum*.

In a world with no fundamental properties, the counterpart to idealism, traditionally the view that reality is fundamentally mental, would plausibly be the contrary view – that when moving down levels, we will eventually reach a point after which all properties are mental. The counterpart to dualism, traditionally the view that the fundamental properties include both mental and physical properties, would plausible be the view that one part of reality bottoms out in mental properties and another part bottoms out in scientific non-mental properties. So rather than reality consisting in one hierarchy of levels, it would consist in two. In one hierarchy, when moving down levels, we will eventually reach a level after which all properties are mental. In the other, we will eventually reach a level after which all

[33] The example is due to Barbara Montero [2006]. On her account of physicalism, a world with nothing but scientific mental properties below a certain level is a world in which physicalism is true. While I disagree, the scenario she describes is still revealing.

properties are scientific and non-mental. I am not sure which terms best describe such counterpart views to idealism and dualism.

1.8 Grounding and Supervenience

As mentioned earlier, a grounding-based formulation of physicalism has an advantage over a supervenience-based formulation, in that it directly explicates the physicalist's commitments. Furthermore, putting aside the question of whether or not supervenience-based definitions of physicalism are compatible with anti-physicalist positions, Schaffer gives two additional reasons why supervenience does not track grounding. First, supervenience is reflexive and can hold symmetrically, while grounding is irreflexive and always asymmetric. For instance, if A and B are identical, A supervenes on B (and *vice versa*) following the standard definition of supervenience, yet nothing grounds itself according to Schaffer, so B does not ground A (nor *vice versa*). Secondly, necessary entities, say numbers, trivially supervene on any and all entities, yet necessary entities need not be grounded in any nor all entities.[34]

Nevertheless, Schaffer himself is open to the possibility of the grounding relation entailing supervenience. As he says:

> There is an interesting question about the modal consequences of grounding. This opens up the possibility of using supervenience for *something* – the right sort of supervenience failure can show grounding failure.[35]

[34] Schaffer [2009], p. 363. I am not entirely convinced that this once and for all closes the door on attempts to define physicalism in terms of supervenience, in a way that brings it into line with a grounding definition. First, it might be that the definition of physicalism in terms of supervenience can be modified in a way that anticipates the differences between supervenience and grounding. For instance, considering that supervenience, unlikely grounding, is formally a non-asymmetric relation, the supervenience physicalist might simply add on to his definition that the non-physical asymmetrically supervenes on the physical. Second, the reasons why supervenience does not track grounding may simply be irrelevant as far as physicalism is concerned. For instance, it might be that numbers are necessary and trivially supervene on any and all things, yet they might also be entities about which physicalism is simply not concerned, perhaps because physicalism is taken as a claim about the empirical world and its properties. Nevertheless, as I mentioned earlier, even if such moves are viable, if supervenience based definitions of physicalism are successful only in so far as they manage to track a grounding based definition, it is not obvious why we shouldn't simply go for a grounding based definition in the first place.

[35] Schaffer [2009], p. 364. Many famous arguments against various forms of physicalism, such as the zombie argument (Chalmers [1996], chap. 3) and the swampman argument (Davidson [1987]), exactly target the claim that the non-physical supervene on the physical (in the right way at least).

If it can be shown that grounding entails supervenience this would be helpful in two ways. First, it will help to label some non-physicalist positions as being such, namely those that deny NP-P supervenience. Second, it helps to determine what goes in what we may call the *complete fundamental grounding base* for a given kind of entity E. The complete fundamental grounding base for E is the complete set of fundamental entities on which E ontologically depends. This is to be distinguished from grounding bases that contain non-fundamental entities and from grounding bases that are incomplete and does not contain all entities on which ontologically E depends. If grounding entails supervenience then, if a set of entities is not a supervenience base for E, then that set is either not a grounding base for E at all or an incomplete grounding base for E.

But why should we think that grounding entails supervenience in the first place? I can think of three reasons. First, the grounding relation is an intimate metaphysical relation, perhaps the most intimate relation that can hold between distinct entities, and it seems right that such a relation should be modally strong and support supervenience. Second, when A supervenes on B, the grounding of A in B is a possible explanation of this (remember, supervenience is a symptom of some deeper metaphysical relation which explains it). And one obvious way grounding can explain why supervenience holds is by entailing that it does so. Third, the kinds of relations that are naturally seen as examples of grounding relations are typically thought to entail supervenience. For instance, composition and realization are plausibly instances of the grounding relation in that wholes/realized properties may be taken as grounded in their parts/realizers, and wholes/realized properties are typically thought to supervene on their parts/realizers. For these reasons it is plausible to assume that the grounding of A in B (henceforth A-B grounding) entails that A supervenes on B.

1.9 Grounding Physicalism and Standard Supervenience

Which kind of supervenience is entailed? To settle this question it will be helpful to turn to the debate on supervenience physicalism. Grounding physicalists and supervenience physicalists share the goal of trying to state the necessary and sufficient conditions for physicalism. Where they differ is in what they take to be sufficient for physicalism. Supervenience physicalists believe that some form of

NP-P supervenience will be sufficient for physicalism (putting aside the NFM constraint). Grounding physicalists, on the other hand, reject this, since they reject that supervenience in general is sufficient for grounding. Nevertheless, there is no reason to think that grounding physicalists and supervenience physicalists should disagree on which kind of supervenience is necessary for physicalism. This is why, when trying to determine the modal consequences of grounding, it is helpful to turn to the debate over how best to formulate supervenience physicalism.

Consider first a form of physicalism based on standard supervenience (following the definition of standard supervenience given earlier).

(Standard Physicalism): All non-physical properties standardly supervene on physical properties, and no mental property is fundamental.

I take it that we have good reasons for thinking this formulation is too strong. The problem is that physicalism based on standard supervenience rules out that there are merely possible worlds that contain anti-physicalist properties, in particular, fundamental properties not capable of being incorporated into physics.[36] Say that there is a fundamental unscientific property *P* in the actual world. A definition of physicalism based on standard supervenience would rightly judge physicalism to be false in our world (a possible world exactly like our world except that it does not contain *P* will be physically indiscernible from the actual world, but discernible with respect to non-physical properties). But what if there are no anti-physicalist properties in our world and the non-physical standardly supervene on the physical? Then physicalism is also false if there are possible worlds, physically indiscernible from our world, except for some extra fundamental unscientific property. This seems like the wrong result. Physicalism, after all, is a contingent thesis about the actual world – it does not aspire to rule out that physicalism is false in all other possible worlds. Turning to the grounding relation, the same point applies. The

[36] By merely possible, I mean possible and not actual. Note that merely possible fundamental mental properties, while anti-physicalist properties, pose no distinct problem for physicalism based on standard supervenience. If these fundamental mental properties are scientific, i.e. incorporable into completed physics, any world containing them will be discernible from the actual world with respect to physical properties, and so won't be relevant from the perspective of evaluating the truth of physicalism in our world. If they are not scientific, then they falsify standard physicalism for the same reason non-mental unscientific properties falsify it.

mere possibility of a world where the physical does not ground the non-physical, due to there being fundamental unscientific properties in that world, should intuitively not rule out that the physical grounds the non-physical in the *actual* world. What this shows is that standard NP-P supervenience is too strong as a requirement for physicalism, either in its supervenience based or grounding based formulation.

1.10 Grounding Physicalism and Minimal Supervenience

In the light of this, Frank Jackson has suggested that physicalists are committed only to the claim that any world that is a *minimal* physical duplicate of the actual world with respect to the physical, rather than physically indiscernible *simpliciter*, is a world that is indiscernible with respect to everything else. As Jackson further explains:

> […] a minimal physical duplicate of our world is a world that (a) is exactly like our world in every physical respect [...], and (b) contains nothing else in the sense of nothing more by way of kinds or particulars than it *must* to satisfy (a). Clause (b) is a 'no gratuitous additions' or 'stop' clause.[37]

We can use this notion of a minimal physical duplicate to define what I'll call *minimal supervenience*:

(Minimal Supervenience): A minimally supervenes on B if and only if, for any two objects *x* and *y*, if *y* is minimal duplicate of *x* with respect to B, *x* and *y* are indiscernible with respect to A.

While Jackson takes duplicates to be worlds, there is nothing preventing our definition of minimal supervenience from being applied to ordinary sized objects as well. For instance, the chemical properties of a body of water *x*, minimally supervenes on the micro-physical properties of the water if and only if, any object that is a minimal duplicate of *x* with respect to its micro-physical properties, is indiscernible from *x* with respect to chemical properties. According to Jackson,

[37] Jackson [1998], p. 13.

physicalism is true if and only if the non-physical minimally supervenience on the physical, though as I argued earlier, an added NFM constraint is needed. So we get:

(Minimal Physicalism): All non-physical properties minimally supervene on physical properties, and no mental property is fundamental.

A physicalism based on minimal supervenience does better than one based on standard supervenience. First, minimal physicalism comes out false, as it should, if there are anti-physicalist properties in the actual world. If there are fundamental mental properties in our world, the NFM constraint is violated. And if there are fundamental unscientific properties in our world, the non-physical properties would fail to minimally supervene on physical properties. Second, physicalism can be true in the actual world, despite there being possible worlds with such anti-physicalist properties. Physicalism, on the minimal physicalist account, only requires that all *minimal* physical duplicates of the actual world are indiscernible with respect to non-physical properties, and worlds with extra fundamental unscientific properties are not minimal physical duplicates. These possible worlds are therefore irrelevant as far as the truth of physicalism in the actual world is concerned.

The grounding physicalist should accept that NP-P grounding entails minimal NP-P supervenience, rather than standard NP-P supervenience, for the same reason a supervenience physicalist should base her physicalism on minimal, rather than standard, supervenience. Intuitively, NP-P grounding in the actual world is unaffected by there being duplicates of the actual world with extra anti-physicalist properties. And this intuition is not threatened by NP-P grounding entailing minimal NP-P supervenience. At the same time, if the actual world contains anti-physicalist properties, grounding physicalism would fail, and this is revealed by a failure either to satisfy the NFM constraint (if the anti-physicalist properties are fundamental mental properties) or a failure of minimal supervenience on the physical (if the anti-physicalist properties are fundamental unscientific properties).

No doubt there is more to be said about the relationship between grounding and supervenience. For one thing, there are reasons to think that minimal NP-P

supervenience (along with the NFM constraint) is insufficient for physicalism.[38] If so, this would indicate that NP-P grounding entails a form of supervenience stronger than minimal supervenience. Of course, this does not change the fact that failure of minimal NP-P supervenience is sufficient for failure of physicalism and NP-P grounding, and for purposes to come, using failure of minimal NP-P supervenience to indicate failure of grounding is good enough.

1.11 Strong and Global Supervenience

There is a further question of whether the supervenience entailed is strong or global supervenience. The relation between the two has been the subject of a fairly complex debate in the literature on supervenience. Rather than rehash that I will proceed with the assumption that grounding entails global supervenience and leave it an open question whether it also entails strong supervenience. There are two reasons for this. Firstly, supervenience definitions of physicalism are typically formulated in terms of global supervenience and a formulation of physicalism is what is at issue. And secondly, strong NP-P supervenience raises some issues about semantic externalism and the wide content of mental states that are best avoided here.

1.12 The Modal Strength of the Relevant Supervenience

What is the modal strength of the supervenience entailed by grounding? Is it metaphysical or merely nomic? Before answering this question, an ambiguity needs to be cleared up. What we might call the textbook distinction between metaphysical and nomic necessity I take to be roughly this: Something is metaphysically necessary at a world *w* if and only if it is the case in every possible world, *regardless of what laws of nature hold in those worlds*. Something is nomically necessary at a world *w* if and only it is the case in every possible world *with the same laws as w*.[39] So for instance, it might be nomically necessary that, in the actual

[38] This is due to the possibility of what John Hawthorne calls 'blockers' (Hawthorne [2002]). Blockers are properties that block the instantiation of supervenient non-physical property, even when the relevant supervenience base is present. According to Hawthorne, physicalists should reject the possibility of blockers, since they do not believe that the tie between physical and non-physical can be so easily disrupted. But minimal physicalism is compatible with the existence of possible worlds that are (i) physical duplicates of our world, and (ii) contains blockers that prevent the instantiation of certain non-physical properties contained in our world.

[39] There are further distinctions that can be made within nomic necessity, corresponding to which laws are held fixed. For instance, we might talk of biological necessity, where the fixed laws are

world, water boils at 100 degrees Celsius in 1 atmosphere pressure, since, in every possible world with the same laws as the actual world, water boils at 100 degrees Celsius in 1 atmosphere pressure. However, it is arguably not metaphysically necessary as there are possible worlds with different laws of nature, where water boils at a different temperature in 1 atmosphere pressure or boils at 100 degrees Celsius under different pressures. That water is H2O, however, is the case in every possible world (where water exists at least), regardless of the laws that hold in those worlds.[40]

We can use this textbook distinction between metaphysical and nomic necessity to define corresponding notions of minimal metaphysical and nomic supervenience:

(Minimal metaphysical supervenience):

> A minimally supervenes on B with metaphysical necessity in world *w* iff for any world *w**, *regardless of the laws in w**, if *w** is a minimal B-duplicate of *w*, *w* and *w** are indiscernible with respect to their A-properties.

(Minimal Nomic supervenience):

> A minimally supervenes on B with nomic necessity in world *w* iff for any world *w**, *with (all and only) the same laws as w*, if *w** is a minimal B-duplicate of *w*, *w* and *w** are indiscernible with respect to their A-properties.

The distinction between minimal metaphysical and nomic supervenience is important for the same reason it was important to determine whether the supervenience entailed by grounding is standard or minimal[41] - if failure of NP-P supervenience is to be used as a test to indicate failure of NP-P grounding, and by implication failure of physicalism, we will need to know whether the supervenience entailed is metaphysical or merely nomic. For instance, if the grounding of the non-physical in the physical entails only nomic supervenience of the non-physical on the

---

biological laws, causal necessity, where the fixed laws are causal laws, non-causal necessity, where the fixed laws are non-causal laws, and so on. All these I take to be varieties of nomic necessity. More on that shortly.

[40] These are just examples. Some would deny that water is H2O in every possible world or that there are worlds with different laws from the actual world where water boils at different degrees.

[41] From now on, unless otherwise noted, supervenience is to be understood as minimal supervenience.

physical, it will be no strike against physicalism if metaphysical supervenience of the non-physical on the physical fails to hold. Now, seeing that grounding is a metaphysical relation it would be natural to think that the kind of supervenience entailed by grounding is metaphysical supervenience. This conclusion also seems to be backed up by the literature on supervenience physicalism. For instance, David Chalmers says that physicalism (or in Chalmers's terminology, materialism) is true, '…if for any *logically possible world W* that is physically indiscernible from our world, all the positive facts of our world are true of W' (my italics).[42] And Frank Jackson's minimal physicalism basically expresses the same idea when he says that physicalism is the claim that *any* world which is a minimal physical duplicate of the actual world, is a duplicate *simpliciter*.[43] However, contrary to first appearance, it is not clear that these views are endorsing the textbook definition of metaphysical NP-P supervenience, as a necessary condition for the truth of physicalism. The problem is that the supervenience base for the non-physical may itself contain the laws of the actual world. This, for instance, is explicit in Jackson's formulation of physicalism, where a physical duplicate of our world is a world that is identical with respect to both physical properties *and* physical laws. In this case, the textbook definition of metaphysical supervenience simply collapses into a form of textbook nomic supervenience. If A supervenes on B with nomic necessity, it follows that any world with the same laws as the actual world that is indiscernible with respect to B will also be indiscernible with respect to A. But those are the very same worlds that are indiscernible with respect to B *and all actual laws*. In other words, if A metaphysically supervenes on B and all actual laws, this is equivalent to A nomically supervening on B. So there is an ambiguity in the notion of metaphysical supervenience. It can be metaphysical supervenience of A on B in the textbook sense, where the relevant possible worlds are all B-indiscernible worlds, regardless of the laws that hold in those worlds. Or it can be metaphysical supervenience of A

[42] Chalmers [1996], p. 42. A couple of notes: Chalmers formulates physicalism in terms of logical supervenience, though to him, metaphysical and logical supervenience comes to the same thing. Also, while Chalmers seems to say that metaphysical supervenience of everything on the physical is merely sufficient for physicalism, he clearly takes it to be a necessary condition as well, as his main argument against physicalism rests on supervenience failing to hold between the non-physical and the physical, which in turn, is supposed to show that physicalism is false. Obviously that would not follow if supervenience of the non-physical on the physical was merely sufficient for physicalism as physicalism might then be true for other reasons.

[43] Jackson [1998], p. 12.

on B *and all actual laws*, which is equivalent to the textbook definition of nomic supervenience.

With this ambiguity cleared up, we can return to the question of what the modal strength of the supervenience entailed by grounding is. The answer is that it depends on what goes in the grounding base for a given entity. For instance, a grounding base for determinable colours such as red, green, and blue, plausibly includes the set of determinate colours, such as scarlet, lime-green, and navy-blue - it is in virtue of having a determinate colour that objects have determinable colours. Does this grounding base also include any laws? Probably not. That determinate colours ground determinable colours seems to be independent of whatever laws obtain. Rather, it follows from the very nature of these entities themselves. To give it a name, we can call the kind of grounding where the grounding base does not contain any laws *metaphysical grounding*. So the supervenience that is entailed by the grounding of determinable colours in determinate colours is proper textbook metaphysical supervenience – any possible world, regardless of the laws that hold in that world, that is indiscernible from the actual world with respect to determinate colours, will be indiscernible with respect to determinable colours. But now consider the further claim that colours are dispositions to reflect light at certain wave-lengths, and that the grounding base for dispositions includes the set of categorical properties. Again, we might consider whether this grounding base include laws in addition to categorical properties. And now the answer is arguably yes. For instance, if the laws of optics had been very different, the same categorical properties would have underlain entirely different dispositions to reflect light. So the grounding base for dispositions plausibly includes both categorical properties *and* the laws of nature.[44] Hence, we can call the grounding involved here *nomic grounding*. If the grounding base for colour-properties includes the laws of our world, the supervenience entailed by this grounding is merely a form of textbook nomic supervenience – any possible world, with the same laws of nature as the actual world, that is indiscernible from the actual world with respect to categorical properties, will be indiscernible with respect to colours.

[44] Again, these are merely examples to illustrate a point.

To sum up then, what these examples show is that there is no univocal answer to the question of which kind of supervenience is entailed by grounding – it will depend on the kinds of entities in question.

1.13 Grounding of the Non-Physical in the Physical: Which Supervenience is Entailed?

Of course, what we are interested in is not any old grounding base. Seeing that, according to physicalism, the set of physical properties coincides with the set of fundamental properties (and no mental property is among them) and grounds everything else, what we are interested in is the total fundamental grounding base for the non-physical. Are there any laws in *this* grounding base? On most physicalist accounts, the answer is yes. For one thing, the fundamental physical laws will go in the grounding base for the non-physical. This is particularly clear if we consider physicalists that are also functionalists. For instance, Jackson explicitly takes the supervenience base for the non-physical to be physical properties *and* laws and this supervenience-base plays the role that the grounding base for the non-physical plays for a grounding physicalist.[45] Another example is Jaegwon Kim who argues that mental properties are definable in terms of their causal relations to other properties and realized by physical properties. Talking about a mental property *M*, Kim says:

> To functionalize *M* is to make *M* non-rigid, and this is easily seen: M is defined in terms of its causal/nomic relations to other properties, and since these relations are contingent – *contingent on the laws that prevail in a given world* – it is a contingent fact whether a given property satisfies the causal/nomic specification that is definitive of *M*.[46]

In other words, to fix the domain of mental properties, on Kim's account, it is not enough to fix the domain of non-mental properties, you also have to ensure that the prevailing laws are held fixed. Otherwise the realizers of mental properties cannot be guaranteed to enter into the same causal relations they do in the actual world and hence, cannot be guaranteed to enter into the causal relations that are distinctive of

[45] Jackson [1998], p. 11.
[46] Kim [1998], p. 99, my italics.

the mental properties they realize in the actual world. The grounding base for mental properties on Kim's account therefore includes laws.

Are there forms of physicalism that do not include laws in the grounding base for the non-physical? A technical variant would be a form of physicalism that allows only for physical properties and no non-physical properties. Such a physicalism would not include laws in the grounding base for the non-physical for the simple reason that there is nothing non-physical to ground and hence no grounding base for it. What of non-functionalist forms of physicalism? Though they may not include laws in the grounding base for the non-physical for all the same reasons that functionalists do, it is still natural to think that physicalists in general will take fundamental physical laws to be part of the grounding base for the non-physical. For instance, consider the view that physical micro-properties ground non-physical macro-properties, not because the micro-properties satisfy any causal role distinctive of the macro-properties, but because micro-parts (and their micro-properties) compose macro-wholes (and their macro-properties). However, composition itself may depend on fundamental physical laws. For instance, the criterion for when some parts compose a whole may be that the parts 'hang together' in some suitable way. But whether some parts hang together will be dependent on the laws of nature. For instance, eliminate the strong nuclear force and subatomic particles would no longer combine into atoms. So on this account, it would not be enough to fix the physical properties of the actual world in order to fix the non-physical properties, you would also need to fix, at least some, laws of nature. And so, the grounding base for the non-physical according to this particular non-functionalist form of physicalism would include laws. And presumably something similar will be the case for most other forms of physicalism. The physical laws of our world make a difference to facts about non-physical properties and so, cannot be left out of the grounding base for the non-physical if this is to be able to support supervenience. And as a result, the supervenience entailed by grounding the non-physical in the physical can only be nomic supervenience, or at least some form of it, of the non-physical on the physical.

An exception to this would be forms of physicalism that wishes to do without laws at all, at least in the traditional sense of laws. Alexander Bird, for instance, has

argued that properties have *dispositional essences*.[47] What this means is that the causal profile of a property, which specifies the causal relations into which the property can enter, could not possibly have been different than it actually is. If the causal profile is different, it is a profile of a different property. This is known as *dispositional essentialism* (or simply, as I will sometimes refer to it, essentialism). Bird also believes that laws of nature can be generated directly from properties and their dispositional essences. In other words, laws of nature should not be thought of as entities in the world that *govern* the behaviour of properties, rather they flow directly from these properties. The upshot of such a view, when coupled with physicalism, is that the grounding base for the non-physical does not contain laws, at least not laws in the traditional sense if being some entities in the world in addition to properties. Still, to say that it follows from this that NP-P grounding entails metaphysical NP-P supervenience would be misleading. First of all, while a dispositional essentialist may not posit laws in the traditional sense, she will still have a revisionary notion of laws as being derived from the causal profiles of properties. For instance, if it is part of the causal profile of a property *P* that it has (or bestows) the power to cause effect E, then it will be a law that *P* causes E. On this account, to fix the properties of a world is *ipso facto* to fix the laws of that world, and *vice versa*. This means that the set of worlds that are indiscernible with respect to the properties of the actual world just *is* the set of worlds with the same laws as the actual world. And similarly, if physicalism is true, the set of worlds that are indiscernible from the actual world with respect to physical properties will be exactly the set of nomically indiscernible worlds. Metaphysical NP-P supervenience thus collapses into nomic NP-P supervenience on a dispositional essentialist account.[48]

In conclusion, there are no obvious candidates for forms of physicalism that endorses supervenience of the non-physical on the physical in the textbook sense of metaphysical supervenience, and even if there were forms of physicalism that entail

[47] In Bird [2007].
[48] It is not compulsory for essentialists to be revisionists about laws of nature. One can maintain that properties have dispositional essences and hold that these essences are determined by the laws of nature. The result is much the same though, as the distinction between nomic and metaphysical supervenience collapses.

this, it should be clear by now that this is not a commitment of physicalists in general.

1.14 Concluding Remarks

In this chapter I have argued that physicalism is best taken as the view that completed physics provides a complete list of all fundamental properties, none of which are mental. Since being non-fundamental is to be ultimately grounded in the fundamental, this means that all non-physical properties are grounded in the properties posited by completed physics. While this is a better definition of physicalism than one in terms of supervenience, there are nevertheless interesting modal consequences of physicalism. In particular, it entails that the non-physical minimally and nomically supervenes on the physical.

Chapter 2

Emergentism and Reductive vs. Non-Reductive Physicalism

Abstract: In this chapter I first discuss how to distinguish emergentism from physicalism. I argue that emergentists posit the existence of fundamental causal laws which should be considered non-physical, and which should therefore be excluded from a physicalist grounding base for non-physical properties. As a result, some non-physical properties have causal powers not grounded in the physicalist base in an emergentist world, and this is incompatible with physicalism. I then argue that emergentism is incoherent if dispositional essentialism is true. Second, I consider how to draw the distinction between reductive and non-reductive physicalism. I argue that we need a notion of what I call broadly physical properties, which can function as a reduction base for other non-physical properties. These broadly physical properties can be generated from fundamental physical properties via the relations of mereological fusion, conjunction, and disjunction.

2.1 Emergentism

One of the troublesome positions when it comes to giving a satisfactory formulation of physicalism is the position defended by the British Emergentists.[49] The reason is that while the position of emergentism is typically thought of as being in opposition to physicalism, traditional formulations of physicalism, in particular supervenience based definitions, have had difficulties excluding emergentism from physicalism. In the following I will discuss how best to understand the doctrine of emergentism and to what extent the truth of a grounding-based definition of physicalism rules it out.

Roughly, emergentism is the view that while everything is ultimately composed of parts with purely physical properties, some wholes have emergent properties with genuinely novel causal powers. For instance, humans may have emergent mental properties, organisms may have emergent biological properties, molecules may have emergent chemical properties, and so on. What is required for emergence to occur is a certain degree of complexity at the physical level. Once such complexity is reached among the physical properties, new properties emerge. This emergence is

[49] Here I focus on C.D. Broad's formulation of the view from Broad [1925]. For other proponents of emergentism, see Alexander [1920] and Morgan [1923]. Philosophers who have pointed out some of problems associated with distinguishing emergentism from physicalism include Crane [2007], Kim [1992b], and Wilson [2005].

governed by what C. D. Broad called *trans-ordinal laws*. These are non-causal laws that connect the physical properties in the emergence-base with the higher-level emergent properties. In addition, higher-level emergent properties have novel causal powers not found at any lower level – not only powers to cause other emergent properties, but also powers to cause physical properties at lower levels. In so far as the causal powers a property has are dependent on the prevailing causal laws, emergentists are therefore also committed to the existence of fundamental causal laws that govern the (causal) behaviour of emergent properties, in addition to their commitment to fundamental trans-ordinal laws. The fact that these causal laws are fundamental is what ensures the genuine novelty of emergent properties. I will call such laws *emergent causal laws*.[50] [51]

It is arguably a part of the emergentist position that physical emergence-bases (i.e. the base from which emergent properties arise) are also supervenience-bases for emergent properties.[52] That is, if emergentism is true in world *w*, then any world that is indiscernible with respect to physical properties and their relations will include all emergent properties of *w*. On a supervenience-based account of physicalism, according to which NP-P supervenience (that is, supervenience of the non-physical on the physicals) is sufficient for physicalism, this makes it difficult to see how emergentism is to be distinguished from physicalism. The same problem arises for grounding physicalists if emergent properties are grounded in the physical properties they emerge from. If they are, what is to separate emergentism from physicalism? Whether or not the British Emergentists took emergent properties to be fundamental or grounded in physical properties is a historical question.

[50] Broad does not use the term 'emergent causal law'. Instead he contrasts trans-ordinal laws with what he calls intra-ordinal laws. These are causal laws that connect properties within a level. It is not obvious to me, however, that this distinction tracks the distinction between trans-ordinal laws and emergent causal laws as I understand them. For one thing, emergent causal laws not only (causally) connect properties at the same level, but also (causally) connect properties at the emergent level with properties at lower level. Nevertheless, even if emergent causal laws can be trans-ordinal in the sense that they connect properties at different level, they are not examples of what Broad has in mind when he talks of trans-ordinal laws since, to him, trans-ordinal laws do not *causally* connect physical properties with emergent properties. Rather, the connection is constitutive or metaphysical.

[51] Sometimes emergentism is characterised in epistemic terms, as the view that facts about emergent properties cannot, even in principle, be deduced from complete facts about the physical. Correspondingly, physicalism is given an equally epistemic interpretation and involves the denial of this claim. Here, however, I focus on a metaphysical interpretation of emergentism, just as I took physicalism to be a metaphysical position in the previous chapter.

[52] See the quote from Broad below. For a recent defence of this claim, see Horgan [1993].

Personally, I believe there is some evidence for the latter. Consider, for instance, what Broad says about emergent chemical properties:

> No doubt, the properties of silver-chloride are completely *determined* by those of silver and of chlorine; in the sense that whenever you have a whole composed of these two elements in certain proportions and relations you have something with the characteristic properties of silver-chloride, and that nothing has these properties except a whole composed in this way.[53]

Now contrast this view about emergent chemical properties with what Broad says about trans-ordinal laws immediately after:

> But the law connecting the properties of silver-chloride with those of silver and of chlorine and with the structure of the compound is, so far as we know, an *unique* and *ultimate* law.[54]

It seems clear that Broad is pointing out what he takes to be an important difference between, on the one hand, emergent chemical *properties* and, on the other, the trans-ordinal *law* that connects them with the underlying physical properties. The former are, as Broad puts it, completely determined by their constituent physical properties, the latter is unique and ultimate. I believe a plausible interpretation of this is that Broad is making the point that emergent properties are dependent upon or grounded in physical properties, whereas trans-ordinal laws are fundamental. But if NP-P grounding is sufficient for physicalism (along with the constraint that there be no fundamental mental properties), and emergent properties are among the non-physical properties grounded in physical properties, then it seems that emergentism will count as a form of physicalism on a grounding physicalist account as well.[55]

[53] Broad [1925], p. 64.
[54] Broad [1925], pp. 64-65.
[55] An interpretation of emergentism according to which emergent properties *are* fundamental, would be easily dealt with from the perspective of both supervenience and grounding physicalism. If the emergent properties are mental, emergentism would count as a non-physicalist view due to violating the NFM constraint. If the emergent properties are unscientific, there would be fundamental properties not treated by completed physics, and physicalism would also fail. If the emergent properties are non-mental and scientific, emergentism would simply be a form of macro-physicalism.

2.2 Using the Nomic/Metaphysical Distinction to Sort Emergentism from Physicalism

One popular way to solve this problem has been by way of the distinction between metaphysical and nomic supervenience. For instance, Robert Howell, after having argued that physicalism requires metaphysical NP-P supervenience, goes on to say: "Since properties emerge only given the existence of emergence laws, […], they seem to emerge nomologically but not metaphysically."[56] This particular way of drawing the distinction between physicalism and emergentism has been criticised on the grounds that it relies on a non-essentialist view of properties.[57] The objection is the following: On the dispositional essentialist account of properties, metaphysical NP-P supervenience collapses into nomic NP-P supervenience, since the set of worlds indiscernible from our world with respect to physical properties, just *is* the set of worlds indiscernible from our world with respect to physical properties *and* the laws of nature of our world. In this sense, emergentists are as committed to metaphysical NP-P supervenience as a physicalist is. But while this may be right, the discussion from the previous chapter clearly shows that there is a deeper problem, independent of the debate over essentialism, as physicalists themselves are committed only to a form of nomic NP-P supervenience. So contrary to Howell's claim, metaphysical NP-P supervenience cannot be necessary for physicalism. If the difference between physicalism and emergentism is to be found in the modal strength of the NP-P supervenience these positions entail, the difference must be located *within* the realm of nomic possibility. For instance, it may be that the emergentist allows for certain failures of nomic NP-P supervenience which, in turn, would falsify physicalism. If so, this would also indicate that physicalists and emergentists differ on what they take to be the grounding base for non-physical properties.

2.3 Physical and Nomic Supervenience

Consider again the interpretation of Broad's claims, according to which emergent properties are grounded in physical properties. It is clear, however, that the physical properties cannot make out the *complete* grounding base for emergent properties. For one thing, emergent properties require for their instantiations that there are

[56] Howell [2009], p. 90. In the quote, 'emergence laws' is referring to trans-ordinal laws.
[57] See Wilson [2005], pp. 436-437.

trans-ordinal laws that connect them to physical properties. This failure of grounding is revealed by a failure of supervenience – a minimal duplicate of our world with respect to physical properties *only* will not contain any trans-ordinal laws, and so, will be a world in which emergent properties fail to instantiate. If emergent properties are to be grounded in physical properties, it must therefore be the case that trans-ordinal laws are included in the grounding base. The same point applies to the causal powers of emergent properties, the existence of which requires that there are emergent causal laws. So these powers cannot be grounded only in physical properties and trans-ordinal laws. Again, this failure of grounding is revealed by a failure of supervenience - a minimal duplicate of our world with respect to physical properties and trans-ordinal laws, but in which there are no emergent causal laws, will be a world where emergent properties instantiate but have no causal powers (or at least not all the powers they have in our world). These points suggest a solution to the problem of distinguishing emergentism from physicalism: If somehow the physicalist can exclude trans-ordinal and emergent causal laws from the relevant supervenience base for non-physical properties, then the truth of physicalism will entail that emergentism is false.

The most obvious way of drawing the physicalism-emergentism distinction in such a fashion is by way of what I'll call *physical supervenience.* Minimal physical supervenience is defined as follows:

(Minimal Physical Supervenience):

> A minimally supervenes on B with physical necessity in world *w* iff for any world *w*, with (all and only) the same physical laws as w*, if *w** is a minimal B-duplicate of *w*, *w* and *w** are indiscernible with respect to their A-properties.

Note how nomic and physical supervenience comes apart if there are fundamental non-physical laws in our world. Nomic supervenience makes a claim about worlds with (all and only) the same laws as *w*, whatever these laws might be. But if there are fundamental non-physical laws in *w*, these worlds will be different from worlds with (all and only) the same physical laws of *w*. This means that physicalism based on nomic supervenience has different truth conditions from physicalism based on

physical supervenience. If there are fundamental non-physical laws in our world, in virtue of which certain non-physical properties are instantiated and have certain causal powers, then physicalism based on nomic supervenience might be true (provided there are no fundamental unscientific or mental properties in our world). However, physicalism based on physical supervenience would be false (a minimal duplicate of all physical laws and physical properties in the actual world, would either fail to include the non-physical properties partly grounded in fundamental non-physical laws, or include them but fail to include their causal powers).

2.4 Physical Supervenience and Emergentism

As far as using this distinction between physical and nomic supervenience to distinguish supervenience physicalism and emergentism, the idea is this: Emergentists are committed to the existence of fundamental non-physical laws. If such laws exist, physicalism based on physical supervenience fails, yet physicalism based on nomic supervenience *simpliciter* is true. Since emergentism is intuitively and historically incompatible with physicalism, a definition of the latter in terms of physical supervenience is superior to one in terms of nomic supervenience, and will allow us to distinguish supervenience physicalism from emergentism.

A similar move is available to the grounding physicalist. For the supervenience physicalist, the way to sort emergentism from physicalism is by restricting the supervenience base for non-physical properties to only physical properties and physical laws. Likewise, the grounding physicalist can restrict the grounding base for non-physical properties to physical properties and physical laws. Call the kind of grounding where the laws included in the grounding base are only physical laws for *physical grounding*. The argument would then go like this: Emergentists are committed to the existence of fundamental non-physical laws which partly ground some non-physical properties and their causal powers. If such laws exist, physicalism based on physical grounding fails, yet physicalism based on nomic grounding is true. Since emergentism is intuitively and historically incompatible with physicalism, a definition of the latter in terms of physical grounding is superior

to one in terms of nomic grounding, and will allow us to distinguish physicalism from emergentism.[58]

2.5 What is a Non-Physical Law?

Of course, the way of sorting emergentism from physicalism described above assumes that, indeed, the emergentist *is* committed to fundamental non-physical laws. In the literature, the focus tends to be on trans-ordinal laws, which are often assumed to be non-physical laws, without argument.[59] It seems to me though, that if we actually try to establish some criteria for what makes a law a non-physical law, emergent causal laws are more problematic from a physicalist perspective.

Assuming that we are not entitled to simply stipulate that the laws posited by emergentists are non-physical, this raises the question of what makes a law a physical law. The necessary and sufficient condition for a property being physical is that the property is a posit of completed physics. Correspondingly, a non-physical property is one that fails to be a posit of completed physics. We might try to use the same definitions for physical and non-physical laws. So a law is physical iff it is a posit (or part of) completed physics, and a law is non-physical if it is not.

One reason a law could fail to be a posit of completed physics would be that it is a non-fundamental law. But this obviously does not apply to trans-ordinal and emergent causal laws, since both are fundamental laws on the emergentist account. A second reason might be that trans-ordinal or emergent laws are unscientific and simply incapable of being incorporated into scientific theory. Consider, for instance, a causal law according to which *P*'s cause *Q*'s. If this law manifests entirely arbitrarily and *P* only sometimes causes *Q*, even when background conditions are held fixed, then it would likely be such that it would not satisfy the criteria for a scientific posit, and so, would not be incorporable into scientific theory. But neither trans-ordinal laws nor causal emergent laws are like that on the emergentist

[58] A small caveat here: While the fundamentality of trans-ordinal and emergent causal laws prevent them from being grounded in physical properties and laws, it does not, strictly speaking, prevent them from supervening on physical properties and laws, since they may supervene for other reasons. For instance, if they are metaphysically necessary, they trivially supervene on anything. I take it, however, this is quite implausible, and in any case, it would only cause problems for a supervenience based definition of physicalism, not the grounding based definition I endorse.
[59] As examples, see Yates [2009], p. 116, McLaughlin [2007], p. 161

account. Consider first trans-ordinal laws. If certain physical conditions are present and property *E* emerges, then whenever those physical conditions are reproduced in our world, *E* will emerge. Trans-ordinal laws thus do not manifest arbitrarily. In fact, the methodology for discovering trans-ordinal laws seems to satisfy the hallmarks of scientific methodology. They are uncovered empirically and inductively, and their manifestations are reproducible. In this sense there is nothing unnatural about them. As Broad himself notes:

> There is nothing, so far as I can see, mysterious or unscientific about a trans-ordinal law […]. A trans-ordinal law is as good a law as any other; and, once it has been discovered, it can be used like any other to suggest experiments, to make predictions, and to give us practical control over external objects.[60]

Presumably the same applies to emergent causal laws. There is nothing in the emergentist account which suggests that such laws are any different from other causal laws of science.

Thirdly, it might be argued that trans-ordinal and emergent causal laws lack the generality that genuine physical laws exhibit. For instance, a law such as Newton's law of gravitation applies to every object in the universe (with mass at least). Trans-ordinal and emergent causal laws, on the other hand, act only on certain very specific arrangement of physical matter (in the case of trans-ordinal laws) or specific emergent properties (in the case of emergent causal laws).[61] So if physics posit only laws with a high degree of generality, trans-ordinal and emergent causal laws will not be posits of completed physics, despite being fundamental. It is one thing, however, to say that the fundamental laws of nature posited by current physics exhibit a high degree of generality, it is another to say that future physics could not *possible* posit laws that act only to very specific entities. The latter claim is what needs to be justified if the specificity of trans-ordinal and emergent causal laws is enough for them to count as non-physical. While generality is no doubt a

[60] Broad [1925], p. 79.
[61] By a law 'acting upon' some properties I mean that these properties have powers to produce certain effects in virtue of this law obtaining. The production may be causal (as in the case of an emergent causal law) or non-causal (as in the case of a trans-ordinal law).

feature of current laws of physics, it is not 'built in' to physics from the get-go that this must be so.

Finally, and this seems to me to be the more promising strategy, a law may fail to be a posit of completed physics because it does not act on physical properties. Since physics describes only fundamental properties and their interactions, such a law, even if fundamental, would simply be left out of completed physical theory. Emergent causal laws are like this. They act only on emergent non-fundamental and non-physical properties, ensuring that these properties have novel causal powers, not found at the level of the physical properties from which they emerge. Since they do not act on physical properties, such laws simply fall outside the scope of completed physics. If emergent causal laws count as non-physical in this way, they should be excluded from a physicalist grounding base for non-physical properties. This arguably makes physicalism false in an emergentist world, since that world would contain non-physical properties with causal powers not grounded in the physical properties and laws. This is also revealed by failure of supervenience – a world that is a minimal duplicate of all the physical properties and physical laws of an emergentist world would differ at the level of non-physical properties, since the non-physical properties instantiated in the duplicate world fail to have all the causal powers they have in our world.

Note that this argument for the non-physicality of emergent causal laws does not obviously apply to trans-ordinal laws, since trans-ordinal laws *do* act on physical properties. Still, we might wonder if trans-ordinal laws are really the kind of laws that could be posited by completed physics. If completed physics posits only laws that govern how physical properties interact with other *physical* properties, then trans-ordinal laws would not be posited by completed physics, since trans-ordinal laws govern how physical properties interact with other non-physical properties (as noted, the interaction here is non-causal). Nevertheless, there are independent reasons for why the physicalist should be hesitant to exclude them from a physicalist grounding base on this basis. To see this, note that there might be fundamental non-physical laws that the physicalist would want in the grounding base for non-physical properties. For instance, say that composition is a law-governed relation between parts and wholes and that the fundamental physical

objects are micro-objects. Say also that, in our world, the law of composition L obtains and acts on certain collection of micro-level parts, which consequently make out macro-level wholes. In some other worlds, there is no such law and no wholes exist. If we exclude L from physicalist grounding base for non-physical properties, the existence of composite wholes is enough to make physicalism false in our world. This, however, seems awfully rough. Surely someone can consistently claim that composition is law-governed and occurs in our world, and still be a physicalist. Physicalism, it seems, should therefore allow for laws like L in the grounding base for non-physical properties. But if L is allowed in the grounding base, despite the fact that it connects physical and non-physical properties or objects, then why shouldn't trans-ordinal laws be allowed as well? Personally, I do not believe there is a problem here for the physicalist. Consider again Broad's example with the properties of silver-chloride being completely determined by the properties of silver and chloride. If we take away the claim that silver-chloride has novel and fundamental causal powers, do we then have a view that should be considered anti-physicalist? I do not think so. Physicalists need not worry about the mere existence of silver-chloride, as long as the properties and powers of silver-chloride are fully grounded in the underlying physical properties and powers. And as long as there are no emergent causal laws, this will be the case. Thus, trans-ordinal laws alone do not pose a threat to physicalism and, should they exist, the physicalist can happily accept them in the grounding base for non-physical properties.

2.6 Emergentism and Dispositional Essentialism

In the debate over how to distinguish emergentism from physicalism, dispositional essentialism (or simply essentialism) requires special treatment. As I proposed above, emergentism is incompatible with physicalism due to the fact that emergentists are committed to the existence of fundamental emergent causal laws, in virtue of which emergent properties have novel and fundamental causal powers. The physicalist should exclude these laws from the grounding base for non-physical properties, on the basis that they are non-physical laws. This has as a consequence that physicalism is false in an emergentist world, since, in the latter, some non-physical properties would have causal powers not grounded in the relevant base. However, if the relevant base allows for trans-ordinal laws, this move is unavailable

to the essentialist. According to emergentism, physical properties plus trans-ordinal laws necessarily suffice for the instantiations of some non-physical emergent properties. But if essentialism is correct and properties have their causal powers essentially, this means that physical properties plus trans-ordinal laws necessarily suffice for the causal powers that makes physicalism false according to the proposal suggested earlier.

A response might be to insist that, contrary to my earlier claim, this shows that trans-ordinal laws must be excluded from the physicalist grounding base after all. But even if some criteria could be produced that would prohibit trans-ordinal laws from the physicalist grounding base, this does not effectively solve the problem. For what if the essentialist position is that properties have *all* their powers, both causal and non-causal, essentially? If so, trans-ordinal laws would reduce to essential dispositions or powers of physical properties. But then, to exclude trans-ordinal laws from the relevant grounding base would be to exclude physical properties themselves from the relevant grounding base. And this is surely not an option for the physicalist.

2.7 Howell: Trans-Ordinal Laws Contaminate the Physicalist Grounding Base

A different approach to this problem is to argue, as Robert Howell does in a recent paper, that if some properties are essentially disposed to give rise to emergent properties, this in itself disqualifies them from being genuine physical properties. As Howell explains:

> Let's grant to the emergentist that there is a genuinely new emergent property E which emerges necessarily from C. [*Where C is some configuration of physical properties (author's insert)*] […] By hypothesis, E is a fundamental, mental property, which means that C is essentially characterized by the disposition to produce a fundamental mental property. […] How could the dispositional property to give rise to a new, non-physical property itself be physical? Even if one wanted, […], to call C itself physical, the property that is nothing but the disposition to give rise to something non-physical surely is not.[62]

[62] Howell [2009], p. 93.

In other words, combining emergentism and essentialism leads to a contamination of the physicalist grounding base. It either fails to contain any genuinely physical properties at all (because the properties in the base give rise to anti-physicalist properties such as fundamental mental properties), or it necessarily includes something non-physical (the property that is nothing but the disposition to give rise to anti-physicalist properties). In either case, if emergentism and essentialism is combined, a physicalist grounding base is unavailable to the physicalist.

I think Howell's argument is problematic for a couple of reasons. First, from the perspective of the conclusions and terminology developed in the previous chapter, it is unnecessarily complicated. If emergentism is true and emergent property *E* is fundamental and mental, then physicalism will already be falsified on the grounds that emergentism violates the NFM constraint. Whether or not the properties from which *E* emerges are proper physical properties simply becomes an irrelevant issue. Second, the argument targets only emergentism about mental properties. Though Howell does make the point that, in general, a group of properties C being disposed to give rise to a non-physical property entails that C cannot be physical, he surely cannot be right about this. After all, it is entirely uncontroversial that physical properties give rise to all kinds of non-physical properties. For instance, put a bunch of fundamental entities with fundamental physical properties together in the right way and you can create tables, chairs, mountains, etc. But no one would argue that the particles composing a chair are not physical entities, just because the thing they compose isn't strictly speaking a physical entity. Even if you add the essentialist assumption that certain particles are *essentially* disposed to give rise to, say, chairs (when put together in the right way), this still seems like a thing the physicalist need not worry about in the slightest. Therefore, what Howell means must be that physical properties cannot be essentially disposed to give rise to properties the existence of which would *falsify* physicalism. Obviously, properties such as *being a table* or *being a mountain* do not fall in that category. The question now is, will Howell's reasons for thinking that emergentism about mental properties is incompatible with physicalism apply to emergentism about other kinds of properties, for instance, chemical properties? The answer is: Only if you think that the existence of fundamental chemical properties automatically falsifies physicalism. As I argued earlier, this is confusing physicalism with micro-

physicalism, something we should be careful not to do. If Howell insists that fundamental chemical properties falsifies physicalism, then he should endorse a 'no fundamental chemical properties' constraint on physicalism (or the physical), in addition to a NFM constraint, and likewise for all the other kind of properties one could be an emergentist about. In any case, if one endorses such a plurality of constraints on what counts as physical or what physicalism can accept, the first problem with Howell's argument reappears – if the existence of fundamental chemical properties falsify physicalism, and emergentists about chemical properties are committed to the existence of said properties (given essentialism), emergentism about chemical properties is automatically disqualified as a form of physicalism. There is no need to consider whether properties can have essential dispositions to give rise to chemical properties and still be considered physical. Thus, my initial response to Howell takes the form of a dilemma: If physical properties are essentially disposed to give rise to anti-physicalist properties, then the existence of these properties is enough to falsify physicalism. If physical properties are essentially disposed to give rise to non-physical, but not anti-physicalist properties, then I can see no reason for thinking that the grounding base for such properties is contaminated from a physicalist perspective, just like a physicalist grounding base is not contaminated by 'giving rise to' tables and chairs.

2.8 Emergentism is Incoherent if Dispositional Essentialism is True

Another problem is that Howell assumes that emergent properties are fundamental. This, however, is contrary to my interpretation of emergentism, according to which emergent properties are non-fundamental properties with fundamental causal powers. Nevertheless, Howell's assumption is a natural one given essentialism. According to emergentism, emergent properties have fundamental causal powers due to the existence of fundamental causal laws. This is compatible with emergent properties being non-fundamental, only given the assumption that a property and the powers it has can come apart. But of course, on an essentialist account, this is impossible. According to essentialism, properties have their powers essentially and the emergent causal laws that, on the non-essentialist account, partly or wholly determine the causal powers of emergent properties, are reducible to these powers (or some of them). Properties thus come with their causal powers 'built in', on the essentialist account, and in so far as the causal powers of emergent properties are

fundamental, this makes emergent properties fundamental as well. This puts pressure on the very idea of emergentism. If essentialism is true and emergent properties are grounded in physical properties and trans-ordinal laws (where trans-ordinal laws themselves may reduce to certain non-causal powers of physical properties), then their causal powers must be grounded in that base as well. There simply is no room for the idea that a base can ground the instantiations of certain properties without grounding their powers, if these properties have their powers essentially. This leaves the emergentist with a choice. On the one hand, she can give up the claim that emergent properties are grounded in physical properties and accept them as fundamental (which is what Howell assumes). If so, emergentism about mental properties would automatically falsify physicalism and emergentism about, for instance, chemical or biological properties would simply be forms of physicalism (though not forms of micro-physicalism). On the other hand, the emergentist can insist that emergent properties are non-fundamental properties grounded in physical properties (and trans-ordinal laws). But if so, it seems they cannot have fundamental causal powers since to ground a property, according to essentialism, is to ground its powers. This means that emergent properties fail to be novel in the sense claimed by the emergentist. To my mind, this makes emergentism a form of physicalism. The emergentist position would be that there are non-physical higher-level properties, yet these properties have no powers that are not grounded in powers bestowed by physical properties. And this is exactly what the physicalist claims as well.

2.9 Concluding Remarks on Physicalism and Emergentism

In the end then, whether dispositional essentialism is true or not, physicalism is incompatible with emergentism. If essentialism is false and properties have their powers contingent on the laws of nature obtaining, then emergent properties have causal powers not grounded in physical properties and physical laws. This is incompatible with physicalism. On the other hand, if essentialism is true, the traditional idea of emergentism is incoherent. Either emergent properties are fundamental and have fundamental causal powers, in which case emergentism about mental properties violate the NFM-constraint and will count as a non-physicalist position, and emergentism about chemical, biological, and other special scientific properties, will be forms of physicalism, though not forms of micro-

physicalism. Or emergent properties are not fundamental and have no novel causal powers, in which case emergentism should count as a form of physicalism.

2.10 Reductive and Non-reductive Physicalism

The standard distinction between reductive and non-reductive physicalism found in the literature goes something like this: Reductive physicalists claim that all properties, in particular mental properties, are identical to physical properties. Non-reductive physicalists, on the other hand, claim that some properties, in particular mental properties, are distinct from physical properties, though they are grounded in (or supervene upon, according to supervenience physicalists) physical properties. While common, this way of drawing the distinction between reductive and non-reductive physicalism obviously won't work given the terminology developed in the previous chapter. In fact, the phrase 'reductive physicalism' is a contradiction in terms if defined as the view that mental properties are physical. Physical properties are fundamental properties, so if mental properties are physical, then mental properties are fundamental – and this is incompatible with physicalism. Fortunately, turning to the literature on reductive physicalism, it becomes clear that self-proclaimed reductive physicalists are not actually trying to identify mental properties with *fundamental* physical properties. Rather, mental properties are usually claimed to be identical to complex neural properties. But such properties are arguably not fundamental, and as such, not strictly speaking physical. Now, physicalists will of course insist that neural properties are grounded in fundamental physical properties. But this simply makes reductive physicalism a form of non-reductive physicalism as it is defined above – on the reductive physicalist account, mental properties are non-physical properties that are grounded in physical properties.

2.11 Non-Physical Properties and Broadly Physical Properties

Seeing that mental properties are non-physical on both the reductive and non-reductive physicalist account, if there is a distinction to be drawn between the two positions, it must be found in the *kind* of non-physical properties they each take mental properties to be. In other words, we are looking to draw a distinction within the domain of non-physical properties. It should be a consequence of this distinction that paradigm examples of reductive and non-reductive physicalism differ on the

kind of properties they take mental properties to be. As a paradigm example of a reductive physicalist position we may take the identity theory, according to which mental properties are type-identical to neural properties of the brain. As a paradigm example of a non-reductive physicalist position we may take role functionalism, according to which mental properties are type-identical to functional properties, which in turn are realized by neural properties of the brain (in humans at least).

Some non-physical properties stand in a particular close relation to physical properties. These are the non-physical properties which can be 'built out of' physical properties. Such properties are what I'll call *broadly physical properties*. They are physical in the same sense that an ordinary material object, say a table, is a physical object: it is ultimately and entirely composed out of (fundamental) physical entities. Broadly physical properties are the kind of properties which requires the instantiation of *specific* physical properties to instantiate. Contrast this with the kind of non-physical properties which do not require the instantiation of any particular physical properties, though which may still be 'physicalistically acceptable' if, in the actual world, they are grounded in the physical.

Until now, I have used the notion of physical properties as a term for the fundamental properties posited by completed physics, and the notion of non-physical properties as a blanket-term for all properties not physical. To stay in line with the terminology used in the literature, however, it is now better to dub the properties posited by physics *fundamental physical properties*. *Physical properties*, in turn, I take to be properties that are either fundamental physical or broadly physical properties. *Non-physical properties* I take to be properties that are physicalistically acceptable, but neither fundamental physical nor broadly physical properties (obviously this excludes anti-physicalist properties from being non-physical properties).

2.12 Generating Broadly Physical Properties from Fundamental Physical Properties

Which tools does the physicalist have in her toolbox with which she can build broadly physical properties? One place to start is with the standard Boolean

operations of conjunction and disjunction.[63] In other words, broadly physical properties can be generated by conjoining and disjoining physical properties.[64] Doing so, however, would never take us outside the level of fundamental physical properties. If *P* and *Q* are fundamental physical properties of fundamental physical objects, then the conjunctive property *P and Q* and the disjunctive property *P or Q*, will also be properties of physical objects. This makes these properties unsuited as reduction-bases for mental properties. Mental properties aren't properties of fundamental objects on the reductive physicalist account, they are properties of complex biological organisms such as humans and other animals. The same point applies to other properties about which one might be a reductive physicalist. For instance, chemical properties cannot be generated by conjoining or disjoining fundamental physical properties, as chemical properties are (presumably) not properties of fundamental objects. So conjunction and disjunction cannot be the only tools at the physicalist's disposal, if she is to build the broadly physical properties which can serve as reduction-bases for higher-level properties.

What is needed is a relation that will allow the physicalist to move up levels (or down, if the fundamental properties are higher-level properties), in order to generate broadly physical properties. For material objects, moving up (or down) levels is a matter of composing (or decomposing) objects. For instance, we move from the subatomic to the atomic level by composing atoms out of subatomic particles. Such composition occurs when the parts are fused together mereologically. The relation of mereological fusion can be applied to properties as well, and we can talk of composite properties having other properties as proper parts. As an example, consider the structural property such as the property of *being an H2O-molecule*. This property is a higher-level property since it is a property of a non-fundamental compositionally complex object (a molecule). It has the properties *being a hydrogen-atom* and *being an oxygen-atom* as proper parts, in the sense that an

[63] I'm leaving out negation. Even if there are such properties as negative properties (which is controversial), it is unlikely that physical properties are closed under negation. For instance, an immaterial substance has the negative property of *not being a boson*, as well as countless other properties like it (*not being an electron*, *not being a proton*, etc), yet I take it we do not wish to say that an immaterial substance has countless physical properties. For the same reason, we should think that negations of fundamental physical properties need not generate broadly physical properties.

[64] Of course, this is not to say that any conjunction or disjunction of fundamental properties is a genuine property. The claim is just that if a conjunction or disjunction of fundamental physical properties form a genuine property, then this property is a broadly physical property.

object instantiating the property will have proper parts, two of which instantiate the property *being a hydrogen-atom*, and one which instantiates the property *being a hydrogen-atom*.[65] The suggestion is that, in addition to conjunction and disjunction, mereological fusion is a relation that generates broadly physical properties from fundamental physical properties. Of course, these relations do not exclude each other and can be applied together. For instance, a conjunction of disjunctions of physical properties will be broadly physical, a conjunction of structural properties with fundamental physical properties as parts will be broadly physical, a structural property with conjunctions of fundamental physical properties as parts will be broadly physical, and so on. On the other hand, non-physical properties are those properties that are grounded in fundamental physical properties, but which cannot be generated by applying the relations of conjunction, disjunction, and mereological fusion.

2.13 Applying the Distinction between Non-Physical and Broadly Physical Properties

Does this distinction between non-physical and broadly physical properties give the right result when applied to the paradigm cases of reductive and non-reductive physicalism? I think it does. The neural properties with which the identity theorist wishes to identify mental properties are presumably quite complex and will be properties of, if not the entire brain, then a significant part of it. It will be something like the property of having certain parts (neurons) with certain properties (*being a neuron*, *firing*, etc.), standing in certain relations. Presumably the property of *being a neuron* is itself a structural property (the property of having molecules with certain properties as parts, with those parts standing in certain relations), the property of *being a molecule* is a structural property (the property of having atoms with certain properties as parts, with those parts standing in certain relations), and so on, until we get the fundamental physical properties. In this way, the neural properties which serve as a reduction-base for mental properties according to the

[65] Of course, these parts must also instantiate certain relational properties for the structural property to be instantiated. I take it that it is one of the defining characteristics of structural properties that they require certain relations to obtain between their non-relational parts. If so, there are plausibly examples of composite properties that are not structural. For instance, the mass of a composite object may be construed as the property of having certain parts with certain masses. In this case, however, it is not clear that the parts need to instantiate any particular relational properties. I will discuss this example in further detail in the following chapter.

identity theorist can be entirely decomposed into fundamental physical properties. And in this sense, these neural properties *require* the instantiation of certain physical properties and relations - they are broadly physical properties. Contrast this with the claims made by the functionalist. According to the functionalist, structural neural properties realize mental properties, yet they are not identical to any one of them in particular. In fact, mental properties could be realized by anti-physicalist properties, for instance, fundamental mental properties. That is, the instantiation of the functional properties with which the functionalist wishes to identity mental properties, requires neither the instantiation of any particular physical properties, nor even the instantiation of any physical properties at all. Nevertheless, as it happens, the higher-level functional properties of our world are grounded in (and therefore supervene upon) the physical laws and properties of our world. They are non-physical properties.

This gives us our distinction between reductive and non-reductive physicalism. Reductive physicalism is the view that all non-physical properties, in particular mental properties, are identical with composite or structural properties that are ultimately and entirely composed of fundamental physical properties (and relations between them). Non-reductive physicalism, on the other hand, allows that some non-physical, in particular mental properties, are distinct from properties ultimately and entirely composed of fundamental physical properties, but nevertheless grounded in them.[66]

2.14 General Worries about Composite Properties

Admittedly, the discussion above skirts several issues about composite or structural properties. First, it might be that one thinks there are no structural properties because one thinks that the kind of composition involved in generating them is unintelligible, and that there is no other way of accounting for them (as David Lewis argues, at least if properties are universals[67]). Second, it might be that structural properties should be accounted for in a non-mereological fashion. For

[66] Of course, it is assumed here that physicalism is true. That is, neither broadly physical nor non-physical properties are subject to emergent causal laws, and have only powers that are ultimately grounded in the powers of fundamental physical properties. Also, there can be no anti-physicalist properties.
[67] Lewis [1986a]. For a reply see Armstrong [1986].

instance, it might be that structural properties are themselves simple, and that the relation between a structural property *S* and the relevant properties instantiated by the parts of the object that instantiates S is one of 'brute' necessitation, not composition (what Lewis calls a "magical" conception of structural properties). If so, it would be false that structural properties are composite properties with physical properties as parts, and structural properties wouldn't count as broadly physical, though of course, they may nevertheless count as non-physical if they are grounded in fundamental physical properties. In either case, the distinction between reductive and non-reductive physicalism that I am proposing won't work. This leaves two options: Either there is no distinction to be drawn at all or the distinction can be drawn in a different manner. The former would be surprising given that the use of the reductive/non-reductive distinction is prevalent in the literature. Therefore, until the latter is worked out in detail, I prefer to put aside the worries about composite structural properties and use them to help draw the distinction between broadly physical and non-physical properties.

2.15 Concluding Remarks

In this chapter, I have argued that if properties have their powers contingent on the laws of nature, emergentism can be distinguished from physicalism by way of the fundamental emergent causal laws that the former position is committed to. Such laws act only on non-physical properties and should be considered non-physical laws because of it, despite the fact that they are fundamental. They should therefore be excluded from the physicalist grounding base for non-physical properties. But if they are excluded, and emegentism is true, there would be causal powers at the non-physical level which are not grounded in the physicalist grounding base. And this contradicts physicalism. In the case of dispositional essentialism being true, emergentism fails to be a coherent position. If properties have their causal powers essentially, a grounding base cannot ground a property without grounding its powers. But emergentism is exactly the view that emergent properties are, firstly, grounded in physical properties and trans-ordinal laws, and, secondly, have fundamental causal powers. So whether dispositional essentialism is true or not, there is no risk of emergentism being counted as a form of physicalism.

Finally, I argued that the best way to understand the difference between reductive and non-reductive physicalism is via the notion of composite or structural properties. Reductionism about a certain group of higher-level properties is the view that these properties are identical to composite or structural properties ultimately and entirely composed of physical properties and relations, or any disjunction or conjunction thereof. Non-reductionism denies the identity. While properties in the disputed group are grounded in physical properties, they are not identical to any mereological fusion, disjunction, or conjunction of these properties.

Chapter 3

The Causal Exclusion Argument and the Drainage Problem

Abstract: In this chapter I begin by clearing up some terminological issues concerning properties, causal powers, and causation. Then I go on to set up the Exclusion problem. I claim that it is best seen as a paradox involving four independently motivated, yet jointly incompatible claims – the claims that there is mental causation, that the physical domain is causally closed, that there is no systematic overdetermination, and that mental properties are distinct from physical properties. This leads to the Exclusion argument, which is the argument that uses the first three claims to establish the denial of the fourth. I discuss to what extent this argument generalizes to higher-level properties and answer affirmatively, which leads to the problem of causal powers draining away if there is no fundamental bottom level (the so-called drainage problem). I argue that reduction of properties across levels is of no help since it comes with similar problems. Finally, I argue that to solve the drainage problem, it must be the case that causal powers reduce across levels. This, in turn, allows for a rejection of the claim that there is no systematic overdetermination.

3.1 Causes

In the previous chapter I talked rather freely of properties and powers. Before proceeding, some brief clarifications are needed.

The *relata* of causal relations I take to be particulars or tokens, rather than types. So for instance, we say such things as 'the football caused the window to break' (an object causing an event) or 'the football hitting the window caused it to break' (an event causing another event). Of course, objects and events can be types. For instance, football games all involve the same *type* of object (*football*) and window-breakings are often preceded by the same *type* of event (*window-hitting*). Yet, when a football hits a window and breaks it, I take it that it would be unnatural to think that what is (literally) causing the window to break is the object-type *football* or the event-type *window-hitting.* Surely, it's the *particular* football that happened to fly through the air and hit a window at a particular place and time that's the cause. Sometimes we say things that might suggest otherwise, for instance, when we say 'war causes misery' without having any particular war in mind. However, such a claim is plausibly interpreted as the claim that all particular wars cause misery. In

other words, by saying that war causes misery one is not necessarily committed to there being an event-type (*war*) that is literally causing misery.

It is equally unnatural to think of properties as causes. Say that one is a nominalist about properties and thinks that properties are sets of concrete objects (Resemblance Nominalism) or tropes (Trope Nominalism). It would be quite strange to insist that what is doing the causing in a case of causation is a *set* of particulars. Members of a set might be causes but the sets themselves just do not seem like the kind of entities that are apt for being causes. Positing universals as properties rather than sets does not make properties any better suited as causes. The problem is that the universals instantiated by the football that crashes through the window have multiple locations. They are located wherever they are instantiated (at least on an Aristotelian account of universals). So if properties can be causes and properties are universals, it might be that the very properties instantiated by the book on my shelf is at this very moment causing a window to break in Sydney. But this is not how we typically think of causes. Causes have specific locations. Whatever it is that is the proximate cause of the window breaking, it is *there*, in the close vicinity of the window and nowhere else.

3.2 Causal Powers

To have a causal power is to be a potential cause. For instance, a rock has the power to cause windows to break because it *can* cause windows to break, assuming certain conditions obtain. So if types and properties cannot be causes, it follows that they are not the kind of entities that *have* causal powers. What have causal powers are the entities that can be causes, i.e. particular objects and events, not object-types, event-types, or properties (in the following, when I talk of objects and events, I take these to be particulars unless otherwise stated). But if a property does not itself have causal powers, what exactly is the relationship between properties and causal powers? It seems obvious that there is a close relation in so far as we are inclined to think that what a thing can *do* depends on what it *is*. Following fairly standard terminology, the job of properties is to *bestow* or *confer* causal powers on objects and events (I will use the former term throughout). The bestowing relation that holds between a property and the powers of the particulars that instantiate it is arguably an example of the grounding relation, with properties grounding the

powers of particulars - particulars have their causal powers *in virtue of* the properties they instantiate.[68] This way of describing things is useful since it allow us to pinpoint the aspects of causes that are causally relevant. A football may be the cause of a window-breaking but not everything about the football is causally relevant. The colour, for instance, seems to have no causal relevance, whereas the mass and velocity of the football does. The latter properties bestow the power to cause window-breakings on the football, the former does not.[69]

3.3 Causation

What of causation itself? Anyone familiar with this topic knows that there is little consensus in the literature on how to understand causation. Since having a power to cause an effect is to be a potential or possible cause of that effect, what we might say is that causation involves the *manifestation* of a causal power. This manifestation requires certain conditions to obtain, which are exactly those conditions that upgrade a thing from being a merely potential cause of an effect (i.e. having a power to cause the effect), to being an actual cause of that effect. A property is causally efficacious with respect to the effect if the power manifested is a power bestowed by the property. I take this be a stipulation or platitude about the connection between causal powers and causation, or as I will sometimes refer to it, causal efficacy. For this reason, I also take it to be neutral on the metaphysics of powers and causation. To have a power is to be a potential cause and to manifest a power is to be a cause, *whatever causation amounts to*. If, for instance, one is attracted to a so-called production view of causation where causation involves (perhaps metaphorical) pushing and pulling, or 'omph', then one might also be attracted to a metaphysically robust understanding of powers and their manifestations. Objects and events have real powers, the manifestation of which

[68] I prefer to remain neutral on the metaphysics of property-instances, or property-tokens, though I do believe that much of the discussion to come will make no sense if instances of properties are concrete objects (as they would be on some nominalist account of properties). I will not argue for that here. Since properties require instantiation to bestow powers, some would prefer to say that powers are bestowed by property-instances. I have no objections to this way of talking.

[69] Sometimes the term 'property' refers to what is known as *abundant properties*. As the name suggests, abundant properties are plentiful and unlike *sparse properties*, which are the properties that ground the causal powers of individuals (perhaps among other things), abundant properties carry no real ontological commitment. Unless otherwise noted, when I talk of properties, I have in mind sparse properties. For more on the distinction between sparse and abundant properties, see Lewis [1986], p. 59-60.

involves this kind of pushing and pulling.[70] On the other hand, if one is attracted to a counterfactual account of causation, according to which C causing E is simply a matter of E being counterfactually dependent on C, this is likely to spill over to one's account of powers. On this account, C having a power to cause E might likewise reduce to the truth of certain counterfactuals, for instance the counterfactual 'were C to occur, E would occur and be counterfactually dependent on C'. I expect similar stories can be given for other accounts of causation.

Despite the reasons mentioned earlier for taking particulars to be causes and having causal powers, for reasons of simplicity, it will still be useful to talk as if properties (understood as types) are causes and have causal powers. In what follows, if I say that a property *P* has a power to cause effect E, this will be shorthand for saying that property *P* bestows a power to cause E on the objects and events that instantiate it. Likewise, if I say that a property *P* is a cause of effect E, this will be shorthand for saying that property *P* bestows a power to cause effect E on an object or event and that this power is manifested, or that the object or event is a cause of effect E in virtue of instantiating property *P*.

3.4 The Causal Exclusion Problem

The Causal Exclusion problem (or simply, the Exclusion problem) is best seen as a paradox involving four distinct claims that are independently motivated and *prima facie* plausible, yet jointly incompatible.

3.5 Mental Causation

First claim: Mental properties are sometimes causes, both of physical properties and other mental properties. As an example of mental causation, say that Jones instantiates an intentional property *P*, for instance by having the belief that it is sunny outside, and that this intentional property bestows on Jones the power to cause him to reach for his sunglasses (a broadly physical behavioural effect, or effect with broadly physical properties) and the power to cause him to believe that his lawn will dry out (an intentional effect, or effect with an intentional property). If these powers manifest, the intentional property *P* is a cause of him reaching for his

[70] This is not to say everyone endorsing a production view of causation is necessarily committed to a realist view of powers and their manifestations.

sunglasses and believing that his lawn will dry out. The former is a case of mental-to-physical causation, the latter a case of mental-to-mental causation. To deny that there is mental causation, and thereby deny that examples like the one just described ever occurs, is to embrace epiphenomenalism, the view that mental properties are never causally efficacious. This is a view that few contemporary philosophers share. As Kim says:

> […] epiphenomenalism strikes most of us as obviously wrong, if not incoherent; the idea that our thoughts, wants, and intentions might lack causal efficacy of any kind is deeply troubling, going as it does against everything we believe about ourselves as agents and cognizers.[71]

Thus, the following principle is heavily supported by what most of us takes as facts about human agency:

(Mental Causation): Mental properties sometimes cause other mental properties and other physical properties.

3.6 The Causal Closure Principle

Second claim: The physical domain is causally closed. This is motivated by empirical arguments. To deny it would be to accept that sciences dealing with physical properties, either fundamental or broadly physical, cannot give a complete causal account of physical effects and the explanatory success of such sciences indicate otherwise.[72] Sometimes this is known as the claim that physics is complete – physics would never have to move outside the physical domain to give causal explanations of physical effects. Often the causal closure of the physical is expressed as the following principle: Every physical effect has a sufficient physical cause. Note though that formulated in this simple fashion, the principle of causal closure is too weak to capture the idea behind it, as it does not rule out that a physical effect is the result of a chain which includes both non-physical, or even anti-physical properties, and physical properties. For instance, say that physical property *P* causes non-physical property *N*, which causes physical effect E, and that

[71] Kim [2005], p. 70.
[72] For some examples of this evidence, see Papineau [2001].

E has no other causes. This is clearly a counter-example to the claim that the physical domain is causally closed since E cannot be completely causally accounted for, without moving outside the domain of physical properties (to *N*). However, assuming that causation is transitive (a standard assumption in the philosophy of causation), it is not a counterexample to the principle of causal closure as it is expressed above, since E *does* has a sufficient physical cause.[73] A solution is to require that for any time *t* and any physical effect E, if E has a cause at *t*, E has a sufficient physical cause at *t*.[74] The example just given would be incompatible with the truth of this principle, since there is a time *t* (the time at which *N* occurs), where E has a cause but no sufficient physical cause.

While an improvement over the initial simple formulation of Closure, I still think this amended version of the closure principle is too weak to fully capture what philosophers, in particular physicalists, have in mind, when they look to endorse the claim that the physical domain is causally closed. Consider the following example: Assume that the world is ripe with non-physical properties sufficiently causing physical properties. As a matter of extraordinary coincidence, for every one of these physical effects caused by a non-physical property, there is an independent physical property which happens to also sufficiently cause the effect. So we do not have a counterexample to the proposed formulation of the causal closure of the physical domain since every physical effect *does* have a sufficient physical cause. Yet, some might argue that if the physical domain is causally closed, it cannot be by pure accident. Or rather, what physicalists mean when they say that the physical domain is causally closed or that physics is causally complete is stronger than any principle expressing this, which is compatible with causal closure by pure accident. In any case, the empirical support for closure certainly seems to justify a stronger claim. Putting radical sceptical worries aside, the fact that, *in the past*, for any time *t* and any physical effect E, E had a sufficient physical cause at *t* (if it had a cause at *t*), justifies us in inferring that, *in the future*, for any time *t* and any physical effect E, E will have a sufficient physical cause *t* (if it has a cause at *t*). But this is just to say that the empirical support for closure justifies the claim that it is *guaranteed* that,

[73] See Lowe [2000] for more on this point.
[74] This is how Kim formulates the closure principle in one of his recent writing on the Exclusion problem. See Kim [2005], chapter 2.

for any time *t* and any physical effect E, if E has a cause at *t*, E has a sufficient physical cause at *t*. And in turn, this is to say that it is no accident that the closure principle holds true.

The obvious way of incorporating into the closure principle that it is non-accidentally true, is by giving it some modal strength. Given the causal laws of our world, it is *necessary* that, for any time *t* and any physical effect E, if E has a cause at *t*, E has a sufficient physical cause at *t*. In other words, the closure principle supervenes on the causal laws of our world. This rules out that a physical property *P* can be sufficiently and accidentally causing a physical effect, that would have been sufficiently caused by non-physical property *N* in any case, since we would not have to change any causal laws in order to make non-physical property *N* the sole sufficient cause of E. It would be enough to simply remove physical property *P*.[75] We might worry that this makes the formulation of closure too strong, since closure would then be incompatible with the existence of possible worlds with extra causal laws that overrule some of the laws of our world. For instance, say that it's a law L of our world that *P*'s sufficiently causes *E*'s, where *P* and *E* are physical properties, and that nothing is interfering with this law in our world. But now imagine a possible world *w* that is a duplicate of our world with respect to causal laws but with an extra law L* that interferes with L. L* does this by ensuring that whenever *P* is instantiated, non-physical property *N* both causes E and prevents *P* from causing E. Intuitively, physicalists should allow, or at least be neutral about, the possibility of such a world. But the closure principle, as stated, rules it out. Note how this problem is similar to the problem physicalists faced when defining physicalism in terms of standard supervenience. The problem was that physicalists were in danger of ruling out the possibility of worlds that are identical to the actual world with respect to physical properties and laws of nature, but which contain extra anti-physical properties. The solution was to commit physicalists only to minimal supervenience of the non-physical on the physical. This makes the possibility of worlds with extra anti-physical properties irrelevant from the perspective of physicalism. And the same move can be made here - the closure principle *minimally* supervenes on the causal laws of our world. This makes the

[75] Of course, this does not rule out that a physical and a non-physical property can each be sufficient causes of the same effect.

possibility of worlds with extra causal laws interfering with the laws of our world irrelevant. The principle that best captures the idea of the causal closure of the physical is therefore the following:

(Closure): In any world that is a minimal duplicate of our world with respect to causal laws, for any time *t*, and any physical property E, if E has a cause at *t*, E has a sufficient physical cause at *t*.

In the following, however, I will often abbreviate this to the simpler claim that physical effects always have sufficient physical causes

3.7 The Causal Exclusion Principle

Third claim: If an effect E has a sufficient cause C at time *t*, then all other properties distinct from C are excluded from being a cause of E at *t*. This is known as the exclusion principle. Before trying to motivate it, let us consider what is involved in denying it.

To deny the exclusion principle would be to accept that the world contains *causal overdetermination*. Causal overdetermination occurs when an effect has two (or more) sufficient causes and where these sufficient causes are not themselves related as cause and effect. A standard example of this from the literature involves two snipers each firing a riffle at a person, call him Target. As it happens, the two bullets fired hits and kills Target at exactly the same time. Also, each bullet would alone have been sufficient for killing Target. In this case, it seems that we have two distinct objects (or events), each a sufficient cause of Target's death. Are such cases impossible? Perhaps not. And as a result, perhaps the Exclusion principle as stated is too strong. What is striking about the sniper case, however, is that it involves an extraordinary coincidence. Even with the added assumption that both snipers are highly skilled and specifically instructed to try to shoot at the same time, it would still require a substantial amount of luck on the snipers part, were their bullets to hit at *exactly* the same time. This suggests that the kind of overdetermination exemplified by the sniper-case, in so far as it occurs at all, occurs only on rare occasions. We can take this conclusion on board and appropriately amend the exclusion principle as originally stated:

(Exclusion): *In general*, if an effect E has a sufficient cause C at time *t*, then all other properties numerically distinct from C are excluded from being a cause of E at *t*.

This version of the exclusion principle allows for the possibility of sniper-cases. What the principle *does* prohibit is *systematic overdetermination*. If certain effects are overdetermined in a systematic manner, overdetermination would be frequent and commonplace (at least assuming that the effects in question are commonplace). And this is exactly what the exclusion principle rules out.

Why believe this principle? First, there are methodological considerations, such as arguments from simplicity or Ockham's Razor. If an effect has a sufficient cause, it seems redundant to posit *additional* causes. And an ontology that allows for (or even worse, embraces) such additional causes seems both unparsimonious and *ad hoc*. This is also reflected in our practice concerning causes and effects. For instance, if the police find the culprit that caused the death of the victim, they typically close the investigation. They do not continue to look for additional causes of the victim's death. Second, if overdetermination indeed requires extraordinary circumstances, then it seems incredible that some effects could be overdetermined as a rule, rather than an exception. Third, and this problem is rarely discussed in this context in the literature, it seems that only certain kinds of effects can be genuinely overdetermined at all. What is required is that the effect is of a kind that fails to accumulate. I will call such effects *non-cumulative effects*. What I mean by this is that, unlike cumulative effects, they do not come in quantities. An example will illustrate the difference I have in mind. Consider again Target that gets hit by two bullets simultaneously. The two bullets cause many effects of course, in addition to the death of Target, one of which is the exertion of a certain amount of force F on Target's body. The exertion of force F, however, is not overdetermined, since neither bullet is a sufficient cause of it. If only one of the bullets had hit target, the total amount of force exerted on Target's body would have been less than F. In fact, it seems that the total amount of force exerted could not possibly be overdetermined. An additional object exerting force on Target's body would simply add to the total amount of force exerted, preventing *this* effect (the total amount of

force exerted) from being overdetermined. Exertion of force is the kind of effect that accumulates given additional causes. Contrast this with the effect of Target dying. An additional bullet fired at Target would not cause him to die an additional time - while you can die in a variety of ways, you cannot die more than once (barring resurrection). The effect thus fails to accumulate given additional causes and this is what allows for *this* effect (the dying) to have more than a single sufficient cause.[76] If only non-cumulative effects can be overdetermined, this already puts a natural restriction on the occurrence of causal overdetermination. More importantly, anyone wishing to deny the exclusion principle in order to allow for the systematic overdetermination of cumulative effects needs to explain how such overdetermination is even possible. As far as solving the exclusion problem by rejecting the exclusion principle, this is relevant since at least some effects of mental properties are cumulative. I will say more about such strategies later.

3.8 Distinctness

Fourth and final claim: Mental properties are numerically distinct from physical properties. One powerful motivation for this claim comes in the form of the argument from multiple realization. In one of the early formulations of the argument, it is presented as a challenge to the reductive physicalist. As Hilary Putnam writes about the prospect of locating a broadly physical property which can serve as the reduction base for pain:

[76] A couple of points: First, there is a similar way of dealing with sniper cases, and other cases involving non-cumulative effects, according to which they are not cases of genuine overdetermination. For instance, the death caused by both bullets might be ever so slightly different from the death that would have been caused by one of the bullets on its own. In this case, the very specific death that occurred was not overdetermined by the two bullets, since neither of the bullets on its own was a sufficient cause of *that* death. Of course, if sound, this line of reasoning merely supports that the original stronger version of the exclusion principle is motivated as well. Second, it might be argued that cases of force being exerted, and other cases involving cumulative effects, do involve *some* form of overdetermination. Consider again the two bullets that each exerts part of a force on Target's body, the total amount of which is F. Now, it may be true that neither of the bullets on its own is a sufficient cause of exerting force F. However, each bullet does cause the effect of some indeterminate force being exerted. And *this* effect is also caused by both bullets in conjunction. So if we consider effects at a relatively determinable level, even cumulative effects, such as exertion of force, become overdetermined. Personally, I find it implausible to claim that the determinable event caused by both bullets in conjunction is identical, at the token-level, to the determinable event caused by each bullets individually (though they are all tokens of the same determinable event-type of exerting some indeterminate force). So I do not think that the two bullets overdetermine even their relatively determinable effects. However, whether the argument succeeds or not is irrelevant for present purposes. If it does succeed, the exclusion principle could simply be rewritten so as to be specifically about causes of relatively determinate effects.

> He [*the reductive physicalist*] has to specify a physical-chemical state such that any organism (not just a mammal) is in pain if and only if (a) it possesses a brain of a suitable physical-chemical structure; and (b) its brain is in that physical-chemical state. […] It is not altogether impossible that such a state will be found. […] But it is certainly an ambitious hypothesis.[77]

However, even if it should turn out that, as a matter of fact, all creatures *do* share a relevant broadly physical property *P* when instantiating mental property *M*, the argument can be strengthened to prevent the identification of *P* with *M* by invoking possible worlds with creatures instantiating *M* without *P*. If such worlds are possible, *M* cannot be identical to *P*. And, the argument goes, surely such worlds *are* possible. Hence, claiming that mental and physical properties are distinct is motivated.

3.9 The Exclusion Argument

Despite each of the four statements being independently motivated, any three of them can be used to establish the denial of the fourth. This is particularly worrisome to non-reductive physicalists, since they tend to accept all four. Distinctness is obviously a constitutive claim of the position, and Closure is universally accepted by all physicalists. Mental Causation and Exclusion are plausible claims in general, independently of whether or not one is a physicalist. In the literature on mental causation, this has lead to the *Exclusion Argument* against non-reductive physicalism, which I take to be the argument that uses Mental Causation, Closure, and Exclusion to show that Distinctness is false. That is, Mental Causation, Closure, and Exclusion together establish that mental properties are identical to physical properties. This argument has been developed and defended in detail by Kim in various writings:[78] From Closure we know that all physical effects, such as behavioural effects, each have a sufficient physical cause. The Exclusion principle tells us that, in general, if an effect E has a sufficient cause C, then no property distinct C from is also a cause of E. But Mental Causation tells us that mental properties are sometimes causes of physical properties. It therefore follows that mental properties must be identical to physical properties. So Distinctness, and by implication non-reductive physicalism, is false.

---

[77] Putnam [1967], p. 228.
[78] See in particular Kim [1998] and Kim [2005]. See Papineau [2001] and [2002] for a similar version of the argument.

3.10 Tidying up the Argument

There are a couple of obvious loop-holes in the way I have set up the Exclusion argument that should be closed straight away. First, the non-reductive physicalist cannot plausibly exploit the fact that the Exclusion principle allows for special or rare cases of overdetermination by claiming that cases of mental causation are such cases. Mental causation is a commonplace phenomenon and as such, does not exhibit the rarity of the cases that the Exclusion principle is meant to allow. Another problem is that mental and physical properties are not independent in the way the overdetermining causes in standard cases of causal overdetermination are. This is due to the fact that non-reductive physicalist is not likely to construe the co-instantiation of a given mental property with the physical properties that are competing with it to cause the relevant physical effect as an extraordinary coincidence. Rather, the instantiation of the mental property is determined, in part at least, by those very physical properties. But then the resulting overdetermination of physical behavioural effects by physical and mental properties is systematic. And this is exactly what the exclusion principle prohibits. Some strategies for dealing with the exclusion problem attempt to turn the table on the exclusion principle by exploiting the very fact that mental causation involves systematic overdetermination. Roughly, the idea is that if overdetermination is problematic because it involves extraordinary coincidences, then psycho-physical overdetermination is unproblematic since it is no coincidence that the overdetermining causes are co-instantiated.[79] While I am sympathetic to this response, it is important to note that this is insufficient to reject the exclusion principle given the motivations I presented for it. Non-coincidental overdetermination may explain why it is not so incredible after all that some effects are systematically overdetermined. However, it fails to explain why such overdetermination does not involve redundant causation or why accepting such redundant causation is not unparsimonious and *ad hoc*. If anything, insisting that the overdetermination is no coincidence makes such views seem even more unparsimonious and *ad hoc*. Worse, if some effects of mental causes are cumulative, then overdetermination of these effects, coincidental or not, seems to be impossible. And for at least some effects of mental causes, this is plausibly the case.

[79] See Bennett [2003] and Kallestrup [2006].

For instance, say that Jones instantiates the property *being in pain* because he accidentally put his hand on a hot stove. This property has powers to cause certain types of effect when instantiated in individuals like Jones, including screaming. Now, if this property has a physical base *P* that, in the circumstances, is a sufficient cause of this type of effect, then we might wonder why the pain and its physical base do not each cause a distinct token of this effect and why these token-effects do not accumulate. If an instance of pain causes Jones to scream at a certain volume, why doesn't the physical base for this mental property cause Jones to scream even louder? After all, the effects of pains typically accumulate in this fashion (putting your entire hand, rather than just a small part of it, on a hot stove will typically cause a greater scream). To accept that the effects do accumulate is not an option of course, since this is to accept that the total effect is not sufficiently caused by the physical property and this violates the closure principle. Merely saying that psychophysical overdetermination is non-coincidental does nothing to solve this puzzle.

Second, the non-reductive physicalist may object that the Exclusion argument does not show that *all* mental properties are identical to physical properties, just that the mental properties that cause *physical* effects are identical to physical properties. Hence, there might still be irreducible mental properties that cause only other irreducible mental properties (which themselves have only powers to cause other irreducible mental properties, and so on). Putting aside the fact that it is not clear which mental properties would have such causal features, there are further reasons to think that such a move will not succeed. Kim has argued that, on a non-reductive physicalist account, cases of mental to mental causation will invariably involve mental to physical causation, since causing a mental property will invariably involve causing its physical grounding base.[80] Kim, in effect, establishes this by way of an extended Exclusion principle which applies to causal and non-causal relations alike.[81] The argument proceeds as follows: Say that mental property *M* putatively causes mental property *M**. If non-reductive physicalism is true, *M** will have a physical grounding base *P** which metaphysically determines *M**. But now,

[80] Kim [2005], pp. 36-41. Kim talks of supervenience bases rather than grounding bases but it is clear that he takes the supervenience relation to express some form of metaphysical determination.
[81] Kim introduces this principle in his [2005], p. 17.

Kim asks, if *M** is grounded in *P**, it seems redundant to claim that, in addition, *M* is a cause of *M**. After all, *P** is there and is guaranteeing the instantiation of *M**, regardless of whether or not *M* is there. Thus, *P** excludes *M* from being a cause of *M**. The exception, Kim notes, is if *M* is itself a cause of physical grounding base *P**. *M* would then be causing *M** by causing the grounding base for *M**. But since this grounding base is physical, this means that mental to mental causation involves mental to physical causation. And as the Exclusion argument shows, this is only possible if mental properties are identical to physical properties. So the non-reductive physicalist cannot even settle for the moderate view that there are irreducible mental properties which cause only other irreducible mental properties.

3.11 Solving the Exclusion Problem

Of course, an attempt to solve the Exclusion Problem by rejecting one of the four inconsistent claims should preferably be supplemented by further arguments, either to the effect that the motivation for the rejected claim can be discarded, or to the effect that they motivate a different claim than we thought they did. An example of this strategy can be found in Kim's earlier writings on the Exclusion problem. Kim accepts that Putnam's challenge of finding a single broadly physical property which all organisms in pain share cannot be met. But rather than inferring from this that pain is distinct from any broadly physical property, he argues that such generic mental properties should be eliminated and replaced with what he calls structure- or species-specific mental properties. In other words, there is no mental property *pain* which can be shared by organisms with very different physiological set-ups. Instead there might be *pain-in-humans*, which can be shared by humans, provided that the neural base for pain in humans is sufficiently homogenous.[82] Kim's strategy is thereby partly eliminativist, partly reductionist. He is an eliminativist about mental properties whose reduction-base consists of heterogonous broadly physical properties. But he is a reductionist and realist about species-specific mental properties with a single neural physical realizer (or set of homogenous neural realizers). Such properties are reducible to their neural base. In addition, he allows

[82] David Lewis proposes a similar strategy in his [1980].

that there are generic mental *concepts* such as the concept of pain *simpliciter*, though this concept fails to pick out a genuine property.[83]

Whether or not one accepts Kim's proposed strategy for dealing with Distinctness and its motivation, it is in any case what I will call a *dialectically satisfactory approach* to the Exclusion Problem. By this I mean that Kim attempts to deal with the motivation for the claim he rejects. Putative cases of organisms with vastly different physiological set-ups sharing a mental property are really cases of organisms having different species-specific mental properties, which fall under the same generic mental concept. Hence, the proposed evidence for Distinctness fails to adequately support the statement. By contrast, an unsatisfactory approach to the problem would be one which settles on three of the claims to establish the denial of the fourth, *while not addressing the motivations for the rejected premise*. A paradox is not satisfactory dealt with by simply rejecting one of the inconsistent claims – we also want to know *why* it is that a claim which seemed plausible nevertheless turned out to be false or how the evidence for the rejected claim is actually evidence for a different claim which generates no paradox. I take it to be a constraint on any proposed solution to the Exclusion Problem which involves a rejection of one of the claims, that it is dialectically satisfactory in this way.

3.12 The Generalization Problem

Before moving on to a discussion of positions which attempt to solve the Exclusion Problem by rejecting one of the claims constituting a paradox, let me first discuss an attempt to approach the issue in a different manner. We might look at this as a sort of meta-objection to the Exclusion problem, since it can be seen as an attempt to reject the problem, without targeting any particular claim. First part of the objection is this: If the Exclusion argument is successful against non-physical properties such as mental properties, it can be applied to other higher level properties, such as chemical and biological properties. This is known as the Generalization problem since the problem of mental causation threatens to generalize to all higher level properties. It is sometimes taken to constitute a

[83] See Kim [1992a].

reductio against the Exclusion argument and the Exclusion problem. We find this line of thought in several writings. For instance, Lynne Rudder Baker tells us that:

> […] not only has the problem of mental causation proved to be intractable but even worse, the same reasoning that leads to scepticism about mental causation also leads to scepticism about almost all supposed 'upper-level' causation, and hence to scepticism about explanations that mention 'upper-level' properties, including explanations offered by the special sciences […].[84]

This is echoed by Ned Block who says in response to the Exclusion argument that: 'it is hard to believe that there is no mental causation, no physiological causation, no molecular causation, no atomic causation but only bottom level physical causation.'[85] These, however, are peculiar worries, in so far as the Exclusion argument has as a premise that there *is* mental causation. Likewise, applying the Exclusion problem to higher level properties in general would presumably involve a premise to the effect that there *is* higher level causation. If the Exclusion argument shows anything about higher-level properties in general, it is that higher level properties are *reducible* to physical properties, since this is the conclusion of the argument with respect to mental properties. But for some higher level properties, this is uncontroversial. We have already seen how a property of chemistry such as *being an H2O-molecule* is broadly physical, since it is ultimately decomposable into fundamental physical properties. Of course, not all higher-level properties are like that. For instance, the property of *being metallic* is arguably a functional chemical property, defined by its causal role of conducting electricity, and with many distinct realizers (*being gold, being iron, being copper*, and so on). There's a real issue over whether or not the higher level properties that are not obviously broadly physical must be reduced in order to secure their causal powers, just as there is an issue over whether or not mental properties must be reduced in order to secure their causal powers. What is not clear, however, is that this is enough to establish a *reductio* against the Exclusion argument. After all, reduction in the special sciences is a live option.

---

[84] Baker [1993], 94.
[85] Block [2003], p. 5. Other authors have argued the same point. See Bontly [2002] and Noordhof [1999].

3.13 The Drainage Problem

Let us therefore put the initial formulation of the generalization problem aside and instead focus on Block's claim that the Exclusion Argument has as a consequence that there is only causation at the 'bottom level'. I take it that by 'bottom level', Block has in mind the fundamental physical level. The idea is this: For any property at a higher level, there will be a property at the level below which is causally excluding it. Repeat this step down through the levels and it will turn out that the only properties not causally excluded are the properties at the lowest level. But now imagine that there is no bottom level, which, as Block argues, is a real possibility. In other words, broadly physical objects and properties might decompose endlessly such that, for every level, there is a level below it. In this case, no properties would be causes since all properties are relatively higher level properties. This is known as the Drainage problem, since all causal powers threaten to drain away into a bottomless hole. But if it is an open question whether there is a bottom level, and the Exclusion argument shows that there is only causation at the bottom level, this means this Exclusion argument leaves it an open question whether there is any causation *at all*. And *this* is surely a *reductio* against the Exclusion argument.

3.14 Kim on the Drainage Problem: Levels versus Orders

In earlier writings, Kim has responded to the Drainage problem by stressing a distinction between orders and levels. Recall, levels are individuated mereologically. Molecules are at a lower level than cells because molecules compose cells, cells are at a lower level than animals because cells compose animals, etc. A property *P* is a higher level property relative to property *Q*, when the object instantiating *P* is mereologically complex relative to the object instantiating *Q*. So for instance, *being a molecule* is a lower level property relative to *being a cell* and a higher level property relative to *being an atom*. In contrast to this, a higher *order* property is a second order property – it is the property of having some first order property. Sometimes dispositional properties are construed in this manner. For instance, if a set of different molecular structures are the structures that are causing objects to break when struck, then being fragile might be defined as the second order property of having one of these first order molecular structures. From this it follows that the relation between lower and higher order properties is an *intra-level* relation since a property *P* and the property of having *P* will always be

properties of the same object. The object that has one of the molecular structures in the relevant set, say a glass, will be the very same object that has the second order property *being fragile*. Why is this important? It is important because the Exclusion principle only kicks in when an effect has two (or more) sufficient causes. And as Kim goes on to argue, it is generally not the case that there are properties at different levels competing to cause the same effect. For instance, molecules cause effects that no individual atom cause, atoms cause effects that no sub-atomic particle causes, and so on.[86] In the case of higher-order properties, on the other hand, a first and a second order property are in direct competition to cause the relevant effect. If a certain molecular structural property *S* realizes fragility in virtue of having the causal power to cause objects to break when struck, then it seems that *being fragile* is causally excluded from being a cause of *that* effect, unless *being fragile* reduces to *S*. According to Kim then, worries about causal exclusion will typically be worries about exclusion within a level, rather than across levels. Likewise, reduction as a cure for these worries will be intra-level reduction. For instance, the relevant reduction-base for a mental property will presumably be a complex neural property, both of which are properties of the same person or cognizer. By stressing this distinction between levels and orders, Kim has supposedly blocked the causal drainage from higher to lower levels, while leaving the force of the argument intact with respect to mental properties and other second order properties.

3.15 Composite Overdetermination

It is not clear, however, that this response by Kim effectively deals with the drainage problem. Granted, as Kim says, higher level objects often have causal powers in virtue of their properties that no lower level object has. For instance, a higher level object can have powers in virtue of its mass that none of its parts has. Hence, there is no singular property at the lower level that is threatening to causally exclude this higher level property and by implication, no immediate threat of causal drainage across levels. But why think that causal competition must be between two *singular* properties, one higher level and one lower level property? Perhaps a higher level property can be causally excluded by a *plurality* of lower level properties. For

[86] Kim [1997], p. 292.

instance, say that an object has the power to exert downward force F on a surface in virtue of its mass, and that each of its two parts has the power to exert a downward force of half an F. While it is true that none of the properties of the individual parts are competing with the mass of the whole to exert F, it might be argued that the masses of the two parts taken collectively *are* competing with the mass of the whole to cause that effect. In other words, we are facing systematic composite overdetermination – overdetermination by a whole and the parts that compose it. And if the mass of each of the parts are collectively exerting force F on the surface, then it seems that it would be redundant to claim that the mass of the whole is also exerting F on the surface. In this case, it certainly looks as if the causal powers of the higher level property drain to a level below.

For now, let's grant the proponent of the original Exclusion argument that the argument successfully shows that all non-physical properties (with causal powers at least), such as mental properties, must reduce to physical properties. In other words, the Exclusion argument shows that the world contains nothing but physical properties, some of which are higher level composite properties with fundamental physical properties as parts. We can now formulate a new version of the Exclusion problem in terms of higher level composite properties, based on the following four claims which appear to be mutually inconsistent.

I: Higher level composite properties sometimes cause effects (call this claim *Higher Level Causation*, or HLC for short).

II: Those effects are also sufficiently caused by the sum of lower level properties that compose these higher level properties (call this claim *Lower Level Causation*, or LLC for short).

III: In general, if an effect E has a sufficient cause C at time *t*, then all other properties distinct from C are excluded from being a cause of E at *t* (Exclusion).

IV: Higher level properties are distinct from the sum of lower level properties that compose them.

Note that, unlike the original Exclusion problem where epiphenomenalism about mental properties is an option, a rejection of HLC, the claim that higher level

properties sometimes cause effects, is blocked off by the drainage problem. The view that there is no bottom level is the view that all properties are composite since any property will be decomposable into further parts. So if composite properties are epiphenomenal, then there is no causation anywhere and this is not an option.

Note also that the Closure principle in the original formulation of the Exclusion problem is replaced by LLC, the claim that a sum or plurality of lower level properties causes the effect also caused by the higher level property they compose. In effect, LLC plays the role that the Closure principle plays in the original formulation of the Exclusion problem. But while the Closure principle finds empirical support in the explanatory success of natural science, it is not clear that the same support extends to LLC. After all, the issue here is not causal competition between a physical and a non-physical property but between some lower level physical properties and another physical property at a higher level. Nevertheless, LLC can be empirically motivated in a different way. Consider mass and the law of gravitation. I take it that this law is empirically supported as a universal law. As such, the law of gravitation does not discriminate between objects that are parts of a whole and those that are not. Say that an object O has a mass of 1 and that the law of gravitation dictates that the object exerts a downward force F when placed on a surface. Now imagine a situation where O is a proper part of another object O*. If the mass of O* is to causally exclude the mass of O from exerting a downward force, this would mean that when O is a proper part of O* and placed on a surface, the law of gravitation applies only to the mass of O*, not to the mass of O. In other words, the relation that O enters into when it becomes part of O* somehow trumps the law of gravitation - before O becomes a proper part of O*, its mass is subject to the law of gravitation. And then when O becomes a proper part of O*, this law no longer applies to O. In this case, composition would violate one of the basic laws of nature. But this sounds incredible and it is doubtful that anyone would want to defend such a claim. Surely an object is subject to the law of gravitation, whether or not it is a proper part of another object. Further, the mass of every non-overlapping proper part of O* each exert a downward force according to this law, the total of which is exactly the amount of force exerted by the mass of the entire object. So the claim that the lower level masses are sufficiently causing the exertion of the downward force also caused by the higher level mass they compose is empirically

motivated, at least in this specific case. Of course, it does not follow from this one example that lower level properties *always* cause the effects also caused by the higher level properties they compose. I will return to this issue shortly and also argue that as far as the drainage problem is concerned it does not matter if LLC is true for all composite properties.

In any case, if one endorses a rejection of the claim that mental and physical properties are distinct, in order to solve the Exclusion problem for mental properties, it is natural to assume that a rejection of the analogous claim that composite properties are distinct from the sum of their parts would be the preferred solution to the Exclusion problem for higher level properties. In other words, the way to avoid the causal competition between composite properties and their parts is therefore to reduce composite properties to their parts – composition is identity.[87] If higher level properties are identical to lower level properties, then there cannot be any causal competition between them. This view comes with problems of its own. For one thing, it requires that identity can hold one-many, a notion some philosophers object to, since it holds between a singular higher level property and a plurality of its parts.[88] Another worry, which is more directly relevant for Kim, is that the levels view of reality is ruled out if composition is identity. If higher level properties just *are* lower level properties, then there is no genuine metaphysical division of reality into distinct mereological levels. Rather, all supposedly higher and lower levels would collapse into one level, which would encompass all of reality. However, recall that Kim's initial defence of the Exclusion argument against the drainage problem was based on higher level objects having causal powers in virtue of their properties which no lower level object has. Endorsing that composition is identity in order to avoid causal competition between properties and the sum of their parts will therefore invalidate this defence. On the other hand, the higher level defence is problematic, as we have seen, since it illegitimately assumes that causal competition is always between two singular properties or objects, whereas, in fact, there seems to be nothing wrong with it being between, on the one

[87] A weaker view would be one according to which composition isn't identity but merely analogous to it. This is Lewis's view in his [1991]. It is not clear, however, that anything short of genuine identity between higher and lower level objects will block the application of the Exclusion principle.
[88] See Van Inwagen [1994].

hand, a singular property or object, and on the other, a plurality of properties or objects.

3.16 Kim on the Drainage Problem: Inter-level Reduction

So far, we have seen how the Exclusion problem can be extended to apply not only to mental properties (and other non-physical properties) but also to broadly physical higher level properties, and how Kim's distinction between levels and orders cannot be exploited in a manner that will help to dissolve this problem. In some of Kim's most recent writings on the Exclusion argument, he seems to realize this and his solution to the drainage problem based on higher level irreducible properties appears to have been replaced by a strategy that advocates reduction across levels.[89] If so, this would make Kim's response to the Exclusion problem for higher level properties the same as his response to the Exclusion problem for mental properties. Using water as an example, Kim starts out by noting that H2O is the micro-structure of water at the atomic level. This fact is what justifies the identity between *being water* and *being H2O*. But H2O itself has a micro-structure, call it R, at a level below the atomic level. This will be at the level of sub-atomic particles. So *being water* is also identical to *being R*. R has a micro-structure, call it S, at the level below the sub-atomic level. This will be, say, at the level of quarks. So *being water* is also identical to *being S*. These steps can repeated until we reach a fundamental or bottom level or forever if there is no fundamental level. This gives us a long, potentially infinite, line of identities:

Being water = being H2O = being R = being S…

Kim then goes on to proclaim: '*Voilà!* These are the identities we need to stop the drainage.' It is important, however, that for the identities to do that job, the proposed solution must be interpreted in the right way. In particular, the identities must be between properties at *different* levels and not just between same-level properties that are specified in terms of their increasingly smaller parts and their properties. To bring out this difference, consider how a reduction of a material object might proceed. Say that my dining table, call it Table, is composed of two

[89] Kim [2005], pp. 68-69. For a discussion of how Kim's response to the drainage problem has changed over the years, see Moore 2010.

parts; a top, call it Top, and a bottom, call it Bottom. Top is itself composed of four pieces of wood, W1-W4, and Bottom composed of four legs, L1-L4. With this in mind, we might propose the following line of identities:

> Table = the object composed of Top and Bottom = the object composed of W1-W4 and L1-L4.

This line of identities can be extended to include the object composed of the parts of W1-W4 and L1-L4, the object composed of the parts of these parts, and so on. The line would either stop at the bottom level or go on forever, if there is no bottom level. If composition is not identity, the reduction here is entirely within a level since Table is not reduced to any objects at a lower level. Rather, Table (the higher-level object) is simply specified in terms of increasingly smaller parts. Keeping in mind the Levels view, where we increase mereological complexity by moving *up*, we might say that the reduction proposed here is horizontal. We do not move up or down levels, merely sideways within a level. So if one is worried that Table is causally excluded by Top and Bottom, identifying Table with the object *composed* of Top and Bottom is of no help. However, if composition *is* identity, we can replace the proposed line of identities with the following:

> Table = Top and Bottom = W1-W4 and L1-L4.

Now Table has been reduced to objects at a level below it and the reduction is vertical, rather than horizontal. And now there can no longer be causal competition between, on the one hand Table, and, on the other hand Top and Bottom, or W1-W4 and L1-L4. Table just *is* those lower level objects.

3.17 Causal Drainage and Inter-level Reduction of Properties

Reduction of properties can be either intra- or inter-level in the same way. For instance, say that the following is true: The property *being an H2O molecule* is identical to the property of having three atomic parts, two of which has the property *being a hydrogen atom* and one which has the property *being an oxygen atom* (of course, the parts have to instantiate certain relational properties as well, a complication I will set aside for the sake of simplicity in this example). The

property *being a hydrogen atom* is identical to the property of having three sub-atomic parts, two of which has the property *being an electron* and one which has the property *being a proton*. And the property *being an oxygen atom* is the property of having twenty-four sub-atomic parts, eight with the property *being an electron*, eight with the property *being a proton*, and eight with the property *being a neutron*. Since the parthood relation is transitive, it follows that *being an H2O molecule* is the property of having thirty sub-atomic parts, twelve with the property *being an electron*, ten with the property *being a proton*, and eight with the property *being a neutron*. In other words, the property *being an H2O molecule* is composed of the properties *being a hydrogen atom* and *being an oxygen atom*, which in turn are composed of the properties *being an electron*, *being a proton*, and *being a neutron*. We therefore have the following identities:

> The property *being an H2O molecule* = the property composed of the properties *being a hydrogen atom* and *being an oxygen atom* = the property composed of the properties *being an electron*, *being a proton*, and *being a neutron.*

As with material objects, we can extend the list by specifying the property *being an H2O molecule* in terms of its parts at the level of quarks, at the level below the level of quarks, and so on, until we reach the fundamental or bottom level. If there is no fundamental level, the list would be infinite. But again, unless composition is identity, this will get us only reduction within a level. *Being an H2O molecule* is not reduced to properties at a lower level but merely to different structural properties at the same level – it is horizontal or sideways reduction, not vertical or downward reduction. If the identities Kim has in mind are identities of *this* sort, i.e. identities between same level properties specified in terms of their parts and properties at various levels, then the identities do *not* stop causal drainage from higher to lower levels. We can still worry whether the property of the whole causally competes with the properties of the parts taken collectively. For reduction to block causal drainage, composition must be identity which generates the following identities:

> The property *being an H2O molecule* = the properties *being a hydrogen atom* twice over and *being an oxygen atom* = the properties *being an*

electron twelve times over, *being a proton* ten times over, and *being a neutron* eight times over.

Here the higher level property reduces to genuinely lower level properties, not merely to other structural properties at the same level. Of course, for causal drainage to be effectively blocked, the identities must go all the way down, either to the bottom level, if such a level exist, or indefinitely, if there is no bottom level.

3.18 Against Inter-level Reduction as a Solution to the Drainage Problem

Is such a reduction of higher level properties to lower level properties a viable solution to the drainage problem? As mentioned earlier, with the claim that composition is identity comes the puzzle of how one thing can be identical to many – how can one higher level property be identical to many lower level properties? But even putting that aside, I do not believe this kind of reduction of properties across levels will solve the drainage problem. Here is why: Assume there is no bottom level and that composition is identity. Now take a property *H* at any level L. Say that *H* is composed of lower level properties *R* and *S*. Which object instantiates *H*? Not any object at level L. The reason is that *H* is identical to lower level properties *R* and *S* and *R* and *S* are properties of objects at a level below L, L-1. Now consider these properties. Which object instantiates *R*? Not any object at level L-1. The reason is that *R* is itself composed of properties, say *X* and *Y*, at a level below L-1, L-2 (the same for *S*). Which object instantiates *X*? Not any object at level L-2 since *X* is composed of properties at a level below L-2, L-3. And we could go on like this without ever reaching an object which instantiates a property. In short, if there is no bottom level and composition is identity, no object ever gets to instantiate a property. In the original drainage problem, causal powers threatened to drain away. If we try to avoid this by reducing properties across levels, it seems that properties themselves (or at least their instantiations), threaten to drain away. And since objects are causes in virtue of instantiating properties, the result is the same – if there is no bottom level, there is no causation. So identifying higher level properties with the lower level properties that compose them is of no help at all against Block's argument that causal powers drain away if there is no bottom level.

Note that this new version of the drainage argument against property-composition being identity is independent of the Exclusion argument and assumes none of its premises. In the way I set it up, it does, however, assume a substance-attribute model for properties according to which substances (objects or events) *instantiate* properties. Could the proponent of the Exclusion argument avoid the problem by going for a bundle view, according to which material objects are bundles of properties (either universals or tropes)? Here's an argument to the effect that a bundle view *would* avoid the drainage problem I've sketched above (I present it merely to reject it): If objects are bundles of properties, the instantiation relation would simply be an instance of the parthood relation – an object O has a property if that property is a part of the bundle of properties that is O. Since the parthood relation is transitive, this means that there is no problem with an object having both a composite property and its parts. For instance, if an object O has property *H* as a part and *H* itself has properties *R* and *S* as parts, then O also has *R* and *S* as parts. If properties decompose endlessly, this just means that O decomposes endlessly. And if composition is identity, this just gives us identities all the way down – *H* is identical to lower level properties *R* and *S*, *R* and *S* are identical to even lower level properties *X, Y, Z,* and so on. Unfortunately, this argument is unsound. On a bundle view, properties aren't *just* parts of objects, they are *compresent* parts. That is, they are parts that occupy exactly the same spatio-temporal region. For instance, a red sphere with a mass of 1 has the properties *being red*, *being spherical*, and *having mass 1* and those properties occupy exactly the same spatio-temporal region – the spatio-temporal region where the sphere is located. Without this qualification, the bundle view would have some strange consequences. For instance, the very same object would not just have a one mass, it would also have the mass of any of its proper parts. If a property is a property of an object only if it is a compresent part of the bundle that is that object, this is avoided. As an example, consider again Table that is composed of Top and Bottom. Say that Table has *mass 2* and that Top has *mass 1*. If parthood of Table were sufficient for being a property of Table, Table would also have *mass 1* according to the Bundle view, since the property *mass 1* is part of Top and Top is a part of Table. But parthood of Table is *not* sufficient for being a property of Table. What is also required is that the property is a compresent part of Table. And *mass 1* is not a compresent part of Table as it does not occupy the spatio-temporal region that Table does. Rather, *mass 1* is a compresent part of

Top, and as such, only occupies a proper part of the spatio-temporal region that Table occupies. Hence, *mass 1* is not a property of Table. This solves the problem of Table having many distinct masses.

In solving this problem, however, we've re-introduced the drainage problem for properties. Say that a bundle view is true and start with any composite property H at a level L which has lower level properties *R* and *S* as parts. Which object has *H*? That is, of which object is *H* a compresent part? No object, since *H* is identical to lower level properties *R* and *S* and they are not compresent parts of the same bundle. But which object has *R* as a compresent part? No object, since *R* is composed of even lower level properties, say *X* and *Y,* and *X* and *Y* are not compresent parts of the same bundle. And we could go on asking this question forever. No property would ever get to be a compresent part of any bundle of properties since every property is identical to lower level properties which themselves are not compresent parts of the same bundle. So whether one endorses a substance-attribute model for properties or a bundle view, the result is the same. Properties and causal powers drain away if there is no bottom level.[90]

3.19 Solving the Drainage Problem I

To sum up, the drainage problem presents Kim, and other philosophers who endorse the Exclusion argument against non-reductive physicalism, with an uncomfortable dilemma. On the one hand, if higher level properties do *not* reduce to lower level properties, they appear to be causally excluded by these lower level properties taken collectively. If there is no bottom level and every level has a level below it, this means that causal powers drain away and that there is no causation anywhere. On the other hand, if higher-level properties *do* reduce to lower-level properties and there is no bottom level, then for any object O at any level L, O does not instantiate any property since every property is identical to properties at a level

[90] Note that these arguments against inter-level reduction of properties do not depend on the Levels view being true, despite their reference to levels. Rather, if sound, they show the Levels view to be false. No object at level L instantiates properties (call them L-objects and L-properties) because every L-property is identical to properties at a lower level L-1 and no object on that level instantiates L-properties. But if there are no L-properties then it is difficult to see in what sense there could be any L-objects. And if there are neither L-properties nor L-objects, then it is difficult to see in what sense there is really a L-level. Of course, this is the result we want since, as I mentioned earlier, the composition-as-identity view is the view that only a single level exists.

below L, which are not properties of O. Again, causal powers drain away and there is no causation anywhere. In either case, it looks as if it is left as an open question whether there is causation. And assuming that this is unacceptable, it seems we have a good reason for rejecting the Exclusion argument off-hand. The Drainage problem establishes a *reductio* against the Exclusion argument.

An advocate of the Exclusion argument can try to avoid this in a couple of ways. First, she might insist that, as a matter of contingent fact, there *is* a bottom level. This would defuse the worry about causal powers draining away, at least in our world. It is not clear, however, that there are empirical grounds for making such a claim. At the very least, it seems to be an empirically open question whether or not there is a bottom level.[91] In any case, even if there is a bottom level and causal drainage is blocked in our world by that level, the Exclusion argument would still have the implausible consequence that it is entirely contingent whether or not there is causation. An alternative would be to go for the stronger claim that it is metaphysically necessary that there is a bottom level. But this is quite controversial and to my knowledge, rarely defended in the literature.[92]

Second, a proponent of the Exclusion argument might choose to reject the claim that broadly physical higher-level properties are causally competing with the lower level properties that compose them, even when they are considered collectively. So not only do some higher level properties have causal powers that do not reduce to the powers of any singular lower level property, they also have powers that does not reduce to the powers of any *collection* of lower level properties. How plausible is this? For some higher level properties, a case might be made. These are the structural properties where the specific structure of the parts matter to the causal powers of the property the parts compose. Take two kinds of molecules with different causal powers, butane and isobutene, which nevertheless have the same parts (four carbon atoms and ten hydrogen atoms).[93] What makes these molecules different is the way their parts are structured. In this case, it might be argued that the structure of the whole has causal powers that the mere sum of the parts do not,

[91] See Block [2003] and Schaffer [2003] for some discussion of this.
[92] An exception is Williams [2006].
[93] The example is due to David Lewis [1986a].

and those powers are exactly what distinguishes the powers of butane molecules from the powers isobutene molecules. Kim sometimes writes as if this is his view:

> [Micro-based properties] do not supervene on [their constituent properties] taken individually or as group. Rather, they supervene on *specific mereological configurations* involving these micro-properties – for a rather obvious and uninteresting reason: they *are* identical with these micro-configuations.[94]

The structure of the carbon and hydrogen atoms that compose a butane molecule is, I take it, an example of what Kim has in mind when he talks of "specific mereological configurations". But even if it is true of some structural properties that the structure itself has irreducible causal powers, it is unlikely to be true of all. Earlier I discussed the case of mass and how the mass of a higher-level object arguably compete with the masses of its part to exert certain downward force, and that to claim otherwise would be to accept that the law of gravitation is systematically violated. But even so, in that example, the way these lower level masses are structured does not seem to make a difference to the causal powers of the mass of the whole. An object composed of ten parts, each with a mass of 1, will have the same higher level mass (a mass of 10), whether its parts form a straight line, a square, a circle, and so on. If so, the structure of the parts cannot be what blocks the causal drainage from the higher to the lower level with respect to the causal powers of mass. And now we can simply run the drainage argument against the property *mass*. Start with an object O at any level. O does not instantiate mass *M* (where *M* is a determinate mass), since *M* is composed of masses at the level below, which are not properties of O. But no object at that lower level instantiates these lower level masses since they are themselves composed of masses at levels even lower, and so on – mass drains away. And this, I contend, is enough to constitute a *reductio* against the Exclusion argument.

3.20 Solving the Drainage Problem II

If the Exclusion argument is to have any hope of establishing reductionism about mental properties, something has to give. The proponent of the argument cannot consistently hold the following claims:

[94] Kim [1998], pp. 117-118.

(i) It's an open question whether there is a bottom fundamental level.

(ii) It's not an open question whether there is causation.

(iii) The Exclusion Argument, in its original form and applied to higher level properties in general, is sound.

Since we are assuming that (i) and (ii) are true, the problem is with the Exclusion argument. But which premise? It will not help us to reject the premise that there is higher level causation since this will play right into the problem of causal powers draining away (if there is no bottom level, all levels are relatively higher levels). The claim that lower-level properties causally compete with the higher level properties they compose is at least justified in the case of mass. And reduction of higher-level properties to lower-level properties is also facing a form of the drainage problem, since the properties to be reduced would never get to instantiate if there is no bottom level. This leaves the Exclusion principle. However, when setting up the Exclusion problem we saw that this principle is motivated. To reject it in a dialectically satisfactory manner we have to explain why those motivations are either ill-founded or do not really support the Exclusion principle as it stands. Here it seems to me that the latter strategy has more promise. We need a version of the Exclusion principle that allows for the kind of overdetermination that occurs when a higher level property causes an effect that is also caused by the properties that compose it, while still prohibiting the kind of overdetermination that occurs in the standard examples from the literature.

3.21 Inter-level Reduction of Causal Powers

To bring out a relevant difference which might help us to draw the distinction between different forms of overdetermination, let me again focus on the case of mass and how a higher level mass, call it *H*, can cause effects also caused by a sum of lower masses, call it *L*, for instance cause a certain downward force to be exerted on a surface. I have already argued that both *H* and *L* are causing an exertion of downward force F on a surface when object O is placed on it. *H* cannot causally exclude *L* without violating the law of gravitation and *L* cannot causally exclude *H* since this leads to causal powers draining away in a bottomless world. I have also argued that *H* cannot be identical to *L* since this leads to a different form of

drainage problem according to which a property like mass would drain away in a bottomless world. This, however, leads us to the problem of cumulative effects. If *H* and *L* both exert downward force F, then the object that has *H*, and which is composed of the parts that each has one of the masses in *L*, would have a mass of H plus each of the masses in L. And that's obviously not the case. If two parts each have a mass of 1, the object they compose will just have a mass of 2, not a mass of 2 plus the mass of each part. But if (determinate) cumulative effects cannot be overdetermined, how can *H* and *L* both be sufficient causes of exerting force F? How can *H* and *L* both exert *the very same token force*, not just two distinct forces of the same amount, or two token-distinct forces of the same type? If *H* and *L* bestow *distinct powers* to exert the same amount of downward force, this is a mystery. Assume that a composite object O with mass *H* is placed on a surface, that *H* bestows on O the power to exert downward force F under those conditions, and that this power is manifested as a result. If O's two parts each have powers to exert a downward force of half an F, the sum of these powers is token-distinct from the power bestowed by *H*, and these powers also manifest when O is placed on the surface, then there seems to be no good reason why the combined force exerted on the surface wouldn't be F times two (the F exerted by O in virtue of *H* plus half an F exerted by one of O's parts plus half an F exerted by O's other part).

The mystery disappears, however, if we reject the assumption that *H* and *L* bestow *distinct* powers to exert downward force F. If instead, the power *H* has to exert downward force F is the very same power *L* has to exert downward force F, we should *not* expect that total downward force exerted on the surface when O is placed on it, is F times two. When O is placed on the surface, there is just *one* power that is manifested, (a token of the type of power to exert downward force F) and hence just *one* downward force that is exerted. Given the empirical facts of the case (that O exerts downward force F, not F times two), it simply must be the case that *H* and *L* bestow on O the very same power to exert downward force F. We have reduction of powers across levels – the power bestowed by the higher level property reduces to the power bestowed by a sum of lower level properties.

3.22 Reduction of Powers and Causal Overdetermination

Even though the power bestowed by the higher-level mass reduces to the power bestowed by the sum of lower level properties, the exertion of force F is still overdetermined. When object O is placed on a surface, the exertion of downward force F has two distinct sufficient causes, *H* and *L*, since both bestow a power on O that manifests. Also, it is systematic overdetermination since it occurs whenever an object exerts a force according to the law of gravitation (at least provided it has more than one proper part with mass). But it is different from the kind of overdetermination that occurs in the standard examples from the literature. Consider again the case of the two snipers whose bullets happen to hit Target and kill him at exactly the same time. Here, the death is overdetermined as well and the two bullets both have a power of the same type (the power to kill Target). But it cannot be the case that the bullets have the same token of this power. The obvious reason is that the power to kill Target had by one bullet could manifest without the power to kill Target had by the other bullet manifesting. This would happen if one of the bullets was fired slightly later or earlier or if one of the bullets wasn't fired at all. If so, only one of the bullets would be a cause of his death. The other bullet would have a power to cause Target's death that failed to manifest. In the case of the object and the sum of its parts both exerting a downward force on a surface, however, this should be not possible given that the composite object and the sum of its parts have the very same token of the type of power to exert downward force F. But this seems right. A composite object cannot exert downward force F on a surface without its parts also exerting that downward force.

So, while both are cases of overdetermination, the sniper case involves the manifestation of two distinct tokens of a type of power (the power to kill Target), whereas the mass case involves the manifestation of just a single token of a power (the power to exert downward force F). We can call the former *strong overdetermination* and the latter *weak overdetermination.*

(Strong Overdetermination):

> An effect is strongly overdetermined iff it is the manifestation of two distinct token-powers.

(Weak Overdetermination):

> An effect is weakly overdetermined iff it is the manifestation of a token-power that is bestowed by two (or more) distinct properties.

Is this a relevant distinction, as far as a revision of the Exclusion principle is concerned? Given the original motivations for the Exclusion principle, the answer is arguably yes. Firstly, one problem with systematic overdetermination is that it seems to require extraordinary circumstances. It is quite a coincidence if, in a handful or even a single case, two snipers happen to hit a target at exactly the same time. But to assume that this happens in general is incredible. However, in cases of weak overdetermination, it is not at all incredible that there are two sufficient causes of a single effect since the overdetermining causes bestow the very same power, the manifestation of which is the overdetermined effect. Secondly, systematic overdetermination seems to involve more causes than needed which goes against principles of ontological simplicity and appears *ad hoc*. However, as the previous discussion has shown, we *do* need to allow for both *H* and *L* to be causes, since all other alternatives have unacceptable consequences, and arguments from simplicity apply only to positing entities that aren't needed. There is nothing ad hoc about the proposal either. Rather, it follows from the evidence, partly empirical and partly philosophical, that *H* and *L* must have the same token-power to cause exertion of force F. Finally, as we saw, the fact that H and L have the same token-power to cause this effect explains how cumulative effects can be overdetermined at all. The reasons we have for prohibiting systematic overdetermination of the sniper-case variety therefore do not apply to the kind of overdetermination that occurs in the case of higher and lower level masses exerting force. This opens the door for a dialectically satisfactory revision of the Exclusion principle. Recall, a rejection of any of the claims in the Exclusion problem should preferably be accompanied by either an explanation of why the reasons for accepting this claim are ill founded, or an explanation of why these reasons are really reasons for accepting a different claim. And this latter explanation is now provided. Since the worries about overdetermination pertain exclusively to strong overdetermination, they are insufficient to support an Exclusion principle that prohibits both strong *and* weak overdetermination. Hence, the Exclusion principle

should be revised in such a way that it rules out only systematic strong overdetermination:

(Exclusion-Strong):

> In general, if an effect E is the manifestation of token-power P bestowed by C, then all properties that do not bestow P are excluded from being a cause of E.

Given that the higher level composite properties that causally compete with the sum of their parts to cause an effect *do* bestow the same token-power to cause that effect, this means that the Exclusion principle does not apply to either. Hence, the Exclusion problem applied to higher level properties dissolves.

Does the Exclusion problem for mental properties dissolve as well? That depends on whether or not the revised Exclusion principle applies to mental properties. If the causal powers bestowed by mental properties are token-identical to powers bestowed by some physical properties, the principle does not apply and mental properties fail to be causally excluded. In the next chapter I will defend the view that this is in fact the case.

Now, the case of mass and exertion of downward force is special in the sense that the forces exerted add up in a straightforward manner. If the parts collectively exert force F and the whole exerts F, then we know that the total force exerted by the object must be F times two, unless the parts and the whole exert the very same token of force F. But in other cases, things are not quite as simple. Consider a different power an object might have in virtue of its mass, for instance, the power had by a bullet to kill Target. Say that the bullet hits Target and this power manifests. In this case, we can worry that the mass of the bullet (*H*) is causally competing with the masses of each of its parts considered collectively (*L*) to cause that effect. However, we can't check empirically whether those effects have accumulated, since a killing is an event that can happen only once to a person at a time. So how are we justified in claiming that *H* and *L* bestow the same token of the power to cause Target's death? After all, if they do not, Target would still just die once. Still, if we assume that the force-case and the killing-case are to be treated

similarly, then we do have a reason for thinking that *H* and *L* bestow the same token of the power to cause Target's death, since we know that this must be so in the force-case. And this is not an unreasonable assumption. After all, it is presumably no coincidence that *H* and *L* bestow the same token of the power to exert downward force F. For instance, this may come down to the fact that the causally relevant instances of *H* and *L* are co-located or that one composes the other. Whatever the explanation is, however, it will apply to the power to cause Target's death, as well as other powers that *H* and *L* share.

3.23 Concluding Remarks

Let me briefly sum up what has been argued in this chapter. The Exclusion problem for mental properties consists in four mutually inconsistent claims that are independently motivated: Mental Causation, Closure, Exclusion, and Distinctness. Advocates of the Exclusion argument opt for rejecting Distinctness – mental properties (and other non-physical properties) must be reduced to broadly physical properties. In the case of mental properties, these broadly physical properties will likely be complex neural properties. The Exclusion argument, however, faces the Drainage problem when applied to broadly physical higher level. I argued, contra Kim's original discussion of this problem, that at least some higher level properties are facing causal exclusion by the sum of the lower level properties that compose them, and that this leads to the causal powers draining away if there is no bottom level. I also argued, contra Kim's most recent discussion of this problem, that reduction of higher level properties to the lower level properties that compose them leads to a new form of the drainage problem. The conclusion is that some effects caused by higher level properties are also sufficiently caused by the sum of their parts.

As this is overdetermination, the original formulation of the Exclusion principle must be false. However, there is an important difference between the overdetermination that the Exclusion principle is supposed to rule out and the kind of overdetermination involved in cases where higher level and lower level properties cause the same effect. In the latter case, the overdetermined effect is the manifestation of a *single* power bestowed by each of the overdetermining causes. I called this weak overdetermination. The case of higher level mass justifies this

claim – if a higher level mass and the sum of its parts bestow distinct tokens of a power to exert a downward force, the total force exerted by an object would not just be the combined force exerted by each of its part. But it is, so a higher level mass and the sum of its parts must bestow just a single token of the power to exert a certain downward force. This kind of overdetermination does not suffer from the problems that come with the standard cases of overdetermination from the literature, where the overdetermined effect is the result of two distinct powers manifesting. I called this strong overdetermination. Hence, the Exclusion principle should not rule out weak overdetermination. This means that the proponent of the Exclusion argument can block causal drainage, while consistently hold that the Exclusion argument establishes reductionism about mental properties. All this requires is that the overdetermination that occurs if mental and physical properties overdetermine behavioural effects are of the strong kind.

Chapter 4

Motivating and Defending the Subset View

Abstract: In this chapter I put some of the conclusions from the previous chapter to work. Since the exclusion principle prohibits only strong systematic overdetermination, the Exclusion problem would be solved if mental and physical properties weakly overdetermine their common effects. I introduce Sydney Shoemaker's so-called Subset view, according to which this is the case. I discuss and reject the claims that the Subset view commits you to endorse that causes are proportional to their effects or that some form of functionalism is true of mental properties. Instead, I argue that the Subset view is motivated by the very statements that constitute the Exclusion problem. Finally, I consider some potential limitations of the Subset view.

4.1 Mental Causation and Weak Overdetermination

Given the fact that the Exclusion principle must be revised to allow for weak overdetermination, it is tempting to think that mental causation might be understood in a way that makes mental effects weakly overdetermined by mental and physical properties. What is required is that the powers to cause physical effects had by any instance of a mental property, must reduce (at the token-level) to the causal powers had by some instance of a physical property. Of course, this will trivially be the case on a reductive physicalist account of mental properties – if mental properties are identical to physical properties, they have the very same powers these physical properties have, both at the type and token-level. However, reduction of causal powers need not entail reduction of the property that has these powers.

4.2 Reduction of Powers without Reduction of Properties

Let's say that a set of powers form a *causal profile* for a property *P* if and only if it is the largest set of powers had by *P*. In other words, any proper subset of powers had by *P* will not be a causal profile for *P* since it is not the largest set of powers had by P. The property that has all and *only* the powers in the causal profile I will call the *occupant* of that causal profile.[95] Now, note that it is a necessary condition

[95] Note that this is a more demanding notion of occupancy than is sometimes used in the literature, where a property is said to be an occupant of a causal power or causal role if it is the property that

for the identity between two properties that they have the exact same causal profile - if the profile of a property *X* contains at least one causal power not contained in the profile for property *Y*, then *X* and *Y* must be distinct. Mental properties are therefore distinct from physical properties if their causal profiles are distinct from any causal profile occupied by a physical property.

Is it also a sufficient condition for the identity between two properties that they have the exact same causal profile? That is, is it the case that if *X* and *Y* are distinct, the profile of *X* must contain at least one causal power not contained in the profile for property *Y*, or *vice versa*? If you say yes, you hold a causal theory of properties, according to which properties are fully individuated by their causal powers, at least throughout worlds with the same laws of nature. Physicalists and naturalists tend to accept some form of this view, and even if it is rejected as being true in general, even non-physicalists are likely to accept it for physical properties.[96] If so, the difference between mental and physical properties is bound to show up as a difference in their causal profiles. The reason for this is that Distinctness is motivated by the argument from multiple realization – the instantiation of *M* requires no particular physical property to be instantiated as *M* can be realized by different physical properties in different individuals. Now assume, for *reductio*, that mental property *M* has a casual profile that is identical to the profile of some physical property *P*. If so, every individual that instantiates M would thereby also instantiate *P* (since physical properties such as *P* are fully individuated by their causal powers). But if the argument from multiple realization is sound, this is false. So *M* cannot have a causal profile that is identical to the causal profile of some physical property.[97] Note that this argument does not assume that a causal theory of properties is true of mental properties, just that it is true of physical properties

---

has this power or plays that role, or the occupant of a set of powers if it has the powers in that set. On my account, this would be a necessary, but not a sufficient condition for being an occupant of these powers, since it does not rule out that the property has additional powers not in that set.

[96] Physical properties at the same level of mereological complexity at least. The causal profile of a sum of lower-level physical properties may be identical to the causal profile of the higher-level physical property they compose, yet I would deny that the sum of lower-level properties is identical to that higher-level property for the reasons given in the previous chapter. In any case, we might say that the sum of lower-level properties differ from the higher-level property they compose with respect to their logical properties since a sum of properties is many, whereas a composite property is one.

[97] This argument also has as a consequence that the causal profile of *M* cannot include the causal profile of any particular physical property. I establish the same conclusion with a different argument shortly.

(towards the end of the chapter I consider how a rejection of the causal theory of properties for mental properties causes problems for physicalism). In any case, the upshot is that for the purposes of the arguments in this chapter, we can assume that it is a sufficient condition for the identity between two properties that they have the exact same causal profile.

Or in other words, Distinctness (i.e. the claim that mental and physical properties are distinct) is true if and only if, for any mental property *M* and its causal profile, and any physical property *P* and its causal profile, *M*'s profile either contains some power not contained in *P*'s profile, or *P*'s profile contains some power not contained in *M*'s profile.[98] This, however, is entirely compatible with *M*'s profile fully or partially overlapping with the profile of some *P*. If the causal profiles of mental properties only partially reduce to powers contained in the profiles of physical properties, mental and physical properties are distinct since mental properties occupy causal profiles that contain powers not contained in the profile of any physical property. In this case, the causal profiles of mental properties would extend beyond the profiles of physical properties. And even if the causal profile of any mental property fully reduces to powers contained in the profile of some physical property, mental properties are still distinct from physical properties if the profiles of the latter contain additional powers, not contained in the profile of mental properties. In this case, the causal profiles of physical properties would extend beyond the profiles of mental properties.[99]

4.3 In What Way are Mental and Physical Properties Distinct?

Mental and physical properties are distinct then, if mental properties have causal profiles that contain powers not contained in the profile of any physical property or, if there are no such powers in the profiles of mental properties, if the physical causal profiles to which the mental causal profiles reduce contain additional causal

[98] Is identity between causal profiles also a sufficient condition for identity between the properties that occupy these profiles? Consider two objects that have exactly the same causal powers. I take it that many would be inclined to accept that these two objects must have the exact same properties as well, at least if the objects exist in worlds with the same laws. Exceptions would involve properties whose nature is not exhausted by their causal profiles. In other words, properties supervene on their causal profiles with nomic necessity.

[99] Of course, it might be the case that the profiles of mental properties extent beyond the profiles of physical properties and *vice versa*.

powers. Let's first consider the possibility that mental properties are distinct from physical properties in virtue of having causal powers that no physical property has. This, for instance, would be the case according to emergentism about mental properties. It cannot, however, be the case according to non-reductive physicalism. First, assuming that physicalists universally accept the Closure principle (recall, this is roughly that claim that every physical effect has a sufficient physical cause), the non-reductive physicalist must reject that mental properties are distinct from physical properties in virtue of having powers to cause *physical* effects that no physical properties have powers to cause.[100] So if mental properties have causal powers not had by any physical property, they must be powers to cause non-physical effects. However, as we saw in the previous chapter, on the physicalist account, causing a non-physical effect, for instance a mental property, is to cause the physical grounding base for this effect (see chapter 3.6). This means that if a property has a power to cause a mental (or otherwise non-physical) effect, this must be a power to cause a physical effect - the physical property or properties that this mental effect is grounded in. So, assuming physicalism, mental properties cannot be distinct from physical properties in virtue of having powers to cause mental effects that no physical property has the power to cause either, since this as well would violate the Closure principle. From a physicalist perspective then, distinctness between mental and physical properties is best based on the claim that the physical causal profiles of which the profiles of mental properties are part, contain additional powers. Naturally, this is not an argument exclusively about mental properties but can be applied to other non-physical properties as well. Call it *the argument for additional physical powers* (APP). This, it seems to me, is an argument that all non-reductive physicalists should accept, in so far as they accept that mental properties (or other non-physical properties) have causal powers and that the Closure principle is true. I will say more about this later.

Notice that APP is an argument that is concerned with causal powers at the type level only. There is nothing in the argument that allows us to conclude that the token-powers of every mental property are identical to token-powers of some physical property. This is the case even if we consider instances or tokens of

[100] I chapter 6 I will discuss a form of non-reductive physicalism that challenges this.

properties – an instance of a mental property may bestow tokens of certain causal powers on a particular that are distinct from tokens of powers bestowed by a physical property on this particular, despite these tokens being of the same type.

4.4 The Subset View and the Exclusion Problem

That mental properties (and other non-physical properties in general) are distinct from physical properties in the manner concluded by APP, and that the causal powers of these properties reduce to some powers of physical properties, at *both* the type and token level is, in a nutshell, the view known as the Subset view. *Prima facie*, the Subset view applied to mental properties provides an attractive solution to the Exclusion problem. First, the view allows for mental causation since it holds that mental properties exist and have causal powers that sometimes manifest. Second, the view is compatible with the causal closure of the physical domain in the following way: Any case of mental causation involves the manifestation of a mental power (i.e. a power had or bestowed by a mental property). But since any mental power is identical to a physical power (i.e. a power had or bestowed by a physical property), any case of mental causation is also a case of physical causation. Hence, on this view, mental causation does not in any way threaten the principle of causal closure. Thirdly, as we have just seen, according to the Subset view, causation by irreducible mental properties does not violate the Exclusion principle since this prohibits only (systematic) strong causal overdetermination, and on the Subset view mental effects are merely weakly overdetermined. Fourthly, not only is the Subset view a non-reductive view since it claims that mental and physical properties occupy distinct causal profiles, it also fits nicely with the motivation for accepting Distinctness in the first place, which was provided by the argument from multiple realization. Recall, the argument from multiple realization states that a mental property cannot be identical to any physical property since the former can be shared by creatures that otherwise share no relevant physical properties. This fact can be accommodated by the Subset view. A mental property can be shared by two creatures with different physical properties if these physical properties bestow *some* of the same powers (at the type level) and these powers are exactly the type of powers also bestowed by the mental property.

In the following, I will expand on the Subset view as a position in the philosophy of mind and defend it as a solution to the Exclusion problem. The point of departure will be Sydney Shoemaker's influential formulation of the view. Of course, it is one thing to work out a position in detail and evaluate its potential theoretical benefits, it is another to argue that the position is true. Towards the end of the chapter I will show how the Subset view can be motivated.

4.5 Shoemaker's Subset View

Shoemaker does not initially focus his discussion on mental properties but instead on the general notion of realization. Take the relation that holds between two properties when one has a causal profile that is a proper subset of the profile had by the other. Call this *the subset relation*. According to Shoemaker, it is at least a necessary condition for realization that the realized property and its realizer stand in the subset relation.[101] That is, one property is realized by another only when the former has causal powers that are a proper subset of the set of powers had by the latter.[102] Again, the subset relation is here taken to hold between powers at both the type and token level. In Shoemaker's most recent writings, the view is slightly more complicated as he introduces a distinction between forward and backward looking causal powers.[103] Forward looking causal powers are simply causal powers as we have been talking about them so far. They are forward looking in the sense that they are specified in terms of potential future events – i.e. the effects they are powers to cause. Backward looking causal powers, on the other hand, are not specified in terms of effects but rather in terms of the properties or events that are causes of the

[101] Strictly speaking, the subset relation is a relation between sets of what Shoemaker calls *causal features*, to each of which corresponds a causal power that the property bestows on the its possessors (Shoemaker [2007], p. 24). For present purposes, this distinction is irrelevant.
[102] Shoemaker [2001], p. 78. Shoemaker makes an exception for conjunctive properties and their conjuncts which he argues also stand in a subset relation but which fail to stand in a realization relation. Hence, standing in the subset relation cannot sufficient for realization on his account. I will argue later that in any case, realization is best seen as a combination of the subset relation and the grounding relation. Another thing to note is that, in his most recent writings, Shoemaker distinguishes between two kinds of realization, what he calls *same subject property-realization* and *micro-realization*. The former is an intra-level relation since it involves realization of one property by another where both are properties of the same subject or individual. The latter is an inter-level relation that holds between microphysical states of affairs and higher level properties As such, micro-realization is similar to the notion of property-composition I've made use of earlier. In the following, the relevant notion of realization is same subject property-realization since the properties competing to cause the same effects in cases of mental causation are properties of the same subject, for instance a mental property competing with a neural property to cause a behavioural effect. So when I talk of realization, it should be understood as same subject property-realization.
[103] Shoemaker [2007], p. 12.

property that has these powers. For instance, the redness of a tomato is typically caused by exposing the still unripe tomato to sunlight. This is a backward looking power of the tomato. In turn, this colour can then cause certain reddish experiences in someone looking at the tomato, cause the tomato to be plucked, etc. These are forward looking powers of the tomato. According to Shoemaker, the subset relation holds both between sets of forward and backward looking powers of a realized property and its realizer. However, in the case of forward looking powers, it is the *realizer* that has the superset of powers and the realized property that has the subset of powers, whereas in the case of backward looking powers, it is the *realized property* that has the superset of powers and the realizer that has the subset of powers. I will later return to the question of why Shoemaker introduces this complication and provide some reasons for why it is better done away with. For now, I will ignore backward looking causal powers and focus exclusively on forward looking causal powers.

So a realizer has a set of causal powers, a subset of which contains the powers of the property it realizes. Shoemaker then goes on to claim that mental properties are realized by physical properties in this sense. In other words, the Subset view is correct for mental properties. What justifies Shoemaker's claim that realization is best understood as involving the subset relation? Since realization is a technical notion, philosophers are, to a large extent, free to define it as they please. The value of a particular definition will depend on how well it fits the platitudes, if any, regarding the concept being defined, and the theoretical advantages the definition provides. For instance, in contemporary philosophy, the notion of realization typically comes as part of a non-reductive physicalist package, according to which realization is the relation that holds between physical and non-physical properties and which is sufficient for physicalism about these non-physical properties. At least, that is the sense of realization which is relevant for Shoemaker's purposes. It has been a particularly popular notion among functionalists to describe the relationship between physical and functional properties.[104] Often though, the term is introduced without any detailed account of the precise metaphysics of the relation. Defining realization, at least in part, as the subset relation, can be seen as an attempt

[104] See Putnam [1967] and Fodor [1974] for some early formulations of this.

to put some metaphysical meat on these bones. Firstly, realization as the subset relation explains why realization is a non-reductive relation (the realized property and its realizer have numerically distinct causal profiles). Secondly, the claim that non-physical realized properties are grounded in their physical realizers, and thereby acceptable properties from a physicalist perspective, is at least made plausible by the Subset view, seeing that the powers had by the realizer (the superset) necessitates the powers had by the property it realizes (the subset) but not *vice versa*. Hence, realized properties asymmetrically supervene on their realizers and this is a sign of grounding.[105] The fact that the Subset view explains and motivates the platitudes associated with the notion of realization is in itself a theoretical virtue of the view, in addition to its promise of solving the Exclusion problem.

4.6 Shoemaker on the Exclusion Problem

It is worth noting that Shoemaker's proposed solution to the Exclusion problem in light of the Subset view differs somewhat from the one I sketched above. He has two arguments for why the Subset view solves the Exclusion problem. The first is based on the claim that mental properties, in a certain sense, are parts of physical properties according to the Subset view. This, of course, is not the sense in which lower level physical properties can be parts of higher level physical properties – Shoemaker is not claiming that mental properties at a lower level compose higher level physical properties. Rather, the part-whole relation he has in mind pertains to the causal profiles occupied by mental and physical properties. A proper subset of a superset is a proper part of that superset. Since a mental property occupies a causal profile that forms a proper subset of the profile occupied by its physical realizer, this is the sense in which mental properties are parts of their physical realizers. With this in mind, consider a mental property *M*, its physical realizer *R*, and a behavioural effect E, which is the causal product of both. Shoemaker considers whether this is a case of causal overdetermination and goes on to answer in the negative, using the following analogy:

> We can say […] that while the R instance causes E, it does so because it includes the M instance that causes E. We might compare this with the case

[105] I will consider the relationship between the Subset view and physicalism in more detail later.

> in which Smith dies as the result of a salvo of shots fired by a firing squad, but in which the only shot in that salvo that hit Smith was the one fired by Jones – the salvo killed Smith, but it did so because it included a particular shot, Jones', that killed Smith. This is obviously not a case of overdetermination.[106]

Shoemaker's second argument for why the Subset view solves the Exclusion problem draws on insights originally developed by Stephen Yablo. Causes, according to Yablo, should be proportional to their effects. Call this *the proportionality claim*. What this means is that 'roughly, they should incorporate a good deal of causally important material but not too much that is causally unimportant'.[107] Yablo suggests that a way of testing whether a property incorporates enough, but not too much, causal material with respect to an effect, and so, whether it is suited as a cause of that effect, is with the help of two conditions. First, effects should be contingent on their causes – if the cause had not occurred, the effect would not have occurred either. Second, causes should be adequate for their effects – if the cause had not occurred, then, if it had occurred, the effect would have occurred as well. Satisfaction of the former condition (call it the contingency condition) indicates that the property does not incorporate too much unimportant causal material, whereas satisfaction of the latter condition (call it the adequacy condition) indicates that the property incorporates enough causal material.[108] To illustrate this, Yablo applies the contingency and adequacy conditions to a case involving a determinable property competing with its corresponding determinate to cause an effect. Say that Sophie is a pigeon trained to peck at all and only red objects. One day she is presented with and pecks at an object with a specific shade of red, say scarlet. It seems there are several properties that are candidates for causing this effect. For instance, we might ask whether the causally efficacious property was the very determinate property *scarlet*, the less determinate property *red*, or the very indeterminate property *being coloured*. According to Yablo, the determinable property *red* is to be preferred as a cause of Sophie pecking over both *scarlet* and *being coloured*. The reason is that, intuitively, while the object being scarlet is enough, or adequate, for Sophie pecking, *scarlet* is

[106] Shoemaker [2007], p. 13. Note that I've changed Shoemaker's notation from Q and P to R and M.
[107] Yablo [1992], p. 274.
[108] Yablo [1992], p. 274. Yablo's account is more complicated than this and he has other conditions that should typically be met in order for a property to count as a cause of an effect (see pp. 276-277). They do not matter for present purposes so I will ignore them.

too specific to be a proper cause of Sophie pecking. After all, Sophie pecks at many objects that are not scarlet, for instance crimson and pink objects. The fact that *scarlet* is too specific, or incorporates too much causally irrelevant material with respect to Sophie pecking, is revealed by a failure to satisfy the contingency condition. If the object in front of Sophie had not been scarlet, but rather some different shade of red, she would have pecked at it anyway. *Being coloured*, in turn, is not specific enough. After all, there are many coloured objects that Sophie does not peck at, for instance blue and green objects. The fact that being coloured is not specific enough, or does not incorporate enough causally relevant material, is revealed by a failure to satisfy the adequacy condition. If the object in front of Sophie had not been coloured, then if it had been coloured, she might not have pecked (since there are worlds where the object is coloured but blue or green instead of red). *Red*, on the other hand, is just right. Intuitively it incorporates the right amount of causal material with respect to Sophie pecking. This is indicated by the fact that *red* satisfies both the adequacy condition and the contingency condition – if the object in front of Sophie had not been red, then, if it had been red, she would have pecked at it, and, if the object in front of her had not been red, but rather some different colour, she would not have pecked at it. In other words, of the three candidates, the determinable property *red* is proportional to the effect of Sophie pecking, whereas the more determinate property *scarlet* and the less determinate property being coloured are not (or at least not as proportional as *red*).[109]

According to Shoemaker, Yablo's proportionality constraints on causation also vindicate the causal efficacy of realized properties as they are understood on the Subset view:

> Where the only causal features of property P1 that play a role in producing an effect are ones that belong to property P2, of which P1 is a determinate or realizer property, there seems a good sense in which considerations of proportionality favour the instantiation of P2 over the instantiation of P1 as a cause of the effect.[110]

[109] This is very quick and involves several assumptions about the correct interpretation of counterfactuals and ordering of possible worlds that may be disputed.
[110] Shoemaker [2001], p. 81.

For instance, say that we ask whether, on a given occasion, a behavioural effect, such as Jones reaching for an umbrella, is the causal result of a certain mental property, say Jones' belief that it is raining, or its physical realizer. According to the Subset view, the mental property has certain causal powers, one of which is the power to cause Jones to reach for his umbrella. The powers had by the mental property are a subset of the powers had by the physical realizer. So the physical realizer also has the power to cause Jones to reach for his umbrella. However, the physical realizer has powers that are irrelevant with respect to this effect. Because of this, citing the realizer as a cause of Jones' behaviour is to cite a property that is too specific. As in the case of scarlet and the effect of Sophie pecking, this is revealed by a failure to satisfy the contingency condition. Had Jones not instantiated the particular physical realizer that he actually did but still believed that it was raining, he would have reached for the umbrella anyway. Here, Jones' mental property fares better - if Jones had not believed that it was raining, but rather had some different belief, say the belief that it was sunny, he would not have reached for the umbrella. In other words, if Jones' mental property has a power to cause him to reach for his umbrella and this power manifests, then Yablo's proportionality constraints tells us that this property is a cause of the behaviour, regardless of the fact that the physical realizer of this property, which has additional powers that are irrelevant for the effect in question, also has this power.

I find Shoemaker's reasons for thinking that the Subset view solves the Exclusion problem unconvincing and, as I will ultimately argue, they do not do full justice to the position. I will begin by considering the argument from the proportionality of causes and give some reasons for thinking that accepting this doctrine is unattractive from the perspective of the proponent of the Subset view. I will then reject these reasons and show that the Subset view can actually be motivated by endorsing the proportionality claim. After that, I consider to what extent this is compatible with Yablo's original claims and whether or not this commits the proponent of the Subset view to adopt the proportionality constraints. Finally, I will argue that even if the proportionality claim is accepted, this does little work to solve the Exclusion problem.

4.7 The Proportionality Claim and the Closure Principle

Recall that, on Yablo's account, the determinate property *scarlet* fails to be a cause of Sophie pecking.[111] This in itself may raise an eyebrow. Personally I think that it seems awfully hard to deny *scarlet* any causal efficacy *at all* with respect to Sophie's pecking. After all, being scarlet is *a way* of being red and *red* is causally efficacious with respect to this effect. A bigger problem, however, is the following: If the considerations that lead one to draw the conclusion that *scarlet* is causally impotent with respect to Sophie's pecking apply in the case of mental properties and their physical realizer, the conclusion should likewise be that the physical realizer of Jones's belief that it is raining outside is causally impotent with respect to Jones' behaviour (a conclusion that Yablo in fact draws). Based on the quote from Shoemaker above, he appears to accept this (as he says, in some cases we should favour the realized property over the realizer as a cause, indicating that he believes the realizer fails to cause the effect in such cases). But this is an unattractive conclusion. To say that, in a case of mental causation, the physical realizer fails to cause the effect is to embrace the denial of the Closure principle (assuming, uncontroversially, that there is no other sufficient physical cause of the effect). If this is Shoemaker's view (I do not think it is), he has not adequately solved the Exclusion problem, unless he gives a story of why we accepted the truth of this principle in the first place.

4.8 The Proportionality Claim and the Subset View

So adopting the proportionality constraints threatens to violate the closure principle. To see how a proponent of the Subset view might respond to this, let me first address an even more serious problem that, when dealt with, will throw some light on the status of the closure principle. This is the problem that accepting the proportionality claim threatens not just the truth of the closure principle, but the Subset view itself. Say that the proportionality constraints are what motivate the claim that a realized mental property has causal powers in the first place. For instance, a line of reasoning might go as follows: The proportionality constraints tell us that a realized mental or determinable property is a cause of certain effects. If a property is a cause of certain effects, then it must have the power to cause those

[111] Though he does allow that it is causally sufficient for Sophie pecking and thereby causally relevant. See Yablo [1992], p. 272.

effects. So the proportionality constraints tell us that the mental property or a determinable property has causal powers.

On the face of it, this way of getting to the powers of realized properties is not an attractive strategy for the proponent of the Subset view. Again, the problem is that the proportionality constraints arguably provide a unique cause of an effect among the candidates.[112] For instance, in the case of Sophie pecking (where *red* is the cause according to the proportionality claim), any property more determinate than *red* will fail to satisfy the contingency condition, and any property less determinate than *red* will fail to satisfy the adequacy condition (*being scarlet* is too specific and *being coloured* is not specific enough). That the proportionality constraints pick out a unique cause can also be seen by considering Shoemaker's claim that they favour the realized property over its realizer, when the powers responsible for producing an effect are all powers of the former. If so, among the properties that candidate to cause the effect, there will always be a property *Q* that has the smallest set of powers that include just those powers that are responsible for producing the effect. But if satisfaction of the proportionality constraints is what justifies that *Q* has those powers, and if *Q* is the unique property that satisfies the constraints, then it seems that the realizer of *Q* does not have those powers. And this is in direct contradiction of the Subset view, according to which *P* and its realizer stand in the subset relation. So if the proportionality constraints are taken to be the foundation on which the causal powers of realized properties is based, not only does this bring the proponent into conflict with the closure principle, it seems to be incompatible with the Subset view itself.

4.9 Reconciling the Proportionality Claim with the Subset View

To see how this worry can be dispelled, we should note that there might be other reasons to think that a property has causal powers, besides satisfying the proportionality constraints. If so, there might be other reasons to think that the realizer of a causally efficacious realized property also has the powers relevant for

[112] At least assuming that the higher-order determinable property has multiple same-level determinate realizers. If not, then the realizer satisfies the proportionality constraints whenever the realized property, and *vice versa*, and Yablo's account of causation would deem both to be causes whenever one of them is. Of course, denying that a certain higher-order property has more than a single realizer removes one of the main motivation for thinking that these two properties are distinct.

producing the effect in question. This opens the door for a reconciliation between the Subset view and the proportionality claim. For instance, Brian McLaughlin, realizing the potential conflict between the Subset view and the fact that the proportionality constraints generate unique causes, proposes on behalf of Shoemaker that a property *P* has a power to cause an effect if either (i): were the effect to occur under certain conditions, *P* would be the property proportional to this effect, or (ii) *P* necessitates a property that satisfies condition (i).[113] So for instance, *red* has the power to cause Sophie to peck since, were a red object to be placed in front her, she would peck, and *red* would be the property proportional to this effect. *Scarlet*, however, despite failing to satisfy the proportionality constraints with respect to Sophie pecking, will nevertheless also have the power to cause Sophie to peck since *scarlet* necessitates *red*. This point generalizes to cases mental properties and their physical realizers, assuming the latter necessitates the former.

Now, I find McLaughlin's proposal that a property has the power to cause an effect, merely in virtue of necessitating a distinct property with this power, unconvincing. The problem is that the claim is prone to counterexamples. For instance, it is plausibly the case that any object with a colour also has an extension. So *being coloured* necessitates *being extended*. In this case, on McLaughlin's proposal, any causal power bestowed by the property *being extended* would also be a power bestowed by the property *being coloured*. And this is counterintuitive. A better proposal replaces necessitation with grounding: A property *P* has the power to cause an effect E, if it grounds a property *Q* that has this power. That this is so actually follows from the transitivity of the grounding relation. Say that the proportionality constraints reveal that an object has the power to cause Sophie to peck in virtue of instantiating the property *red* (where the 'in virtue of' locution expresses the grounding relation). If the object is red in virtue of being some determinate shade of red, say scarlet, then it follows that the object also has the power to cause Sophie to peck in virtue of having the property *scarlet*. This would be a sort of reversed version of what Kim calls *the causal inheritance prin*ciple, according to which a higher-order realized property inherits the causal powers of its lower-order realizer. Here, however, it may be more appropriate to say that, by

[113]McLaughlin [2007], p. 167. McLaughlin dubs the latter direct power bestowing, the latter indirect power bestowing.

grounding it, the realizer inherits the powers of the property it realizes.[114] The reason *being coloured* fails to inherit the powers of *being extended* is that an object presumably isn't extended in virtue of being coloured. This leaves us with the following argument for the Subset view: Assume that *P* grounds *Q*, that *P* is the unique property proportional to effects E1-E5, and that *Q* is the unique property proportional to effects E6-E10. From this is follows that *P* has the powers to cause E1-E5, and that *Q* has the powers to cause E6-E10. *P* is not proportional to effects E6-E10, but since P grounds a property that is proportional to these effects (*Q*), it nevertheless has those powers. So *P* has all the powers of *Q*. *Q*, however, does not have the powers to cause effects E1-E5 since, one, *Q* is not proportional to these effects (*P* is), and two, *Q* does not ground *P* (grounding is asymmetric and by assumption, *P* grounds *Q*). So *P* has powers that *Q* does not. Hence, the powers of *Q* form a proper subset of the powers of *Q*'s grounding base *P* – we have the Subset view.

4.10 Yablo and the Subset View

Is the argument above compatible with Yablo's view on causation? Earlier (chapter 3.3) I stated that a property *P* is causally efficacious with respect to some effect E, if E is the manifestation of a power bestowed by *P*. If so, Yablo is wrong when he says that the causal efficacy of a determinable property excludes all determinates of this property from being causally efficacious. On the contrary, in so far as determinate properties ground their corresponding determinables and inherits their causal powers, the argument above shows that the causal efficacy of a determinable property *entails* the causal efficacy of all its determinates. Of course, Yablo may be seen as denying this claim about causal efficacy. His view may be that bestowing a power that manifests is insufficient for a property to be causally efficacious. What is required, in addition, is that the property satisfies the proportionality constraints. However, it is not clear to me what would motivate such a claim. First of all, note

[114] This is potentially misleading and it is important to understand it in the right way. When I say that a realizer inherits the powers of the property it realizes, I do not mean to say that the powers of the realized property are fundamental, relative to the powers of the realizer (though the metaphor of inheritance may suggest that). The direction of explanation from the power of the realized property to the powers of the realizer is epistemic, not metaphysical. The idea is that we are justified in thinking that a property has powers to cause certain effects because the property is proportional to these effects, and that this, in turn, justifies us in inferring that any property that grounds this property has those powers as well. This is the sense in which we might say that the realizer inherits the powers of the property it realizes.

that Yablo explicitly denies that satisfaction of the proportionality constraints is necessary for causation (or causal efficacy). What he says is this:

> Without claiming that proportionality is strictly necessary for causation, it seems clear that faced with a choice between two candidate causes, normally the more proportional candidate is to be preferred.[115]

Why is the more proportional property to be preferred? The preference might be purely pragmatic. For instance, it may be that, in general, the more proportional property has a higher degree of explanatory relevance than any determinate of this property. This seems justified in the case involving Sophie the pigeon. Someone trying to determine the cause of her pecking will typically be doing so for one of two reasons: Either to find out how best to make her peck or how best to stop her from pecking. In both cases, citing the determinable property *red* as a cause, rather than *scarlet*, will be more useful. If you want to make Sophie peck, it is easier to do so by placing red objects in front of her than scarlet objects (since red objects will typically be easier to find than scarlet objects). And if you want to stop her from pecking, you will need to remove all red objects from her surroundings, not just all scarlet objects.

So here's one possible answer to the question of why the more proportional property is to be preferred as a cause: It provides a better or more useful explanation of the effect. This, however, raises a new question: Why think that a preference for the more proportional property as an *explanation* of an effect has any bearing on the *causal* status of any determinate of this property? Is it not entirely possible that one property provides the most useful causal *explanation* of some effect, while it is just one of many sufficient causes of that effect? As far as I can tell, the answer is yes – *unless*, that is, the original exclusion principle (the principle that, in general, an effect has no more than a single sufficient cause) is considered a valid constraint on metaphysics. In this case, preferring one property over another as a cause of an effect comes with a heavy metaphysical commitment – it commits one to exclude all other properties from being causally efficacious with respect to that effect. In other words, Yablo may want to deny that a property bestowing a power that

[115] Yablo [1992], p. 277.

manifests is sufficient for causation, since it leads to systematic causal overdetermination. This worry, however, is unmotivated if we accept that systematic weak overdetermination is unproblematic, since this is the kind of overdetermination that follows from accepting the causal efficacy of a determinable property and each of its corresponding determinates (recall, the argument is that the property making out the grounding base, for instance a determinate property, inherits the powers of the determinable property it grounds. So the very same powers are bestowed by several properties). To sum up then, if it is accepted that satisfying the proportionality constraints is what justifies the claim that a determinable property has powers to cause certain effects (because this property would be proportional to these effects), and accepted that its corresponding determinate inherits these powers by grounding the determinable property, then I see no reason to insist that the determinate property fails to be a cause of the effects also caused by the determinable property, despite the fact that the determinate property fails to satisfy the proportionality constraints.

While the focus above is on determinable properties and their determinates, the points apply to mental properties and their physical realizers as well. Say that a mental property *M* satisfies the proportionality constraints with respect to some behavioural effect E. Assuming physicalism, this mental property must have a property *P* as its grounding base, say a structural neural property, which thereby inherits the power to cause E from *M*.[116] And when this power manifests, we should not be afraid to insist that both the mental property and the structural neural property are causally efficacious with respect to E. Of course, this is compatible with thinking that the mental property is a better causal *explanation* of E. As promised earlier, we can therefore dispel the worry that the Closure principle comes out false on the account that causes are proportional to their effects. Even if a mental property satisfies the proportionality constraints with respect to some effect and its physical realizer does not, the physical realizer still inherits the powers of the mental property (since the physical realizer ground the mental property), and so,

[116] Of course, *P* need not, and almost certainly will not, make out the complete grounding base for *M*. At one point, Shoemaker makes a distinction between core and total realizers: "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". We can make the same distinction with respect to core and total grounding bases. To use this terminology, *P* will likely just be a core grounding base for *M*, not a total grounding base.

whenever the mental property is causally efficacious (i.e. whenever one of its powers manifest), the physical realizer will be causally efficacious as well (since it bestows the very same causal powers).

4.11 Is the Subset View Committed to the Proportionality of Causes?

Considering the argument above to the effect that the Subset view is motivated if we accept that causes are proportional to their effects and that properties inherit the causal powers of the properties they ground, we might wonder if this means that the proponent of the Subset view *must* accept the proportionality claim. McLaughlin thinks so, and since he is inclined to reject it, this means that he is inclined to reject the Subset view.[117]

There are a couple of reasons why one might think that the proponent of the Subset view is necessarily committed to accepting the proportionality constraints. One (and this seems to be McLaughlin's reason), without them, the Subset view leads to unacceptable causal overdetermination. The idea, I take it, is something like the following: The Subset view holds that determinable/mental properties can be causally efficacious. It *seems*, however, that whenever a determinable/mental property is causally efficacious with respect to some effect, its underlying determinate property/physical realizer is a distinct sufficient cause of this effect, which would be unacceptable. The only way to solve this problem is by appeal to the proportionality constraints which not only generate unique causes, and so deny that there is genuine causal competition, but also award the causal victory to the determinable/mental property. Hence, the proponent of the Subset view is necessarily committed to the proportionality claim. However, as should be clear from the previous discussion, I do not consider this a sound argument. First of all, it assumes that the original exclusion principle is a valid constraint on a metaphysical theory, and appeals to it to establish that the proponent of the Subset view must necessarily accept the proportionality claim. But as I argued in the previous chapter, the original exclusion should be rejected. Second of all, as I argued above, there may be other reasons, besides the fact that a property satisfies the proportionality constraints, for thinking that a property has powers to cause an effect. For instance,

[117] McLaughlin [2007], p. 165.

when one property grounds another property that satisfies these constraints with respect to some effect, the former plausibly inherits the power to cause this effect from the latter. In fact, as we saw, it seems that the proponent of the Subset view *must* endorse something like this, since the Subset view itself will otherwise be undermined. This means that adopting the proportionality claim does not even prevent the Subset view from leading to causal overdetermination. Of course, since it leads only to weak overdetermination, this is no reason for concern.

A second reason for thinking that the proponent of the Subset view is necessarily committed to the proportionality claim would be if one thinks that the proportionality constraints provide the *only* way to motivate that a class of realized properties have causal powers. Recall, the argument from earlier which motivates the Subset view – first step is to get the causal powers of the realized property with the help of the proportionality constraints, second step is to argue that these causal powers are inherited by any property that grounds this realized property. If the proportionality constraints are rejected, the worry is that the first step cannot be carried out. And if there are no other ways to get the causal powers of realized properties, the Subset view fails since it states that the causal profile of any realized property is non-empty. However, it is doubtful if the proportionality constraints, in general, are needed to motivate the claim that a class of realized properties have non-empty causal profiles. Consider for instance mental properties. When motivating the claim of Mental Causation (the claim that mental properties are sometimes causes), I did not appeal to any specific conditions on what is to count as a cause. Rather, Mental Causation was established pre-theoretically and did not need technical philosophical arguments about the nature of causation to be vindicated. I do not wish to rule out that the proportionality constraints can provide justification for the existence of causal powers of some properties. For instance, it may be that the reason we are inclined to think that determinable properties have non-empty causal profiles is that they satisfy the proportionality constraints with respect to some effects. However, there may still be other ways of establishing these causal profiles and in any case, the proportionality constraints are not needed in the case of the causal profiles of mental properties. For this reason, the proponent of the Subset view should not feel forced to accept the proportionality claim.

4.12 The Proportionality Constraints and the Exclusion Problem

So where does this leave the proportionality claim with respect to the Exclusion problem? Recall, Shoemaker indicates that they help the proponent of the Subset view solve this problem, as they "make it unmysterious that the causal efficacy of a multiply realized property is not preempted by that of its realizer properties".[118] However, as far as I can tell, the proportionality constraints actually do little work here. As we have seen, they do not motivate the claim that a realized property has a non-empty causal profile, at least not in the case of mental properties. If anything, they threaten to make it mysterious why the efficacy of the realizer is not pre-empted by the efficacy of the realized property – if a property must satisfy the proportionality constraints with respect to some effect to be a cause of that effect, it seems that the realizer is excluded from causing an effect whenever the property it realizes satisfy the proportionality constraints. Of course, as I've just argued, there is no real mystery here. However, the mystery is not dispelled by any appeal to causes being proportional to their effects. Rather, what is doing the work is the fact that the causal powers of the realized property are identical to powers had by the realizer. This means that whenever the realized property is causally efficacious (i.e. whenever one of its power manifest), the realizer will be causally efficacious as well. But the claim that the powers are realized properties are identical to powers of the properties that realize is already at the heart of the Subset view. No further assumptions about the nature of causation are therefore needed to provide this solution to the Exclusion problem. In the end, this is all the better for the proponent of the Subset view. The proportionality claim is a controversial one and it is a virtue of the Subset view that this claim does not come as part of the package its proponents offer, and that the claim is not needed to solve the Exclusion problem.

4.13 The Argument from Parts and Wholes

Putting aside the argument from the proportionality of causes, let me return to Shoemaker's other argument for why the Subset view solves the Exclusion problem. Recall, this turns on an analogy between the realization relation and the part-whole relation. Roughly, the argument goes as follows: In a sense, realized properties are parts of their realizers. Parts and wholes do not overdetermine their

[118] Shoemaker [2001], p. 93.

common effects. Hence, a realized property and its realizer do not overdetermine their common effects.[119]

One worry about this argument is that the analogy between, on the one hand, mental properties and their physical realizers, and on the other, parts and wholes, is somewhat metaphorical. Mental properties and their physical realizers do not literally stand in a part-whole relation, rather it is their profiles that do so.[120] So even if one agrees that there is no Exclusion problem for parts and wholes, it might be argued that this tells us little about the Exclusion problem as it pertains to mental *properties* and their realizers.

There is another worry though, that, when addressed, helps to dissolve the first worry as well. This worry is that it is controversial to simply assume that parts and wholes do not overdetermine their effects. As I argued in the previous chapter, exclusion worries naturally arise for composite properties and objects since reduction of higher level properties to lower level properties is problematic (primarily due to the possibility of there being no fundamental bottom level). But if part-whole reduction is ruled out and we assume that a composite property or object is not causally excluded by the sum of its parts or *vice versa*, then parts and wholes *do* overdetermine their common effects, in the trivial sense that they are numerically distinct causes of the same effects. The solution I proposed to this problem was to accept that parts and wholes weakly overdetermine their common effects and that the considerations that lead us to adopt the Exclusion principle fail to rule this out. In other words, allowing for weak overdetermination is the key to resolving the Exclusion problem for parts and wholes. But if this is correct, Shoemaker's argument cannot be quite right. Recall, weak overdetermination involves the manifestation of a power shared by several properties. So the fact that parts and wholes share the powers to cause their common effects is what prevents the Exclusion principle from kicking in. Now, Shoemaker tries to dissolve the Exclusion problem for mental properties by arguing that mental properties stand in a part-whole relation according to the Subset view. They do this because the powers

[119] Shoemaker [2007], p. 13.
[120] If properties are construed as being identical to sets of powers, rather than being the entities that bestow these powers, then realized properties would literally be parts of their realizer. This, however, is not Shoemaker's view, nor is it a view I consider here.

of a realized property are proper parts of the powers of its realizer, which means that they share the powers that are powers to cause their common effects. But the reason there is no Exclusion problem for parts and wholes is exactly that they share the powers that are powers to cause their common effects. So the fact (if it is a fact) that mental properties and their realizers stand in a part-whole relation, is in itself not what helps to solve the Exclusion problem for mental properties, as Shoemaker would have it. Instead, the two Exclusion problems, one for mental properties and their realizers, one for parts and wholes, have a common solution – they are both solved by the fact that the causally competing properties share the powers to cause their common effects. As promised, this also dispels the worry that the analogy to parts and wholes breaks down since mental properties do not literally stand in a part-whole relation to their physical realizers. As it turns out, whether or not this is so is irrelevant as far as a solution to the Exclusion problem is concerned. What matters is that the properties competing to cause an effect share the relevant power.

To sum up then, Shoemaker underplays the potential of the Subset view to solve the Exclusion problem. There is no need to rely on controversial assumptions about the nature of causation or potentially dubious analogies to parts and wholes. The resources needed to solve the Exclusion problem are already part of the Subset view and how it construes the relation between realized properties and their realizers.

4.14 The Subset View and Physicalism

Early in his book, Shoemaker makes the following claim about properties that are not physical, he says the following:

> […] if physicalism is true, all of these properties must in some sense be determined, constitutively rather than just causally, by physical properties or physical states of affairs. […] we can express the determination claim by saying that instantiations of the properties in question are realized in the instantiation of physical properties of some sort or in physical states of affairs of some sort.[121]

I take it that by 'constitutive determination', Shoemaker has in mind a relation along the lines of the grounding relation introduced in chapter 1. We can now see

[121] Shoemaker [2007], p. 1.

why the realization relation cannot simply be identified with the subset relation. The reason is that the subset relation in itself is neutral on whether properties that have the subset of powers are grounded in the properties that have the superset of powers - it is entirely consistent to claim that mental properties stand in a subset relation to some physical properties while denying that the former are grounded in the latter.[122] If the quote from Shoemaker above is correct, however, the realization relation is *not* neutral on whether realized properties are grounded in their realizers. If it was, realization could not express grounding (or 'constitutive determination').

That the subset relation is neutral on issues of metaphysical dependence is, of course, to be expected if the grounding relation is primitive and not reducible to any other relation. Because of this, a claim of grounding between two kinds of properties will always be in addition to any other claim about the relation one takes to hold between the two, for instance a claim that the properties in question stand in a subset relation. Given this (and assuming that grounding and realization is not the very same relation[123]) we might wonder how realization can be an expression of grounding. It can be so only if the realization relation breaks down into several necessary but individually insufficient conditions, one of them being the grounding condition. We already know that the subset relation is another (per Shoemaker's stipulation). And it seems plausible that the two are jointly sufficient for realization on Shoemaker's account. At least, the only properties standing in the subset relation that Shoemaker excludes as also standing in the realization relation are conjunctive properties and this is arguably on the basis that conjunctive properties intuitively do not ground their conjunct properties.

With this in mind, we might ask how the Subset view of realization relates to the general thesis of physicalism. One thing to note is that, like physicalism, the Subset

[122] As mentioned earlier, Shoemaker takes conjunct properties and the conjunctive properties they are conjuncts of, to stand in a subset relation without standing in a realization relation. While Shoemaker is not explicit about what justifies this exception, we can now see that it is motivated in so far as it is intuitively false that conjunctive properties ground their conjuncts. Rather, the metaphysical dependence is the other way around – a red square object has the property *being red and square* in virtue of having the properties *red* and *square*, not *vice versa*. But if conjunctive properties do not ground their conjuncts, then they cannot realize them, despite the two standing in a subset relation.

[123] An uncontroversial assumption, I take it, since realization is exclusively an intra-level relation between properties. The grounding relation can almost certainly hold between other kinds of entities and between entities at different levels.

view can be restricted to a certain class of non-physical properties and be the view that the properties in that specific class are realized by physical properties in accordance with the definition of realization given earlier. In this case, the Subset view is compatible with the existence of additional anti-physical properties that falsify unrestricted physicalism. But even so, there's a further question of whether or not the Subset view is a form of physicalism about the properties it claims are realized by physical properties. Earlier (chapter 1, fn. 30), I mentioned that physicalism about a restricted class of non-physical properties is best understood as the claim that these properties are ultimately grounded in nothing but fundamental non-mental physical properties.[124] In other words, physicalism about realized properties is true if and only if there is a chain of metaphysical dependencies, or "grounding-links", that connects these realized properties with fundamental physical properties, and none of the latter properties are mental. Since, per definition of the realization relation, there is such a link going from realized properties to their realizers, it follows that physicalism about realized properties is true if and only if there is a chain of grounding-links from the realizers of realized properties to fundamental physical properties, and that none of the latter are mental. Or simply: Physicalism about realized properties is true if and only if physicalism about their realizers is true. On the Subset view, realizer-properties in our world are physical properties, though typically they will be broadly, rather than fundamentally, physical. For instance, the realizers of mental properties will presumably turn out to be complex neural properties. Strictly speaking, physicalism may fail for such broadly physical properties, either because there are emergent causal laws in virtue of which these broadly physical properties have causal powers not grounded in the powers of the lower level fundamental properties of which they are composed, or because there are mental properties among the fundamental properties that compose them. So the Subset view is a form of physicalism about realized properties, only if it is assumed that this is not the case. In this respect, however, the Subset view is in the same boat as other physicalist positions. Take reductionism about mental properties, for instance, which holds that mental properties are identical to physical properties. Again, the physical properties that make out that the reduction base for mental properties are typically taken to be

[124] If there is no fundamental level, the counterpart view would be that the class of properties in question will ultimately be grounded in an infinite line of scientific non-mental properties.

broadly physical neural properties. Formulated in this fashion, reductionism is entirely compatible with the failure of physicalism since physicalism may fail for broadly physical properties. Yet, this is usually a possibility that is put aside - the assumption that physicalism is uncontroversially true for broadly physical properties is almost always taken for granted.

Having seen how the Subset view can be a form of restricted physicalism about realized properties, we can move from this to unrestricted physicalism or physicalism *simpliciter* by adding to the Subset account of realization that *all* properties not physical (either broadly or fundamental) are realized by physical properties.

4.15 The Problem of Backward Looking Powers

This conclusion that the Subset view is a form of physicalism relies on the stipulation that realization entails grounding. It also, per stipulation, entails that realized properties and their realizers stand in the subset relation. On Shoemaker's most recent account of the subset relation, however, there is a tension between these two necessary conditions for realization – if two sets of properties stand in the subset relation, they cannot stand in the grounding relation. I will briefly explain why that is and how the problem can be solved.

As I mentioned earlier, Shoemaker's current view is that a realized property has forward looking causal powers that form a proper subset of the forward looking causal powers of its realizer *and* backward looking causal powers that form a superset of the backward looking causal powers of the realizer. Shoemaker's motivation for introducing the notion of backward looking causal powers is based on the following example (which Shoemaker attributes to Richard Boyd): Say that two chemical compounds X and Y have different constituents, but nevertheless have exactly the same forward looking causal powers. That is, X and Y have powers to cause exactly the same effects. According to Shoemaker, it would be counter-intuitive to say that *being compound X* and *being compound Y* is the same property since the ingredients from which the compounds are made differ. So forward looking causal powers cannot be all that individuates a property. What distinguishes X and Y are instead their backward looking causal powers – their

causal origins. X is caused by mixing substances A and B; Y is caused by mixing substances C and D.[125]

This connects with the Subset view in the following way: If a realized property is multiply realizable, it can be caused in more ways than any of its particular realizers. For instance, say that you have two cans of spray-paint, one with scarlet paint and one with crimson. You can cause an object to be red by spraying it with either can, but you can only cause it to be scarlet by spraying it with the can containing scarlet paint. This is why the set of backward looking causal powers at the level of realized properties, will be a superset of the set of backward looking causal powers at the level of their realizers. If this is so, however, it seems that realized properties fail to supervene on their realizers since the instantiation of a specific realizer *R* is compatible with the instantiation of several distinct realized properties with different backward looking causal powers. All that the subset relation requires is that these properties have the same forward looking causal powers (which form a subset of the powers of *R*) and that they share the backward looking causal powers that form the subset that contains exactly the backward looking causal powers of *R*.[126] It does *not* rule out that each of these realized properties have their own distinctive backward looking causal powers that they do not share. Hence, it seems possible that a minimal physical duplicate of our world would fail to include the realized properties of our world but instead include different realized properties that differ from the actual realized properties with respect to some of their backward looking causal powers. But as was argued in chapter 1, if the realized properties of our world are grounded in their realizers, this entails that they supervene upon them. Since it now seems that they do not, we can infer that realized properties are not grounded in their realizers in our world. This, however, means that, contrary to Shoemaker's claim, the realization relation cannot express the kind of metaphysical dependence of the non-physical on the physical that the physicalist is looking to endorse – the stipulative definition of realization is

[125] Shoemaker [1980], p. 267.
[126] This point is raised by McLaughlin in his [2007].

unstable if the subset relation is taken to include both backward and forward looking causal powers.[127]

One way to counter this would be to argue that there are no realized properties that have the same forward looking causal powers but which differ with respect to backward looking causal powers. Of course, this cannot merely be the claim that, as a matter of *actual* fact, there are no such properties. Rather, it would have to be the much stronger claim that such properties are impossible. It is hard to see how such an argument could be made once it is admitted that the compounds brought about by mixing different substances are indeed distinct, despite them having exactly the same forward looking causal powers. It is one thing to insist that properties in general are individuated purely in terms of their forward looking causal powers, it is another to admit that some are not (*being compound X* for instance), while insisting that other must be (realized properties). A simpler solution to the problem would be to go for the former claim and back off from the claim that *being compound X* is a different property from *being compound Y*. And really, this is hardly a major bullet to bite. After all, *being compound X* and *being compound Y* share *all* their forward looking causal powers. That is, once the substances are mixed, there is no possible way of telling the two compounds apart without independent information about their causal origin. It is thus unlike a case of, say, gold versus fools gold where a sample of each may share all of their superficial properties but where expert analysis allows us to distinguish between them. In light of this, it does not seem to me counter-intuitive to insist that *being compound X* and *being compound Y* are the very same property, despite the fact that their causal origins differ.[128]

[127] It should be noted that McLaughlin raises the issue in a different context. In short, McLaughlin thinks that Shoemaker needs the claim that realized properties supervene on their realizers in order to solve the Exclusion problem for mental properties. Hence, McLaughlin's worry is that if supervenience fails, Shoemaker's view fails to deal with this problem. I personally do not think that the solution offered by the Subset view hinges on realized properties supervening on their realizers. As I argued earlier, the crux of the solution offered by the Subset view is that behavioural effects are the manifestation of powers had by both mental properties and their physical realizers. This solution would be available even if realized properties failed to supervene on their realizers, as long as realized properties and their realizers overlap with respect to the powers that are powers to cause behavioural effects. The main worry, as I see it, is rather that the Subset view fails to be a form of physicalism if the relevant supervenience fails.

[128] Shoemaker has stated (in personal communication) that he now accepts that the subset relation is a relation that holds only between the forward looking causal powers of the properties involved.

Without the claim that properties are individuated, in part, by their backward looking causal powers, supervenience of realized properties on their realizers can be restored. Since the forward looking causal powers of realized properties form a subset of the powers of their physical realizers, any physical duplicate of our world will include the properties that have exactly the powers in that subset. And this, in turn, allows for realized properties to be grounded in their realizers.

4.16 The Subset View and Functionalism about Mental Properties

Obviously the Subset view is committed to there being sets of powers that are causal profiles of mental properties. Say that *R1* and *R2* are putative physical realizers for mental property *M*. If so, and given the Subset view, *R1* and *R2* must overlap with respect to some powers that form a subset of each of their causal profiles. But why think this subset of powers is a causal profile for *any* property? Or to out it another way, why think there is any property that is the occupant for this set of powers. After all, it seems that it might just be a subset of powers that are had by *R1* and *R2* and not by any other property. Shoemaker answers this question by adopting a certain form of functionalism. The causal profiles of mental properties are given by folk psychological causal definition. A set of powers, with respect to which the causal profiles of *R1* and *R2* overlap, is a causal profile for a unique property *M* because folk psychology defines *M* as being the property that has *exactly* those powers. Shoemaker's functionalism thus distinguishes itself from traditional formulations of both role and filler functionalism.[129] Roughly, role functionalism is the view that mental properties are second-order properties, defined in terms of causal roles that are then played or occupied by their first-order realizer. To use a crude but common example from the literature, *pain* may be defined as the second-order property of having some first-order property that causes subjects to exhibit aversive behaviour under certain conditions (for instance, then they are subjected to tissue-damage). A first-order property that plays this role is a realizer of pain. Since the definition of *pain* (and mental properties in general) poses no restrictions on the number of first-order properties that can play the relevant role, role functionalism allows for multiple realization. It is unclear, however, how it allows for mental properties to have causal powers since it is part of the description

[129] See McLaughlin [2006] for a comprehensive discussion of the differences between role and filler functionalism.

of a mental property that the causal work is done by its first-order realizer – mental properties are described in terms of causal tasks performed by *other* properties. If so, the notion of realization as it is employed by the role functionalist and the proponent of the Subset view cannot be the same – if mental properties are epiphenomenal, their powers cannot form a proper *non-empty* subset of the powers had by some physical property.

Filler functionalism contrasts with role functionalism in that the filler-functionalist definitions of mental properties explicitly assign causal powers to them. For instance, rather than pain being the property of having some *other* property that causes aversive behaviour, pain is defined, on the filler functionalist account, as being the property that itself causes aversive behaviour. So far this is in line with Shoemaker's view. Shoemaker, however, is not a traditional filler functionalist. The reason is that filler functionalists claim that the functional descriptions of mental properties provide grounds for *identifying* mental properties with physical properties. If pain is the property that causes aversive behaviour and our best neuroscience informs us that neural property *N* is the property that causes aversive behaviour, then we can infer that pain is identical to *N*.[130] One thing to note is that, due to the possibility of multiple realization, filler functionalists regard the traditional functional description of pain (and other mental properties) as improper since there is no *unique* property that causes aversive behaviour. Hence, there is no unique reduction-base for mental properties. This problem is dealt with by replacing generic functional descriptions of mental properties with descriptions relative to certain structures or species. The description 'the property that causes aversive behaviour' may not refer since it fails to pick out a unique property among all actual and possible properties. However, it may succeed in picking out a unique neural property among, say, the properties instantiated by members of the human species. In this case, the property that causes aversive behaviour in humans (call it 'pain-in-humans') can be reduced to this neural property.[131] This also means that, like role functionalism, traditional filler functionalism is incompatible with the Subset view

[130] We find this argument in Lewis [1972].
[131] As we saw in the previous chapter, this is essentially also Kim's solution to the problem of multiple realization that reductionists face.

since the causal profiles of mental properties are not *proper* non-empty subsets of the causal profile of some physical properties.

The difference between Shoemaker's functionalism and traditional filler functionalism is that Shoemaker takes functional definitions to give genuine causal profiles for mental properties. Filler functionalists, on the other hand, take functional definitions to give only *part* of a causal profile, the causal profile that is the profile of the physical property that is the reduction-base for the mental property picked out by the functional definition. In other words, on the filler functionalist account, there is no occupant for the set of powers given by the functional description, since there is no property that corresponds to all and *only* those powers (recall, on my definition, a property is an occupant of a causal profile if and only if it has all and only the powers that make out that profile). To see this, consider again the case of *pain* (or *pain-in-humans*) and its reduction-base neural property *N*. Obviously the functional description of *pain* will be a great deal more complex than merely stating that it is the property that causes aversive behaviour. But whatever the complete functional description will turn out to include, it will surely not include reference to *all* the powers of *N*. For instance, *N* can cause certain other neural properties and will show up in a certain way on a CAT scan, none of which are powers that will be part of the functional description of the mental property *pain*.[132] So if *pain* is identical to the property that causes these effects, then the functional description does not provide a causal profile for *pain* (or any other property for that matter) but merely *part* of a causal profile. Not so on Shoemaker's account. Here the functional analysis gives a causal profile – the profile for the relevant mental property. Of course, it also gives part of a causal profile – part of the causal profile of the physical property that realizes it. But this is simply to repeat that the powers of the mental property form a subset of the powers of its realizer.

It might be argued that the functional description of mental properties will be improper on Shoemaker's account as well since there is no unique property that causes aversive behaviour. Rather, it is caused by pain *and* all the different realizers

[132] At least not if the functional descriptions are provided by common-sense folk psychology.

of pain. Going for structure-specific functional descriptions is of no help here. Even if the functional description manages to pick out a unique *physical* property that causes aversive behaviour among members of a certain species, the same effect will also be caused by a distinct non-physical property *pain*, given that Shoemaker denies psycho-physical reduction. What is needed to ensure a referent for the functional description is rather a stop-clause after the causal roles. For instance, rather than *pain* being the property that causes aversive behaviour, it is the property that causes aversive behaviour *and nothing else* (again, the complete functional description will obviously by much more complex). Understood in this manner, folk psychological functional descriptions pick out unique properties and give causal profiles with mental properties as occupants. They no longer pick out the physical realizers since these have causal powers beyond the powers referenced in the functional descriptions.

4.17 Must a Proponent of the Subset View Accept Shoemaker's Functionalism?

As we saw, both role and filler functionalism, in their traditional forms at least, are incompatible with the Subset view. But what is the relationship between Shoemaker's functionalism and the Subset view? Does one support the other or is Shoemaker endorsements of functionalism and the Subset view entirely independent? In his discussion of the Subset view and problem of mental causation[133], McLaughlin includes Shoemaker's functionalism as part of the package that comes with the Subset view. If one is sceptical about the prospects for functionalism, including the modified kind of filler functionalism that Shoemaker endorses (as McLaughlin is), then this might be seen as a reason for being sceptical about the prospects of the Subset view. I will argue, however, that the Subset view is motivated independently of arguments for functionalism and that there is no reason to think that one position supports the other.

Firstly, note that while Shoemaker may endorse that we get the causal profiles of mental properties from folk psychological definitions, he presumably does not endorse the claim that the causal profiles of realized properties in general are given

[133] McLaughlin [2007].

by folk definitions. This should be clear from the fact that Shoemaker takes determinable properties to be realized by their determinates. Surely there are determinable properties for which it is quite implausible that folk definitions give the causal profile, for instance, highly scientific kinds such as *being an enzyme* or *having wave impedance*. Since Shoemaker takes the Subset view to hold for all realized properties, it cannot be the case that the kind of functionalism he endorses for mental properties is required for the Subset view in general (as I suggested earlier, it may be the case that the causal profile of determinable properties is determined (in part at least) by applying the proportionality constraints to them).

4.18 Motivating the Subset View Independently of Functionalism

Still, it might be thought that while the truth of functionalism (of the kind endorsed by Shoemaker) is not strictly speaking required for the truth of the Subset view, the former still provides a strong reason, perhaps even the main reason, for thinking that the Subset view is true for *mental* properties. If so, denying functionalism would remove a strong, perhaps the main, motivation for accepting the Subset view for mental properties. I do not share this worry since I believe the Subset view is motivated independently of any functionalist commitments. To see how, consider the following four claims that sum up the Subset view (or at least a central part of it):

i) Mental properties have powers to cause physical effects.

ii) These causal powers are also had by physical properties.

iii) These physical properties have additional causal powers, not had by mental properties.

iv) The powers shared by mental and physical properties are identical at both the type and token level. That is, the token-powers had by an instance of a mental property are numerically identical to token-powers had by some instance of a physical property.

And now consider this slightly reformulated version of the Exclusion problem.

i*) There is mental-to-physical causation (Mental Causation).

ii*) Every physical effect has a sufficient physical cause (Closure).

iii*) Mental properties are distinct from physical properties (Distinctness).

iv*) Behavioural effects are not systematically and strongly overdetermined (Exclusion).

When put in this way, there is a striking similarity between the four claims constituting the Subset view and the four claims constituting the Exclusion problem. In fact, as I will proceed to argue, each of the claims constituting the Subset view is entailed by each of the claims constituting the Exclusion problem, at least given certain assumptions a physicalist is likely to accept. If so, the reasons for accepting each of the four claims constituting the Exclusion problem are indirectly reasons for accepting the Subset view.

First, take the claim that there is mental-to-physical causation. A property causing an effect E involves the manifestation of one of its powers (namely the power to cause E). So if there is mental-to-physical causation, then mental properties must have powers to cause physical effects – i) is entailed by i*).

Second, the Closure principle tells us (roughly) that every physical effect has a sufficient physical cause. Any physical effect E caused by a mental property *M* (at time *t*) must therefore also be caused by a physical property *P* (at *t*). But if *P* is a cause of E, then *P* must have the power to cause E as well. Hence, for any power *M* has to cause some physical effect E, there is a physical property *P* which also has that power – ii) is entailed by ii*).

Third, Distinctness tells us that mental properties are distinct from physical properties because the former can be multiply realized. If so, this means that any mental property *M* must have a causal profile that is different from the causal profile of any physical property. Since the powers of any mental property *M* are also had by some physical property *P*, *M* and *P* must be distinct, either because *M* has additional powers not had by *P* or *vice versa*. As I argued earlier, the former cannot be the case – powers had by *M* to cause non-physical effects (for instance mental effects) are really powers to cause the physical grounding base for these non-physical properties and *M* cannot have powers to cause physical effects not had

by any physical property since this would violate the Closure principle. It follows that if Distinctness is true, it must be the case that the physical properties that have the powers also had by mental properties, have additional causal powers – iii) is entailed by iii*).

Fourth, the Exclusion principle tells us that effects are not systematically and strongly overdetermined. On a non-reductive physicalist account, the effects of mental properties are systematically overdetermined by mental and physical properties (every time a power bestowed by a mental property is manifested, the Closure principle tells us that a power bestowed by a physical property must manifest as well). If so, they cannot be strongly overdetermined, i.e. they cannot be the manifestation of two token-distinct powers. So the causal powers contained in the profile of a mental property must reduce to powers contained in the profile of some physical property at both the type and token level – iv) is entailed by iv*).

Since each of the four claims constituting the Subset view is entailed by one of the four claims constituting the Exclusion problem, the justification for the latter claims transmits to the former. And since we are assuming that the four claims constituting the Exclusion problem are individually motivated, this means that the Subset view is motivated. Note that since the four statements that constitute the Exclusion problem are motivated independently of functionalist commitments, this means that the Subset view is motivated independently of such commitments.

4.19 Kim on the Subset View

Some might object to the dialectic of an argument that establishes the Subset view with the help of the statements constituting the Exclusion problem. Kim, for instance, has recently made the following point against the Subset view:

> Consider a mental property M. How does M get to have a physical property, P, as one of its realizers? According to the subset definition, the causal powers of M must be a subset of the causal powers of P. How is that possible? We may assume that most of P's causal powers are powers to cause other physical events but we can allow, at this point, that P's causal powers may include causal powers to cause nonphysical events as well. But for the present strategy to work for the mental causation problem, the causal powers of M must include at least some of P's physical causal powers. This

> amounts to the supposition that M has causal powers to cause physical events. How do we show that? Well, showing that that is possible, or showing how that is possible, is exactly the problem of mental causation. We seem to be back to square one, and very quickly, in a small circle![134]

In other words, Kim is objecting that the Subset view is assuming what it sets out to show, namely that there is mental causation. As such, the claims that constitute the Exclusion problem, in particular MC (the claim that there *is* mental causation), cannot function as evidential basis for the Subset view. Instead, the explanation should be reversed – MC should fall out of the Subset view, which requires that the view is motivated independently of MC. Kim makes the point even clearer elsewhere:

> For a positive resolution of the mental causation problem, I believe we should reach the statement that mental properties are causally efficacious as a conclusion, not start with it as an assumption. What troubles me is that Shoemaker's procedure seems to be the opposite; for him, mental causal efficacy is a "starting assumption".[135]

I believe Kim's point should be resisted here, not just by the proponent of the Subset view but by anyone interested in solving the Exclusion problem. First, consider Kim's claim that a solution to the Exclusion problem (or as Kim calls it, the problem of mental causation) should have the statement that mental properties are causally efficacious, or MC for short, as a conclusion. This seems unreasonable if the Exclusion problem is a paradox with MC as one of its claims. A paradox can be solved (or dissolved) by rejecting one of the mutually inconsistent claims constituting it, or by showing that the inconsistency between the claims is merely superficial or apparent, typically by disambiguating one of the claims. One does not, however, solve a paradox by giving an account or argument that has one of its claims as a conclusion. While such an account might help to strengthen the motivation for accepting that particular claim, it does not help to solve the paradox of which that claim is part. Indeed, the Exclusion problem is a paradox *only* if the claims that constitute it, including MC, are already motivated. So while Kim is correct in saying that MC is a starting assumption for the Subset view, this should

[134] This is from an email correspondence published online at: http://brainbrain.blogspot.com/2006/07/kim-vs-subset-view-of-higher-level.html
[135] Kim [2010], p. 107.

not be held against the view – in fact, any theory that promises to provide a solution to the Exclusion will have already assumed that we have good reasons for thinking that MC holds true. Consider for instance what Kim himself says in earlier writings:

> [..] I doubt that very many of us who have been "worried" about mental causation have actually been concerned about the possibility that our thoughts and desires might turn out to have no powers to move our limbs. […] [*The Exclusion problem*] is the problem of showing *how* mental causation is possible, not *whether* it is possible. […] In raising the how-question, we are assuming, "defeasibly, but firmly" as Burge says, that the whether-question has already been affirmatively answered.[136]

I bring this up, not to make an *ad hominem* point against Kim, but because it strikes me as exactly the right thing to say about MC, or any other statement that is part of a paradox.

4.20 Motivating Reductionism over Non-Reductionism

Once it is accepted that the statements that make up the Exclusion problem are already motivated and that they are all, in that sense, "starting assumptions" for any view that is concerned with solving the Exclusion problem, then it becomes clear that they can provide evidential basis for mind-body theories that incorporate the claims that constitute the Exclusion problem, or other claims that are entailed by them. Of course, this is not to say that none of them can be given up. It may be that there are overruling reasons for thinking that one of the claims should be rejected. For instance, early arguments for psycho-physical reduction were not primarily based on worries about the possibility of mental causation, but were rather arguments from ontological parsimony.[137] If one is convinced by such arguments, then one may prefer to reject Distinctness in any case, whether or not this is necessary to account for mental causation. However, given that the Exclusion principle must be amended to allow for weak systematic overdetermination (as the argument in previous chapter showed) and given that this, in turn, allows for the Exclusion problem to be solved without rejecting any of the claims that constitute it, it is crucial that an argument to the effect that one of the claims of the Exclusion problem should be rejected must be *independent* of wanting to solve this problem.

[136] Kim [1998], p. 61-62.
[137] As an example, see Smart [1959].

Rejecting a *prima facie* plausible statement S, on the basis that one wishes to maintain statements P, Q, and R, is a sound move only if S is incompatible with P, Q, and R. If it is not, then one needs to provide independent reasons to doubt S in order to reject it. As far as rejecting the claim that mental and physical properties are distinct, however, I think it is fair to say that proponents of the Exclusion argument have not done much to provide such arguments. Kim, for instance, is explicit about the fact that he takes the Exclusion argument to be the only real argument for reduction of mental properties.[138] And Braddon-Mitchell and Jackson make the point that if the choice is between a reductive form of functionalism (filler-functionalism) and a non-reductive one (role-functionalism), the fact that the former allows for the causal efficacy of mental properties is what settles the issue in favour of this position.[139] In both cases, reductionism is established as superior to anti-reductionist alternatives on the grounds that the former allows for mental causation. If what I've argued in this and the previous chapter is sound, this is a poor argument for reductionism since the Subset view allows for mental causation as well and denies that mental properties are reducible to physical properties.

What of Kim's proposal to replace generic mental properties, such as *pain*, with structure-specific mental properties such as *pain-in-humans*, *pain-in-dolphins*, *pain-in martians*, and so on? I take it that were it not for the fact that this replacement paves the way for psycho-physical reduction, nobody would have ever seriously proposed it. This brings us back to the question of what motivates reduction. If reductionism is attractive for reasons that have nothing to do with mental causation (for instance, considerations about ontological parsimony) then this may provide motivation for accepting the claim that there are only structure-specific mental properties. And *by implication*, this solves the Exclusion problem. However, in so far as psychophysical reduction is motivated by a desire to solve the Exclusion problem, replacing generic mental properties with structure-specific mental properties fail to be motivated independently of wanting to solve the Exclusion problem. As such the motivation for accepting structure-specific mental properties is a poor one since we don't need to accept those to solve the Exclusion problem.

[138] Kim [2005], p. 125.
[139] Braddon-Mitchell & Jackson [1996], p. 101.

Could a similar objection not be raised against the amendment to the Exclusion principle that I proposed in the previous chapter? Recall, the claim there that was the original exclusion principle is ambiguous as it does not distinguish between strong and weak overdetermination, and that once the original exclusion principle is appropriately amended to prohibit only strong systematic overdetermination, the Exclusion problem can be solved. Now, perhaps the reductionist could argue that this amendment is equally motivated by a desire to solve the Exclusion problem. If so, it seems the reductionist is at least as well off since we just have two different proposals to amend different statements of the Exclusion problem – the reductionist proposes an amended version of the Distinctness (amended from the claim mental properties are distinct from physical properties to the claim that structure-specific mental properties are distinct from physical properties) which they then go on reject (in the grounds that the argument from multiple realization fails to go through), and I propose an amended version of the exclusion principle (which prohibits strong systematic overdetermination but allows for weak systematic overdetermination) which I then argue is compatible with the other statements of the Exclusion problem (and result in the Subset view). However, there is an important difference between the two proposals since I offered reasons for rejecting the original exclusion principle that are *independent* of the desire to solve the Exclusion problem. Part of the reason the original exclusion principle should be rejected is that properties at different levels weakly overdetermine their common effects. The argument for this claim was entirely independent on any reason one might have for wanting to solve the Exclusion problem. Rather, it was based on the possibility of there being no fundamental bottom level and the mundane empirical fact that composite objects have a mass that is the additive sum of the mass of their parts at a given level. So unless the reductionist can show this argument to be unsound or come up with similar independent reasons for amending Distinctness, the proposal to amend the exclusion principle is more attractive.

4.21 Limits of the Subset View

As we have seen, on Shoemaker's account, it is a necessary condition for realization that a realized property and its realizer stand in the subset relation. Since the subset relation is a relation between causal profiles or sets of causal powers, an obvious way in which the view is limited is with respect to epiphenomenal or non-causal

properties, i.e. properties that have no causal powers. Examples of such plausibly include logical or mathematical properties, for instance the property *being a prime number*. Obviously, if such properties have no causal profiles, then they cannot enter into the subset relation and by implication cannot enter into the realization relation. Perhaps there is a way of making sense of realization between logical or mathematical properties but if so, it would have to be spelled out in a different manner than the one proposed by Shoemaker.

Another potential limitation comes in the form of properties that are only partly individuated in terms of their causal powers. In other words, they are properties that have a non-causal aspect in addition to their causal powers. Phenomenal properties may be such properties. Many would argue that they are properties that have a certain phenomenal or qualitative aspect that cannot be accounted for in purely causal terms.[140] Note that one can accept this while still insisting that phenomenal properties have causal profiles. For instance, the phenomenal property *being in pain* may have a non-causal qualitative aspect (the "what's-it-likeness" of experiencing pain), in addition to being the cause of pain-behaviour. If phenomenal properties have causal profiles, then they are not only able to enter into the subset relation with physical properties; we know that they must do so, given the truth of the closure principle and the exclusion principle. However, as we have seen, a further condition for realization is that the realized properties are grounded in their realizers. And there are well-known reasons for thinking that phenomenal properties fail to supervene on physical properties and, by implication, fail to be grounded in physical properties. For instance, according to the so-called zombie argument, there are possible worlds that are duplicates of the actual world with respect to physical properties but in which our human counterparts are "zombies" with no phenomenal experiences. This is not ruled out by the mere fact that any phenomenal property has a causal profiles that is a proper subset of the profile of some physical property. At best, this ensures is that the zombie-world contains "zombie-properties", i.e. properties that have the same causal profile as phenomenal properties have in our world, but which bestow no phenomenality on the creatures that instantiate them.[141]

[140] See for instance Block [1978], Kim [2005], chap. 6 and Nagel [1974].
[141] Of course, if it is denied that phenomenal properties have causal profiles at all, then they are in the same boat as logical and mathematical properties and will not even be able to enter into the

It is, perhaps, too strong to say that phenomenal properties are properties to which to the Subset view does not apply at all. After all, the Subset view does apply to the aspect of phenomenal properties that can be captured in causal terms. However, the Subset view cannot *fully* account for them.

Phenomenal properties then, in so far as they are partly or fully individuated by their phenomenal character, will plausibly fail to be realized by physical properties since they plausibly fail to be grounded in physical properties, and this is revealed by a failure to supervene on physical properties. Still, just because there are properties to which the Subset view either does not apply (because these properties have no causal profiles) or to which it applies but for which physicalism nevertheless fails (because these properties have a non-causal, in addition to a causal aspect), does not mean that the theory should be rejected. Despite the potential limitations, the position still provides an attractive account of how irreducible mental properties can have causal efficacy, given the closure principle and the exclusion principle.

subset relation. Personally I'm not sure what the motivation would be for such a claim once the Subset view is accepted.

Chapter 5

The Subset View and Overdetermination of Causal Powers

Abstract: In this chapter I consider a new version of the Exclusion Problem, and whether the Subset view should be rejected on the basis of it. The starting point is an argument developed by Carl Gillett and Bradley Rives against the existence of determinable properties as they are construed according to the Subset view. Gillett and Rives contend that if determinable properties exist, they would overdetermine the causal powers of objects and that this would be an ontologically unparsimonious position. This argument can be extended to realized properties in general, including mental properties. I ultimately argue that worries about parsimony with respect to overdetermination of causal powers should be given little weight when adjudicating the truth of the Subset view. The reason is that such overdetermination plausibly occurs in other areas. Hence, we should be suspicious about very general metaphysical principles that prohibit it.

5.1 Introduction

In chapter 3.7, I considered some of the reasons for endorsing the original formulation of the exclusion principle. One of those reasons was that a position which entails the denial of the exclusion principle would be unparsimonious. Principles of parsimony, however, apply only to entities that it is unnecessary to posit. And, as I argued, it is necessary to allow for weak overdetermination in the case of some higher-level properties and the lower-level properties that compose them. This lead to an amended version of the exclusion principle according to which only strong systematic overdetermination is prohibited, and which therefore allows for weak overdetermination of behavioural effects by mental properties and their physical realizers. In this chapter I will discuss an objection to the effect that the move from allowing weak overdetermination by properties at *different* levels to allowing weak overdetermination by properties at the *same* level is illegitimate, and that a position that allows for the latter should be rejected on the grounds that it is unparsimonious.

This objection is based on an argument by Carl Gillett and Bradley Rives against the existence of determinable properties. However, their argument applies to all pairs of properties that stand in the subset relation. On the Subset view, this includes mental properties and their physical realizers. I will begin by giving a brief sketch of the distinction between determinable and determinate properties. Then I will present Gillett and Rives' argument and show how it plausibly escapes the conclusions from chapter 3. Finally I will consider an alternative response on behalf of the Subset view.

5.2 Determinable and Determinate Properties

Paradigm examples of determinable properties include properties like *colour*, *mass*, and *height*. Corresponding determinates to these determinables include properties like *red*, *a mass of 2 kilograms*, and *a height of 2 meters*. While much has been written about the various features of the relation between determinables and determinates, the crucial feature is arguably that determinates *specify* their corresponding determinables.[142] *Being red* is a specific way of *being coloured*, having *a mass of 2 kilograms* is a specific way of having *mass*, and *being 2 meters tall* is a specific way of having *height*. Determinate properties can themselves be determinables relative to other properties. For instance, *red* is both a determinable and a determinate property – it is a determinable relative to the property *scarlet* and it is a determinate relative to the property *colour*. It seems plausible, however, that not all properties are determinable. The property *a height of exactly 2 meters* is presumably not a determinable property relative to any determinate, since there is no way of having *height* that is more specific than having *a height of exactly 2 meters*. In what follows, when I talk of determinate properties it is implied that they are maximally determinate properties of this sort, i.e. determinate properties that allow for no further specification.

According to Shoemaker, the Subset view can be applied to determinable properties and their corresponding determinates. That is, any instance of a determinable property is realized by some instance of one of its corresponding determinates. For example, the redness of an apple is realized by some more determinate shade of red,

[142] For a comprehensive discussion of other features of the determinable-determinate relation see Funkhouser [2006].

say scarlet. This means that determinable and determinate properties stand in the subset relation – the token-powers of any instance of a determinable property form a proper subset of the token-powers had by some instance of one of its corresponding determinates. Of course, this makes effects of determinable causes weakly overdetermined since any one of these effects is the manifestation of a power that is bestowed by at least two distinct properties; a determinate property and its corresponding determinable. However, as long as the exclusion principle prohibits only systematic strong overdetermination, i.e. token-effects that are the manifestation of two distinct token-powers, this is unproblematic.

5.3 Gillett and Rives' Objection

The objection raised by Gillett and Rives is that such a view is ontologically unparsimonious. Following Gillett and Rives' terminology, *optimists* about determinable properties (that is, philosophers happy to posit their existence) face a serious metaphysical problem as their position involves overdetermination of the causal powers of the objects in the world. Pessimism, the view that determinable properties do not exist, on the other hand, has no such result. Coupled with a metaphysical principle about ontological parsimony, this, Gillett and Rives contend, shows that pessimism is to be preferred over optimism:

> The worry is that while we should clearly accept that *being one gram* and *being a charge of +1,6 x 10 ^(-19) coulombs* are properties which contribute powers to individuals, it is not clear that we should also take determinable properties such as *being a mass* and *being charged* to contribute powers *in addition* to the determinates that always accompany them. For to do so would be a kind of "double counting", to use David Lewis' phrase, of the causally efficacious properties. We shall call this the 'Parsimony Worry'.[143]

Obviously the Subset view of determinable properties does not fare well in the light of this worry, as it is at the heart of this account that the powers had by a determinable property such as *red* are also had by its corresponding determinate *scarlet*. Thus, as Gillett and Rives point out:

[143] Gillett and Rives [2005], 486-487. Note that Gillett and Rives have a different definition of what it is for a property to be causally efficacious from the one I use. To them, a causally efficacious property is one that actually bestows some power on an object (Gillett and Rives [2005], p. 486). Hence, on Gillett and Rives' account and contrary to my definition, causal efficacy need not involve the manifestation of any power. This is purely a terminological dispute.

> […] the Subset View incorporates a massive 'double-counting' of the properties that contribute powers and an associated overdetermination of causal powers. And both of these features heavily offend against any principle of ontological parsimony.[144]

As a consequence, Gillett and Rives propose that we give up the Subset View, and realism about determinable properties in general, in favour of eliminativism about determinable properties – if determinable properties do not exist, they do not overdetermine the powers of objects.

Why is it a problem for a theory that it incorporates 'double-counting' of properties and overdetermination of powers? Presumably because there is an implicit premise in Gillett and Rives' argument to the effect that the world is such that properties do *not* overdetermine the powers of objects. Or at least, that properties do not *systematically* or *in general* overdetermine the powers of objects, in the way they would do if instances of determinable properties have powers also had by some instances of their corresponding determinates. That is, the Parsimony Worry rests on a metaphysical principle which tells us that, in general, distinct properties do not bestow the same causal powers. We can call this *the principle of No Overdetermination of Powers (NOP)*.

This, of course, is the original exclusion principle with a twist. While the original exclusion principle focuses on causal overdetermination of effects, Gillett and Rives' principle, focuses on the non-causal overdetermination of powers. According to the Subset View, a red object has certain causal powers in virtue of having the property *red*. However, it also has those very same causal powers in virtue of having some determinate shade of red, say the property *scarlet*. These causal powers are thus grounded twice, and are metaphysically overdetermined. And this is what NOP prohibits.

Now, the conclusion from chapter 3, which refutes the original formulation of the exclusion principle, refutes NOP as well. Some higher-level composite properties bestow token powers that are identical to token-powers bestowed by the sum of

[144] Gillett and Rives [2005], 491.

lower-level properties that compose them, so NOP is false. Nevertheless, we can restrict NOP on Gillett and Rives' behalf and apply it only to pairs of properties of the same object, i.e. properties at the same level. Hence, the appropriate formulation of the principle of no overdetermination of powers is the following:

(NOP*): In general, distinct properties do not bestow the same powers on the same objects or individuals.

Since a determinable property and its determinate realizer are properties of the same objects, and so on the same level, this weaker formulation of the NOP-principle is all Gillett and Rives' need to run their argument against the existence of determinable properties - the fact that there is overdetermination of powers by properties at *different* levels does not refute NOP*.

Does this present an *ad hoc* restriction on the initial formulation of NOP? I think not. In so far as there is a substantial metaphysical distinction between levels and orders, it seems reasonable to think that what holds for properties at different levels may not hold for properties at different orders within a level.[145] For instance, worries about causal drainage are exclusively worries about the causal powers of properties at different levels, and does not apply to properties within a level. If so, it would not be obviously *ad hoc* to insist that a refutation of NOP based on examples involving properties at different levels, fails to extend to a refutation of NOP*. In other words, NOP* can be a plausible metaphysical principle even if it is conceded that NOP is not.

5.4 A New Exclusion Problem

As I mentioned above, Gillett and Rives endorse eliminativism about determinable properties in the light of NOP*. Now, that may not be such an unpalatable conclusion. Perhaps we can live with there being no determinable properties. The problem I will consider here is that Gillett and Rives' argument can be easily extended to target mental properties as well, as they are construed by the proponent

[145]By orders here I just mean any kind of division of same-level properties that stand in some interesting metaphysical relation, and not just the division between first- and second-order properties.

of the Subset view.[146] They too have causal powers that are identical to powers had by their physical realizers, and so, have powers that are overdetermined. If one shares Gillett and Rives' 'Parsimony Worry' about determinable properties, it seems that one should be equally worried about mental properties being unparsimonious. I do not mean to suggest that Gillett and Rives' endorse, or need endorse, eliminativism about mental properties. However, if NOP* is accepted, a new version of the Exclusion problem can be generated, with the premises formulated in terms of causal powers, rather than causes:

(Mental Causation):	Mental properties bestow powers to cause physical effects on some individuals.
(Closure):	Any power to cause a physical effect is also bestowed by a physical property.
(Distinctness):	Mental properties are distinct from physical properties.
(NOP*):	In general, distinct properties do not bestow the same powers on the same objects or individuals.

Of course, on the Subset view, the fact that the causal powers of any mental property *M* are also powers of some physical property *P* is part of what makes it the case that *P* is a realizer of *M*. And the Subset view also endorses Distinctness. This, however, is in tension with NOP*. So the proponent of the Subset view cannot accept all four claims.

*5.5 Rejecting NOP**

I will now go on to argue that NOP* should be rejected. The strategy for doing so will be to show that there are cases where overdetermination cannot plausibly be avoided when accounting for the powers of objects, and that these cases are not highly theoretic or exceptional but rather commonplace and systematic. These cases involve the grounding of what I call *joint powers*. Since overdetermination of powers not only occurs, but is commonplace, this means NOP* must be false. This,

[146] If mental properties just *are* determinable properties with physical properties as their determinates, Gillett and Rives' argument need not be extended but would apply directly to mental properties. Some philosophers have explored this possibility, in particular Yablo in his [1992]. For some dissenting views, see Ehring [1996] and Menzies [2008]. I will say more about the problems associated with treating the relation between mental properties and their realizers as an instance of the determination relation in the next chapter.

in turn, makes it much less clear why anyone endorsing the Subset View, should be worried about the overdetermination of powers that follows from her position.

5.6 Varieties of Powers - Simple, Compound, and Joint

I have been assuming throughout that properties ground the causal powers of an object (or event) by bestowing them on that object - objects have their causal powers *in virtue of* instantiating certain properties. For instance, say that Alice the pigeon has been trained to peck at objects if they are scarlet. In this case, the power had by scarlet objects to cause Alice to peck is grounded in the property *scarlet*, which bestows on those objects the power to cause Alice to peck. Call a power that is bestowed by a single property in this way a *simple power*.

Powers are not always simple. In many cases a plurality of properties work together to bestow a single power.[147] How might such powers be grounded? To answer this question, we need to consider how causal powers are individuated. The most obvious answer is that they are individuated in terms of the effects they are powers to cause. This can be summed up in the following slogan: If two objects have powers to cause the same effects in the circumstances, then they have exactly the same powers.

With this in mind, we can make sense of powers being composed of other powers:

(Power Parthood): A power P to cause effect E has power P_1 as a proper part iff P_1 is a power to cause effect E_1, where E_1 is a proper part of E.

And this allows for a straightforward way in which a multitude of properties together can bestow a power:

(Multi bestowing): Properties X_1, X_2, …, Xn together bestow power P if power P is entirely composed of powers P_1, P_2,…, P_n and X_1 bestows P_1, X_2 bestows P_2, …, and Xn bestows P_n.

[147] One example of this is when a plurality of lower-level properties bestow a plurality of power that are identical to a power also bestowed by the higher-level property they compose.

Here is an example to illustrate how this works. Take Vera, a pigeon which has been trained to peck when presented with scarlet objects and lift her wings when presented with round objects. Scarlet round objects thus have the power to cause the total effect of Vera pecking and lifting her wings. This power (the power to cause Vera to peck and lift her wings) is not had or bestowed by any single property and is therefore not a simple power. Rather, it is bestowed by two properties, *scarlet* and *round*, in the way described by (Multi bestowing). The power to cause Vera to peck and lift her wings is entirely composed of two distinct simple powers, the simple power to cause Vera to peck and the simple power to cause Vera to lift her wings. The former is bestowed by *scarlet*, the latter by *round*. And that is how *scarlet* and *round* together ground the power to cause Vera to peck and lift her wings. Call a power that is composed of simple powers in this way a *compound power*. Compound powers thus have the following two characteristics: One, they require the instantiation of a plurality of properties to be bestowed, and two, each property involved in bestowing a compound power bestows a power to cause just a proper part of a total effect (i.e. no single property involved bestows the power to cause the total effect).

I take the example with Vera to show that cases of compound powers are possible. And while that particular example is somewhat artificial, instances of compound powers are commonplace in the actual world as well. A cup of coffee has the power to reflect light at a certain wavelength in virtue of *being black* and the power to burn the tongue in virtue of *being hot* which compounds to the power to reflect light at a certain wavelength *and* burn the tongue, a football has the power to travel a certain distance in virtue of having *velocity* and the power to attract other objects in virtue of having *mass*, which compounds to the power to travel a certain distance *and* attract other objects, and so on.[148] Note that compound powers need *not* pose any problem for NOP* as long as the powers bestowed by each property involved in bestowing a compound power are distinct. And since powers are individuated by

[148] You might think that while such cases are actual, they are as artificial as the case of Vera the pigeon. After all, who would ever think to count the effect of reflecting light at a certain wavelength *and* burning the tongue as *one* effect and the power to cause such an effect as *one* power? I fully concede this point. Compound powers are easy to come by as they are simply mereological sums of distinct powers.

the effects they are powers to cause, powers are distinct if they are powers to cause distinct effects. In other words, if properties X_1 and X_2 bestow powers P_1 and P_2 respectively, which entirely compose compound power P, they do so in a manner that is kosher from the perspective of NOP*, if P_1 and P_2 are each powers to cause *distinct* effects E_1 and E_2. And in all the cases of compound powers mentioned above, the proper parts of the relevant total effects are distinct. Vera pecking is different from Vera lifting her wings, a cup of coffee reflecting certain wavelengths is different from tongue-burning, and a football travelling through the air is different from a football attracting other objects. In neither of the cases mentioned do two distinct properties bestow a power to cause the *same* effect. Rather, in every case, distinct properties each bestow a power to cause a distinct proper part of a total effect - a total effect that neither property on its own bestows a power to cause.

Sometimes, however, properties come together to bestow a power in a very different manner. Imagine Agnes, a pigeon trained to peck at objects only if they are both scarlet and round.[149] Intuitively, the properties *scarlet* and *round* together bestow the power to cause Agnes to peck, just like *scarlet* and *round* together bestow the power to cause Vera to peck and lift her wings. Unlike the latter case, however, the effect of Agnes pecking is not composed of distinct parts which *scarlet* and *round* each bestow a power to cause. For instance, *ex hypothesi*, it is simply not the case that *scarlet* bestows the power to cause Agnes to peck just halfway. Instead, the power to cause Agnes to peck is an all or nothing affair. Either *scarlet* and *round* are both instantiated by an object, in which case the object has the power to cause Agnes to peck, or either *scarlet* or *round* or both are not instantiated by the object, in which case it has neither the power to cause Agnes to peck nor the power to cause any proper part of that effect. I will call a power that is bestowed by a multitude of distinct properties in this way a joint power. Joint powers thus have the following two characteristics: One, they require the instantiation of a plurality of properties to be bestowed (this characteristic is shared with compound powers), and two, none of the properties in this plurality bestow just a power to cause a proper part of the relevant effect (*scarlet*, for instance, does not bestow the power to cause Agnes to peck just halfway).

[149] This example is due to McLaughlin. See McLaughlin [2007].

Before moving on to the question of how joint powers of objects are grounded, note first that, like instances of compound powers, instances of joint powers are commonplace. I (and many others) drink coffee only if it is black and hot, so a cup of coffee has the power to cause me to drink it in virtue of having the properties *black* and *hot*. Yet neither property on its own bestows the power to cause any part of this effect (if the coffee was white and hot or black and cold, it would not have the power to cause any part of me drinking it). A football has the power to break a window in virtue of having both *mass* and *velocity*, yet neither property on its own bestows a power to cause any proper part of this effect (if the football has a mass but no velocity or a velocity but no mass, it would not have the power to cause any proper part of the breaking), and so on.

How are joint powers of objects grounded? For instance, how do the properties *scarlet* and *round* ground the power of scarlet round objects to cause Agnes to peck? One response would be to insist that it is a brute fact that they do so – *scarlet* and *round* together bestow the power to cause Agnes to peck, and there is no deeper metaphysical fact that explains this. However, this seems to me an unhappy conclusion, in so far as we expect that the nature and causal powers of a plurality of properties depends on the nature and causal powers of the individual properties that make out this plurality. This, for instance, is the case for compound powers, which are bestowed by a plurality of properties in virtue of each of these properties bestowing a proper part of this power. In other words, while compound powers are bestowed by a plurality of properties, they are ultimately grounded in the causal powers of singular properties. Since this model fails to apply in the case of joint powers, the question is whether an alternative account is available. I can think of two ways joint powers might be grounded in a way that does not rely on a plurality of properties 'brutely' bestowing them. As I will now argue, both provide counterexamples to NOP*.

5.7 Grounding Joint Powers in Conjunctive Properties

First, it might be thought that the property which grounds the power of scarlet round objects to cause Agnes to peck is the conjunctive property *scarlet and round* (and similarly, what grounds the power of footballs to cause windows to break is the

conjunctive property *mass and velocity*, what grounds the power of cups of coffee to cause me to drink them is the conjunctive property *hot and black*, and so on for other cases of joint powers). This account side-steps the issue over how a plurality of properties can bestow a joint power, since it claims that any joint power is bestowed by a singular conjunctive property. Still, the proposal arguably manages to incorporate the two main characteristics of joint powers. First, if conjunctive properties are grounded in their conjunct properties (a view, I take it, that many of us would be inclined to hold), joint powers *do* involve a plurality of properties, in this case the properties which form the conjunctive property, which in turn bestow and hence ground the joint power. Second, as it is the conjunctive property *scarlet and round* which alone bestows the power to cause Agnes to peck, neither of the properties involved in bestowing a joint power bestows just a power to cause a part of the relevant effect (the conjunctive property *scarlet and round* bestows a power to cause the total effect of Agnes pecking and not only a part of the pecking, and the conjunct properties *scarlet* and *round* do not bestow a power to cause any part of this effect at all).

But while the introduction of conjunctive properties allows for the grounding of joint powers, it also seems to violate NOP*. Again, assume that scarlet objects have the power to cause Alice to peck and that *scarlet* bestows this power. Some of these objects are also round and so, have the conjunctive property *scarlet and round*. Now, say that a scarlet round object is placed in front of Alice and causes her to peck. I take it as undeniable that the salient or relevant cause of this effect is the property *scarlet* (at least in normal contexts). However, it seems awfully rough to insist that the conjunctive property *scarlet and round* is *completely* inefficacious with respect to this effect. The reason, I think, is that there is an intuitive sense in which conjunctive properties, assuming here that they exist, include their conjunct properties. It is the same reason why it seems awfully rough to insist that *scarlet* is completely inefficacious with respect to effects caused by *red*. In some sense, to be scarlet *is* to be red, just like being scarlet and round *is* to be scarlet (among other things – it is also to be round of course).[150] Hence, to *completely* exclude

[150]This is not to say that the sense in which to be scarlet is to be red is the same sense in which to be scarlet and round is to be scarlet. As we saw, one of the platitudes about the determinable-determinate distinction is that determinates specify their corresponding determinables. So to be

determinate/conjunctive properties from being causally efficacious with respect to effects caused by their determinables/conjunct properties just doesn't sound quite right.[151] If conjunctive properties stand in the subset relation with their conjunct properties, however, we can accommodate these intuitions. The conjunctive property *scarlet and round* does cause Alice's pecking since its causal profile includes the causal profile of *scarlet*. However, it is not a salient cause of the pecking since the profile also includes relatively many causally irrelevant powers, for instance, the powers that form the causal profile for the property *round*. Again, this is analogous to my comments on the case of Sophie pecking at all red objects – *red* is the salient or explanatorily relevant property, however, *scarlet* is still a cause of this effect as well. But of course, to say that conjunctive properties and their conjunct properties stand in the subset relation is to accept that they overdetermine some causal powers – the powers bestowed by the conjunct properties are also bestowed by the conjunctive property that it is a conjunct of. So NOP* must be false.[152]

To sum up then, if you introduce conjunctive properties in order to ground joint powers of objects, the powers bestowed by any of the conjunct properties will be bestowed twice on those objects, once by that conjunct property and once by the conjunctive property which grounds the joint power. And since joint powers are commonplace this means overdetermination of powers is commonplace. NOP* must be false.

5.8 A Potential Pessimist Rejoinder

What would it take for the Gillett and Rives to avoid this conclusion? Given the assumption that the conjunctive property *scarlet and round* grounds the power to

---

scarlet is to be red in a *specific* way. But to be scarlet and round is not to be scarlet in a specific way. Rather, to be scarlet and round is to be *partly* scarlet.

[151] Of course, Gillett and Rives do not think determinable properties exist so obviously they do not think they can be causes. They cannot say the same for conjunctive properties, however, since we are assuming here that they are the properties that ground joint powers.

[152] Another reason to think that conjunctive properties and their conjunct properties stand in the subset relation is that this explains why conjunct properties supervene on the conjunctive properties they are conjuncts of. Why is it that an object with the property scarlet and round necessarily has the property scarlet? Again, intuitively the relation of metaphysical dependence is reversed – conjunctive properties depend on their conjunct properties. So the supervenience cannot be explained by reference to the grounding relation. If, however, these properties stand in the subset relation this is no mystery since the causal profile of scarlet is included in the causal profile for scarlet and round.

cause Agnes to peck (which is not bestowed by any of the individual conjunct properties), rejecting the existence of the conjunctive property *scarlet and round* is not an option, as to do so would leave at least one power ungrounded.[153] If one is determined to get rid of the overdetermination of the powers of scarlet round objects it therefore seems that the only option is to insist that scarlet round objects have neither the property *scarlet* nor the property *round.* As silly as this sounds, this would actually be in line with Gillett and Rives' advice regarding determinable properties, i.e. keep the property which bestows the powers in the superset (determinate properties/conjunctive properties) and eliminate the properties bestowing the powers in any subset of this superset (determinable properties/conjunct properties). And granted, a consequence of such a move would be that the power of scarlet round objects to cause Agnes to peck is no longer overdetermined. However, it is highly counter-intuitive to insist that scarlet round objects have neither the property *scarlet* nor the property *round* but only the conjunctive property *scarlet and round,* that a football flying through the air have neither *mass* nor *velocity* but only *mass and velocity,* or that the coffee I'm drinking has neither the property *hot* nor the property *black* but only the property *hot and black,* and so on. For this reason it seems that the proponent of NOP* would do better to look elsewhere to ground joint powers.

5.9 Grounding Joint Powers in Properties Bestowing Conditional Powers

Alternatively, an account of how joint powers are grounded without the introduction of conjunctive properties may be found in Shoemaker's account of conditional powers. A conditional power is a power bestowed by a property on an object only if certain other properties are also had by that object. To use one of Shoemaker's favourite examples, *being knife-shaped* bestows on an object a power to cut wood, only if that object also has the property of being suitably rigid. If the object is made of paper or whipped cream, for instance, it won't have the power to cut wood, even if knife-shaped. Similarly, *scarlet* may be said to bestow the power to cause Agnes to peck on objects conditionally, in particular conditioned on those

[153] I am assuming that this conjunctive property is maximally determinate. Otherwise Gillett and Rives' would reject its existence on the grounds that it is a determinable property. Of course, they would have to replace with the conjunctive property that *is* a maximally determinate of the property *scarlet and round*, in which case the problem reappears.

objects also having the property *round* (and similarly, *mass* bestows on footballs the power to cause windows to break conditioned on those footballs also having *velocity*, *being black* bestows on cups of coffee the power to cause me drink them conditioned on those cups also *being hot*, and so on for other cases of joint powers). Like the account considered in the previous section that grounded joint powers in conjunctive properties, this account succeeds in incorporating the two main characteristics of joint powers. First, joint powers will involve a plurality of properties, the property conditionally bestowing the joint power and the property that this bestowal is conditioned on. And second, neither of the properties involved bestows just a power to cause a part of this effect (*scarlet* bestows the power to cause the total effect of Agnes pecking and *round* does not bestow the power to cause any part of this effect at all but functions instead as a sort of background condition which enables *scarlet* bestow the joint power). The problem from the perspective of the proponent of NOP* is that it seems just as correct to say that *round* bestows on some objects the power to cause Agnes to peck, conditioned on those objects being *scarlet*. After all, there is nothing about the case that suggests *scarlet*, as opposed to *round*, is better suited as a full-blown causally sufficient property, rather than a mere background condition.[154] Instead, the case at hand is entirely symmetrical - *scarlet* and *round* are each equally good candidates for conditionally bestowing the power to cause Agnes to peck. The natural conclusion to draw from this is that *scarlet* bestows conditionally on some objects the power to cause Agnes to peck (conditioned on those objects having the property *round*), and that *round* bestows conditionally on those very same objects the power to cause Agnes to peck (conditioned on those objects having the property *scarlet*). But if *scarlet* and *round* both conditionally bestow the power to cause Agnes to peck on the same objects, and if this example is representative of joint powers in general, then we have a counterexample to NOP*. We have distinct properties bestowing the very same power on the same objects.

[154] This is assuming that there is even a genuine metaphysical distinction to be made between properties functioning as background conditions and causally sufficient properties which may be disputed.

5.10 Further Potential Pessimist Rejoinders

How can the pessimist avoid this conclusion? One way would be to simply deny that *scarlet* and *round* both bestow the power to cause Agnes to peck. As a matter of fact, the proponent of NOP* could argue, only one of them bestows the power to cause Agnes to peck while the other is merely a background condition. This, however, is terribly *ad hoc* and smacks of metaphysical arbitrariness. What, we might ask, determines which property gets to bestow the power to cause Agnes to peck? As noted above, the case appears to be entirely symmetrical and there seems to be no relevant factors that could decide this question. At best it would be arbitrary which property would win out. Of course, this worry has a distinctive epistemic flavour. After all, it may turn out there are factors of which we are simply unaware which *would* decide the question. Nevertheless, it is doubtful that the pessimist would want his argument to hinge on the mere possibility that, despite appearances and as a matter of fact, either *scarlet* or *round* simply does not bestow the power to cause Agnes to peck. Until such factors should reveal themselves to us, it is reasonable to assume that both *scarlet* and *round* bestow this power, and hence that NOP* is false.

Another, initially more promising, strategy that might be pursued would be to deny that powers are *entirely* individuated by the effects they are powers to cause. In particular, the proponent of NOP* may insist that conditional powers are partly individuated by the conditions that must obtain for them to be bestowed. I take it that this is how many of us tend to think of causal powers. For instance, a boulder and a paper clip may both be able to cause a plane of glass to break when placed on it, yet we might balk at the claim that this is the very same type of power, since the boulder will manifest this power in a wide range of both normal and abnormal conditions, whereas the paper clip will only manifest this power when the plane of glass is already under considerable stress. This indicates that powers are partly individuated in terms of the conditions in which they are bestowed, in which case the boulder and the paper clip have different types of powers to cause the very same effect. We can see how this also leads to *scarlet* and *round* bestowing distinct powers, since the conditions that must obtain for those powers to be manifest are different. *Scarlet* bestows the power to cause Agnes to peck in conditions where *round* is instantiated, and *round* bestows the power to cause Agnes to peck in

conditions where *scarlet* is instantiated. These are simply two distinct types of powers bestowed by two distinct properties. And if *scarlet* and *round* bestow *different* conditional powers, grounding joint powers in properties bestowing conditional powers no longer gives us a counterexample to NOP*, since NOP* only tells us that distinct properties do not bestow the *same* powers to the same object. A new problem arises, however. Rather than having more properties than we need to account for the *power* to cause Agnes to peck, we now have more powers than we need to account for the *effect* of Agnes pecking. The manifestation of the power bestowed by *scarlet* (the power to cause Agnes to peck conditioned on *round*) is sufficient for Agnes pecking. And likewise for the power bestowed by *round* (the power to cause Agnes to peck conditioned on *scarlet*). Therefore (again given the symmetry of the case), if Agnes is put in front of a scarlet round object and pecks, we must reasonably conclude that both powers were manifested. So when Agnes pecks, we have two distinct properties that are each causally sufficient *and* efficacious for this pecking. But of course, this is just a case of strong causal overdetermination. In other words, if the properties bestowing a joint power are each bestowing distinct powers, joint *powers* would indeed not be overdetermined. However, the manifestation of joint powers would involve systematic and strong *causal* overdetermination. But this cannot be a better result than the overdetermination of powers that follows if *scarlet* and *round* bestow the *same* power conditionally. Getting rid of overdetermination of powers on pain of accepting overdetermination of effects is hardly an attractive trade-off. A possible way out of this would be to insist that, as a matter of fact, only *scarlet* or *round* gets to be causally efficacious for the effect of Agnes pecking. But again, given the symmetry of the case it is difficult to see why this is not simply an *ad hoc* solution which faces the same problem as insisting that, as a matter for fact, either *scarlet* or *round* does not bestow the power to cause Agnes to peck.[155]

[155] Philosophers sympathetic both to the notion of conditional powers and the idea that powers are partly individuated in terms of the conditions in which they are bestowed, may find it worrying to discover that this leads to strong causal overdetermination whenever joint powers are manifested. Maybe they even find it so worrying that they start to suspect that something must be wrong about an argument that has this conclusion. There's a way to avoid the conclusion though. In the case of Agnes pecking, *scarlet* and *round* bestow different *types* of conditional powers on scarlet round objects (assuming that powers are partly individuated in terms of the conditions in which they are bestowed). However, any token of one of these powers may be identical to a token of the other. So for any particular scarlet round object, it has just one token-power to cause Agnes to peck, though

5.11 Concluding Remarks on the Grounding of Joint Powers

The upshot of this is that joint powers cannot be grounded without NOP* being false. If joint powers are grounded in conjunctive properties, these conjunctive properties will overdetermine the powers bestowed by their conjunct properties. And if joint powers are grounded in properties bestowing conditional powers, the joint powers themselves will be overdetermined by the property bestowing the conditional power and the property this bestowal is conditioned upon. Either way, NOP* is false and an argument resting on NOP* will be unsound.

*5.12 Weakening NOP**

Seeing that NOP* turns out false on both accounts of how joint powers might be grounded, the pessimist about determinable properties may consider whether she can establish her position with less. Perhaps she does not need to commit to the claim that, in general, distinct properties bestow the same powers to the same object. All she needs to establish pessimism is the weaker claim that *realized properties* do not overdetermine the powers also bestowed by their realizers. If this can be established, Gillett and Rives can still run their argument against the existence of determinable properties (as they are construed according to the Subset view), and the Exclusion problem formulated in terms of causal powers would still present a paradox. The trick, of course, is to show that a further weakening of the original NOP principle is justified and not merely *ad hoc*. For instance, to explicitly restrict NOP to realized properties and their realizer and claim that they cannot overdetermine any causal powers would be to beg the question against the proponent of the Subset view. There has to be a substantive and relevant difference between the overdetermination of powers involved in grounding joint powers and the overdetermination of powers involved in positing the existence of realized properties, if the pessimist wishes to allow only for the former.

Here is a potential difference that someone sympathetic to the Parsimony Worry may exploit: If joint powers are grounded in properties bestowing conditional powers, overdetermination of powers cannot be avoided. This is because the

---

this is a token-power of two different types of powers. Of course, this does not help the proponent of NOP* since this token-power will be determined by two distinct properties, *scarlet* and *round*.

overdetermining properties each bestow powers that the other does not. For instance, the joint power to cause Agnes to peck is bestowed both by *scarlet* and *round*. However, since *scarlet* bestows powers that *round* does not, and *vice versa*, both properties must be posited to account for *all* the powers of scarlet round objects. To deny the existence of either *scarlet* or *round* to avoid the overdetermination of their joint powers would leave some powers ungrounded. Contrast this with the grounding of the powers putatively bestowed by realized properties. Since these are a subset of the powers that are bestowed by the realizer, we don't *need* to ground them by positing realized properties. There would be no powers ungrounded even if the existence of realized properties was denied.[156]

The weakened version of NOP* that follows from this line of reasoning would then be something like this:

(NOP**): In general, distinct properties do not bestow the same powers on the same objects or individuals, *except in cases where they do so as a result of grounding all the powers of those objects and individuals.*

This version of NOP would allow the pessimist to accept the overdetermination involved in grounding the joint powers of objects in properties bestowing conditional powers, while at the same time preventing realized properties from overdetermining the powers bestowed by their realizers. Nevertheless, I believe such a strategy brings us to a point where the pessimist is in a dialectically awkward position. To see this, it is important to note that the proponent of the Subset view will have their own reasons for being an optimist with respect to realized properties. This is especially clear in the case of mental properties and philosophy of mind, where eliminativism is generally frowned upon. Indeed, the reasons for thinking that mental properties have causal powers presuppose that these properties exist and that eliminativism about mental properties is false. Now, I do not wish to deny that

[156] If joint powers are grounded in conjunctive properties, this difference between the overdetermination involved in grounding joint powers and the overdetermination involved in positing realized properties would no longer be present. The reason is that, like realized properties, conjunct properties have only causal powers that form a subset of some other properties (conjunctive properties). And since these conjunctive properties must be posited to ground joint powers, we wouldn't need to posit the conjunct properties. As I mentioned earlier, this would be quite counter-intuitive.

principles of parsimony can play a role in constraining metaphysical theories or even deny that it can trump what looks like good reasons for positing entities of a certain kind. However, the proponent of the Subset View is well within her right to find it unfair that the pessimist about realized properties asks her to give up a view that, as I argued in the previous chapter, is motivated by a combination of empirical and philosophical evidence, purely on the basis of a general metaphysical principle that has not only been weakened to allow for overdetermination of powers by properties at different levels, but now has also been weakened to allow for *some* overdetermination of powers by properties at the same level. It may be reasonable to ask philosophers to adhere to a certain general methodological principle, it is much less reasonable to ask them to adhere to this principle, and then go on to list several areas where the principle may be overruled.

In other words, I do not think that proponent of the Subset View should feel embarrassed about endorsing the following formulation of the NOP principle:

(NOP***): In general, distinct properties do not bestow the same powers on the same object, *except in cases where these properties stand in the realization relation and we have good reasons for thinking that the realized properties exist.*

Of course, this leaves it open for someone to argue that a certain class of properties fails to be exempted. For instance, pessimists about determinable properties, such as Gillett and Rives, may wish to deny that we have good reasons for thinking that they exist.[157] Likewise, someone may think that mental properties exist but that they fail to be realized by physical properties, in the sense of realization endorsed by the Subset view. In both cases, the Subset view would be false. What the discussion in this chapter shows, however, is that such arguments must be largely independent of the Parsimony Worry. There is no straight-forward route from principles of parsimony to the denial of the Subset view.

[157] It's worth noting that Gillett and Rives' spend a large part of their paper following this strategy as they discuss, and reject, David Armstrong's argument that an adequate account of laws of nature requires determinable properties.

5.13 Concluding Remarks

I conclude that Gillett and Rives' pessimist argument against determinable properties fails. In its original form, it relies on a metaphysical principle that is false. And a modified version of the argument that tries to get by with a weaker principle places the pessimist in a dialectically awkward position as the proponent of the Subset view can provide a weakened version of NOP herself which is compatible with her position. This also means that the Exclusion problem framed in terms of causal powers fails to get off the ground. To reject the Subset view, principles of parsimony ultimately play little role. Rather, a rejection must involve arguments that directly target the reasons for endorsing the Subset view.

Chapter 6

Alternative Non-Reductive Solutions to the Exclusion Problem

Abstract: In this chapter I discuss two current alternative non-reductive solutions to the Exclusion problem. The first is an argument from causation according to which a difference-making account of causation makes false the closure principle. I defend this solution against the charge that this contradicts the empirical evidence for closure but nevertheless argue that the Subset view is superior since it can explain mind-body supervenience and can take onboard many of the insights of a difference-making account of causation. The second is an argument from the metaphysics of properties according to which adopting an ontology of tropes will allow us to reduce mental causes to physical causes, without contradicting the argument from multiple realization. I ultimately reject this solution on the grounds that it requires us to accept the implausible claim that mental and physical tropes stand in a determinable-determinate relation.

6.1 Introduction

In chapter 3.21-22, I argued that a consequence of the drainage problem is that we must allow for weak overdetermination of some effects by properties at different levels. In chapter 5, I extended this point to apply to certain same-level properties, namely properties having powers to cause joint effects. All in all, these considerations make an Exclusion principle that prohibits this kind of overdetermination highly implausible. Furthermore, the reasons for accepting the Exclusion principle in the first place do not apply to weak overdetermination. Hence, the original formulation of the Exclusion principle is ambiguous: It does not distinguish between strong and weak overdetermination and once disambiguated, it becomes clear that the evidence supports only a formulation in terms of strong overdetermination. In chapter 4, I argued that with this amended version of the Exclusion principle, the statements constituting the Exclusion problem actually entails the Subset view. And since each of these statements is individually motivated, this means that the Subset view is motivated.

This, of course, assumes that once the Exclusion principle is suitably amended, no further disambiguation or amendments of these statements is required. In this chapter, I will discuss two recent dissenting views. One view holds that the Closure principle (roughly the claim that every physical effect has a sufficient physical cause) must be amended; another that the notion of a property employed in MC (the claim that mental properties are causally efficacious) and Distinctness (the claim that mental and properties are distinct) is ambiguous. Both hold that this amendment/disambiguation is the key to solving the Exclusion problem without giving up the claim that mental and physical properties are distinct. This in itself is not a problem for the proponent of the Subset view. After all, it may well be that there are several distinct reasons for thinking that the Exclusion problem fails to present a genuine paradox. What is a problem is that, as I will argue shortly, the non-reductive views that fall out of these alternative solutions to the Exclusion problem are incompatible with the Subset view. Hence, the proponent of the Subset view must show why these alternative non-reductive views are wrong, or at least show that the Subset view is superior to them. It is the purpose of this chapter to do just that.

6.2 Amending the Closure Principle – Causal Sufficiency versus Causal Efficacy

Some philosophers have pointed out that there is an important difference between a property being causally sufficient for an effect E and that property being a cause of E. For instance, Yablo, in his discussion of determinable and determinate properties, insists that pairs of determinates and determinables compete only for the role of being a cause or causally efficacious, with respect to a certain effect, not for the role of being causally sufficient for this effect.[158] If so, causal sufficiency must be a weaker notion than causal efficacy. Yablo then goes on to argue that a determinate property may be causally sufficient for an effect that is actually caused by its corresponding determinable (recall, Yablo's view is that a causally efficacious property with respect to some effect is the property that is proportional (or most proportional) to that effect, and that, in some cases, determinable properties satisfy this condition better than any of their determinates). Yablo brings

[158] Yablo [1992], p. 274.

up this distinction in order to falsify the following version of the Exclusion principle:

(Exclusion*):

> If a property X is causally sufficient for an effect E, then no property distinct from X is causally relevant to E.[159]

This principle obviously cannot be right if a determinate property can be causally sufficient for an effect caused by its corresponding determinable.

In a couple of recent papers (one co-written with Christian List), Peter Menzies has argued that Yablo's points can be applied, even in cases where the properties in question do not stand in a determinable-determinate relation.[160] Like Yablo, Menzies' conclusion is that once the distinction between causal sufficiency and causal efficacy is appreciated, the Exclusion principle formulated in terms of causal sufficiency can be rejected, and like Yablo, Menzies arrives at this via a conception of causes as difference-makers. [161] According to Menzies and List, difference-making can be analysed counterfactually in the following way:

(Difference-making): *C* makes a difference to *E* (and so, is a cause of E according to the difference-making account of causation) if and only if:
1) If *C* is instantiated, then *E* is instantiated.
2) If *C* is not instantiated, then *E* is not instantiated.

This account differs from a standard counterfactual account of causation since Menzies and List offer a non-standard semantics of counterfactuals. On a standard Lewis-Stalnaker interpretation, 1) is true if *C* and *E* are both actually instantiated. On Menzies and List's account, however, it is in addition required that in all possible worlds *relevantly similar to the actual world* in which *C* is instantiated, *E*

[159] Yablo [1992], p. 247.
[160] Menzies [2008] and Menzies & List [unpublished].
[161] The idea that causes are difference-makers is common to a range of theories about causation, including Yablo's theory that causes should be proportional to their effects, traditional counterfactual accounts (Lewis [1973]), interventionist accounts (Woodward [2003]), and contrastive accounts (Schaffer [2005]). Yablo is explicit that the proportionality claim (the claim that causes are proportional to their effects) is an attempt to capture this idea.

is instantiated. Likewise, for 2) to be true, it must be the case that in all worlds relevantly similar to the actual world in which *C* is not instantiated, *E* is not instantiated either.[162] One advantage of this interpretation is that it allows Menzies and List to flesh out the notion of causal sufficiency.[163] For a property to be causally sufficient for effect *E*, it is enough that the first counterfactual is satisfied. That is, *C* is causally sufficient for *E* if and only if, in every possible world relevantly similar to the actual world in which *C* is instantiated, *E* is also instantiated. This explains why, in the Sophie case, *scarlet* is sufficient for Sophie pecking without being a cause of this effect. It is sufficient because in worlds relevantly similar to the actual world in which the object in front of Sophie is scarlet, she pecks at it. It is not, however, a cause of the pecking, since there are possible worlds relevantly similar to the actual world in which the object is not scarlet, but in which Sophie nevertheless pecks, for instance, worlds in which the object is crimson or pink.

To appreciate how this applies to mental properties and their physical realizers, consider a case where a mental property *M* is causing the instantiation of a certain behavioural property *E*, and where a physical property *P* realizes *M*. Assuming physicalism, *M* supervenes on *P*. This makes *P* causally sufficient for *E*, since in all possible worlds relevantly similar to the actual world, in which *P* is instantiated, *M* is instantiated and causes *E*. There are, however, relevantly similar worlds in which *P* is not instantiated but in which the behavioural effect *E* nevertheless occurs, for instance, worlds in which a different physical realizer of causally efficacious mental property *M* is instantiated. In other words, whenever a mental property is causing an effect, the physical property that realizes this mental property will be causally sufficient for this effect.

6.3 The Difference-making Account of Causation and the Exclusion Problem

Menzies view raises some questions, in particular about how to understand the notion of possible worlds that are *relevantly* similar to the actual world. Let me put

[162] Menzies and List [unpublished], p. 6.
[163] Something Yablo does not do, preferring, as he says, to remain neutral on how best to understand the notion (Yablo [1992], fn. 5).

those questions aside however, and instead focus on how Yablo and Menzies propose to solve the Exclusion problem. First, note that both Yablo and Menzies assume that the Exclusion principle operative in the Exclusion problem is Exclusion* - that is, a principle according to which the causal *sufficiency* of a property with respect to effect E, excludes all other properties from being causally *relevant* with respect to E. Causal relevance, whatever this exactly amounts to, includes being causally efficacious with respect to E. So according to Exclusion*, if a property is causally sufficient for an effect E, then no other property can be a cause of this effect.[164] However, the relevance of this Exclusion principle, as far as the Exclusion problem is concerned, is questionable. For one thing, despite the fact that Menzies and List explicitly attribute it to Kim, it is nowhere to be found in Kim's writing (that I have been able to find anyway). Instead, what Kim says is this:

> To acknowledge mental event *m* (occurring at *t*) as a cause of physical event *p* but deny that *p* has a physical cause at *t* would be a clear violation of the causal closure of the physical domain […]. But to acknowledge that *p* has also a physical cause, *p**, at *t* is to invite the question: Given that *p* has a physical cause *p**, what causal work is left for *m* to contribute?[165]

In other words, Kim's worry is not that mental property *M* (or mental event *m*, qua instantiating mental property *M*) is excluded from being causally efficacious with respect to physical effect E due to there being a physical property which is merely causally *sufficient* for E. The worry is rather that the causal efficacy of *M* is threatened due to there being a physical property which is a sufficient *cause* of E. Assuming that this worry is based on the prior acceptance of some version of the Exclusion principle, it cannot be the version that Yablo and Menzies aim to refute. Instead, it must be a version framed in terms of sufficient causes, which is also the version I used when originally setting up the Exclusion problem – in general, a property being a sufficient cause of E excludes other properties from being causes of E.

[164] Yablo [1992], p. 247, Menzies [2008], p. 196, and Menzies and List [unpublished], p. 1.
[165] Kim [1998], p. 37. While Kim frames the problem in terms of events, this can easily be translated into a problem about the causal efficacy of mental properties.

What do Yablo and Menzies have to say about this weaker version of the Exclusion principle? Yablo does not directly address the issue, though as I pointed out in chapter 3, the proportionality constraints on causation arguably *vindicate* it. The reason is that they generate unique causes among sets of causally competing determinate properties and their less specific determinables.[166] So if a property satisfies the proportionality constraints, every determinate and determinable of this property will fail to do so, and so, fail to be a cause of the effect in question. Menzies explicitly considers whether or not to accept an Exclusion principle framed in terms of sufficient causes, as opposed to causally sufficient properties. While he hesitates to adopt this principle outright (as he says, he is uncertain if the principle in general is true or false), he does go on to argue that, in the specific case of mental causation, the causal efficacy of a physical realizer excludes the causal efficacy of the mental property it realizes, and *vice versa*.[167]

Of course, if both Yablo and Menzies accept an Exclusion principle framed in terms of causes or causally efficacious properties, rather than in terms of causal sufficiency, and if this is the principle that Kim (and myself) rely on when generating the Exclusion problem, one might wonder how Yablo and Menzies expect to solve the Exclusion problem by rejecting a stronger principle. The answer is that both Yablo and Menzies frame the Closure principle in terms of causally sufficient properties as well. When setting up the Exclusion problem, this is how Menzies defines the Closure principle:

(Closure*): For every physical effect E that has a cause, there is a physical property *P* that is causally sufficient for E.[168]

---

[166] Yablo is well aware of this, stating, as mentioned above, that determinable and determinate properties compete for the role of being a cause. This is what leads him to deny that *scarlet* is causally efficacious with respect to Sophie's pecking. As I argued in chapter 4, I think this is too strong and not something Yablo should feel forced to commit to, partly because Yablo already denies that satisfaction of the proportionality constraints is necessary for causation (though for different reasons). Menzies and List, on the other hand, clearly take difference-making (as they spell it out) to be both necessary and sufficient for causation.

[167] Menzies [2008], p. 212-215.

[168] Menzies [2008], p. 198. Yablo's version of this is basically identical to Menzies', though he frames it in terms of events and calls it the principle of physical determinism. See Yablo [1992], p. 247.

Since being causally sufficient does not entail being a sufficient cause (or even a partial cause) on Yablo and Menzies' accounts, the truth of Closure* is compatible with the failure of Closure – even if some physical effects do not have physical causes (at time *t*), it may nevertheless be the case that for every physical effect, there is a physical property that is causally sufficient for that effect. In other words, while Exclusion* is stronger than Exclusion, Closure* is weaker than Closure. This leaves room for a solution to the Exclusion problem that does not involve the rejection of any of its premises: Accept Exclusion and Closure* (the weaker versions of these principles), and reject the stronger Exclusion* and Closure. This leaves us with four claims: First, that mental properties are distinct from physical properties (Distinctness). Second, that some physical effects have mental properties as causes, either as sufficient causes or as part of their sufficient causes (MC). Third, that these physical effects (in general) have just a single sufficient cause (Exclusion). And finally fourth, that there is a physical property (or set of physical properties), that is causally sufficient for each of these physical effects (Closure*). Again, if causal sufficiency does not entail being a cause, these four claims are entirely compatible. [169]

*6.4 Replacing Closure with Closure**

Obviously, the crucial, and controversial, step in Yablo and Menzies solution to the Exclusion is the rejection of the stronger principle, Closure. As I stated in chapter 3, however, a satisfactory solution to the Exclusion problem cannot simply amount to a rejection of one (or more) of the problematic premises. It must be supplemented with a story of why we came to believe that premise in the first place, or why the evidence for it is really only evidence for a weaker principle. Unfortunately, Yablo, does not discuss the stronger version of the closure principle, and so, does not give his reasons for thinking it legitimate to reject it and replace it with the weaker Closure*. Let me therefore put Yablo aside and focus on Menzies, who explicitly rejects the stronger version of the closure principle by way of a simple counterexample. Imagine that we have excellent empirical reasons for thinking that an individual instantiating mental property *M* is the difference-maker for whether or not that individual exhibits behaviour B. In this case, Menzies is prepared to take

[169] For similar arguments see Woodward [2008] and Raatikainen [forthcoming].

this at face value and as evidence that the stronger closure principle fails. As he says:

> […] M and a neural property that realizes it cannot both be difference-making causes of B. So our epistemic situation is one in which we have to decide between the well-confirmed that M is the cause of B and the purely conjectural hypothesis that there exists some physical P that is a difference-making cause of B. It would not be irrational under these circumstances to favour the first hypothesis over the second, concluding that the strengthened causal closure principle is false in this case.[170]

A couple of things should be noted here. First, saying that the stronger version of the closure principle is a 'purely conjectural hypothesis' is somewhat unfair. As I pointed out in chapter 3, the stronger version of the closure principle is itself supposedly backed up by empirical evidence, and if it was purely conjectural, it is unlikely that so many have worried about the Exclusion problem, rather than simply giving up the closure principle. Second, it is important to note that Menzies is not merely pointing out that mental properties turn out causally efficacious on the difference-making account of causation. As noted earlier, giving an account of causation that vindicates the causal efficacy of mental properties is all well and good. However, it does not in itself provide a solution to the Exclusion problem – in fact, as I stressed earlier, the causal efficacy of the mental is a starting assumption for generating the problem in the first place. Rather, the crucial point to note is that Menzies' account of causation has as consequence that when a mental property is causally efficacious, its physical realizer is *not*. This is why the case that Menzies asks us to imagine provides a counterexample to the stronger version of the closure principle.

That said, some might object that Menzies is doing little to actually solve the Exclusion problem. After all, it could be argued that what he is doing is using three claims to show that a fourth is false: Given the truth of Exclusion and Distinctness, a case of mental causation provides a counterexample to Closure. But this is old news. The Exclusion problem constitutes a paradox exactly because any three of its premises can be used to establish the denial of a claim that has independent support. Granted, Menzies' account of causation may vindicate the Exclusion principle and

[170] Menzies [2008], p. 216.

thereby provide novel and additional support for that principle, but that in itself does not help to solve the Exclusion problem. Recall, a dialectically satisfactory solution to the problem that involves the rejection, disambiguation, or amendment of one of the premises must come with a story of why the *evidence* for the rejected, disambiguated, or amended premise must be rejected or why it fails to support the premise under attack. While Menzies does not himself address this issue, I believe a case can be made on his behalf that his account has the resources to make plausible the claim that the evidence for Closure is really only evidence for the weaker Closure*.

6.5 The Empirical Evidence for Closure

Firstly, given Menzies' non-standard interpretation of the counterfactuals used to analyse causation, it is easy to see how a causally sufficient property might be mistaken for a causally efficacious property. Consider the following case: Jones has a pigeon and two objects that differ only in their colours - say that one is scarlet and one is lime-green. Jones determines that the pigeon consistently pecks at the scarlet object and fails to peck at the lime-green object. In this case, it seems Jones has some evidence for the claim that scarlet is the causally efficacious property with respect to the pecking. However, the case is entirely compatible with *scarlet* merely being causally sufficient for this effect, since the pigeon may be responding to the property *red*, rather than *scarlet*. Now, it might be argued that if Jones is a thorough researcher, he will take into account that the objects at his disposal also differ with respect to their determinable properties (one is red, the other green), and therefore will take into account the possibility that, in fact, it is *red* causing the pecking, rather than *scarlet*. Presumably, Jones would then test for this by presenting the pigeon with a variety of red but non-scarlet objects and check if pecking ensues. In other words, it is unlikely that Jones would make a mistake with respect to which property is causing the pecking. But now consider a scenario where Jones is part of a context in which some types of properties are typically ignored. For instance, say that Jones is a neuroscientist, that the 'objects' at his disposal are test-subjects, and that the property tested for causal efficacy with respect to some effect E is neural property *N*. As before, Jones discovers that whenever his test-subjects instantiate *N*, it consistently results in effect E (say, a certain kind of behaviour) and that whenever *N* fails to be instantiated, the effect fails to occur as well. Again, it seems

that in this case, Jones has some evidence for the claim that *N* is causing E. But now add that *N* is a realizer for mental property *M* and that, in fact, it is *M* that makes the difference between the effect occurring or not. Should we expect that Jones (or his colleagues) considers this possibility? I think not. It is not unlikely that Jones would fail to do so and therefore fail to test for the causal efficacy of mental property *M* with respect to E. As a neuroscientist, Jones is looking for *neural* causes and potential non-physical psychological causes will simply fall outside the scope of his research.[171] Even if Jones discovers that some test-subjects not instantiating *N*, but rather *N**, still exhibit the relevant behavioural effect E (just as it can discovered that a pigeon pecks at other red but non-scarlet objects), this is unlikely to lead to the conclusion that there might be a causally efficacious higher-order property, realized by causally inefficacious but nevertheless causally sufficient properties *N* and *N** (inefficacious and sufficient with respect to E, naturally). Instead, Jones is more likely to conclude that *N* and *N** are distinct properties causing the same type of effect. In other words, scientific disciplines are typically methodologically "closed off" from other disciplines. For instance, a chemist will look for chemical causes, a biologist for biological causes, a neuroscientist for neural causes, and so on. However, they will generally not consider potential causally efficacious properties treated by other disciplines. If scientific disciplines in general are closed off from causally efficacious properties treated by other disciplines, then we should expect that they systematically mistake properties that are merely causally sufficient for properties that are causally efficacious. And if so, it is no surprise that natural science appears to produce evidence for the stronger closure principle.

Secondly, the success of natural science, which was one of the main motivations for accepting Closure, is not discredited by the considerations above. Say that, indeed, neuroscience mistakenly and systematically classifies merely causally sufficient neural properties, with respect to certain behavioural effects, as causally efficacious. This will in no way impugn on the ability of neuroscience to make successful predictions about which effects will occur given the instantiation of certain neural properties. What is required for this is that there are stable correlations between the

[171] Even if the possibility of a mental cause did occur to Jones, consider what would happen if he was to ask the neuroscience department that employs him for additional research funding to investigate whether behaviour has psychological, rather than neural causes. He would almost certainly be turned down.

instantiation of certain neural properties and certain effects. And if certain neural properties are causally sufficient for behavioural effects caused by the mental properties they realize, this is guaranteed by satisfaction of the first counterfactual in the difference-making analysis of causation.

To sum up then, given Menzies account of causation and causal sufficiency, the empirical evidence arguably supports nothing more than the weaker formulation of the closure principle, Closure*. And replacing Closure with Closure* has as a result that the Exclusion problem fails to constitute a paradox. Therefore, a case can be made that Menzies provides a dialectically satisfactory solution to the Exclusion problem.

6.6 Difference-making and the Subset View

As I mentioned earlier, the truth of the Subset view is compatible with there being alternative sound solutions to the Exclusion problem. But is Menzies' view one of them? The crucial question is this: Is it compatible with Menzies' account that mental properties and their physical realizers stand in the subset relation? That is, can Menzies accept that the token-powers had by an instance of any mental property are identical to token-powers had by the instance of a physical property that realizes it? What I have to say here is very much the same as I said when discussing Yablo's view and whether or not it is compatible with the Subset view. If a property is a cause of an effect when it bestows a power to cause E that manifests, then Menzies' view is clearly incompatible with the Subset view, for it is part of the latter view that whenever a mental property manifests a power and causes an effect, its physical realizer manifests the very same power, and thereby also causes the effect. And Menzies *denies* that a physical realizer causes the effects also caused by the property it realizes.

The incompatibility, however, can be more or less severe. Consider the claim that mental and other non-physical properties and their physical realizers stand in the subset relation. In other words, the token-powers had by any mental or otherwise non-physical property are also had by the physical property that realizes it. Menzies can accept this as long as he denies that manifesting a power to cause E is sufficient for causing E. And presumably he would deny this since a property can manifest a

power to cause E without being a difference-maker for E, the latter condition being necessary for causation on Menzies' account. This, it seems to me, is not a severe disagreement (though as I argued in chapter 4, I find it hard to see the motivation for denying that realizers do cause behavioural effects if they manifest powers to cause these effects). After all, the proponent of the Subset view and Menzies would then agree on the metaphysics of the realization relation and mental causation, they just disagree on how best to classify the properties involved. According to Menzies, whenever a mental property manifests a power to cause an effect for which it is also a difference-maker, its physical realizer manifests the very same power, though only the former deserves the title of being a cause of E. According to the Subset view, whenever a mental property manifests a power to cause an effect for which it is also a difference-maker, its physical realizer manifests the very same power, and both deserve the title of being a cause of E. There might even just be two notions of causal efficacy – one demanding notion according to which causes must satisfy the criteria for difference-making, another less demanding according to which manifesting a power to cause E is enough. Both are legitimate notions of causal efficacy and can peacefully coexist. Correspondingly, the amendment of the Closure principle that Menzies' endorses is not a major one. While it is false that every physical effect has a physical cause in the more demanding sense, it will be true that every physical effect has a physical cause in the less demanding sense.[172] While I am inclined to think that the criteria for difference-making reveal explanatory features of a property, I will not argue against the above interpretation of Menzies' view, nor does it seem to me to pose a serious challenge to the Subset view.

On the other hand, Menzies view may be interpreted in a different manner that is more troublesome from the perspective of anyone endorsing the Subset view. In particular, Menzies may deny that physical realizers have powers to cause effects

[172] Note that while manifesting a power to cause E is weaker than causing E on the difference-making account of causation, it is stronger than being causally sufficient for E. On Menzies' account, blinking my eyes is causally sufficient for my hair growing (any world relevantly similar to the actual world in which I blink are worlds where my hair grows). However, the event of blinking my eyes presumably involves no property that manifests a power to cause the growth of my hair. This means that on the interpretation of Menzies' view proposed here, he could accept the following Closure principle: For every physical effect E, there is a physical property P that manifests a power to cause E.

caused by the mental properties they realize. This would present a more substantial disagreement with the Subset view. Now the disagreement would not just be over how to classify the properties involved in a case of mental causation, it would be over both the metaphysics of mental causation and the realization relation – when a mental property manifests a power to cause E, it's physical realizer does not. Hence, standing in the subset relation cannot be necessary for realization. This interpretation of Menzies' view is the more interesting one since it is clearly one that the proponent of the Subset view cannot accept. In the following, I will consider some reasons for rejecting it.

6.7 Against Menzies' Account

Since the latter interpretation of Menzies' view is incompatible with the Subset view, it is in the interest of the proponent of the Subset view to argue against it, or at least show that the Subset view presents a more attractive alternative. I see at least two potential problems for Menzies' account. The first is that it leaves mental-on-physical supervenience unexplained. On the Subset view, the reason why a mental property *M* is always instantiated when its physical realizer *P* is instantiated is that the two properties stand in the subset relation. But this explanation is unavailable to Menzies since *M* has token-powers that *P* does not. Unless Menzies has an alternative account of realization that explains the supervenience, it seems that he is forced to posit it as a brute modal fact. This, surely, is an unattractive consequence of his view.

A related worry is that the view sounds very much like a form of emergentism. Note that if we accept Menzies' definition of causal sufficiency, it arguably follows from emergentism that there are causally sufficient physical properties for effects that are caused by mental properties. Say that *P* does not realize *M* but is instead the emergence base for *M*. Since there are strong modal connections between emergent properties and their emergence bases, it will plausibly be the case that in possible worlds relevantly similar to the actual world in which *P* is instantiated, the effects of *M* will occur (since *M* emerges from *P* in those worlds). On the other hand, assuming that *M* has multiple emergence bases, there will be worlds relevantly similar to the actual world in which one of these is instantiated instead of *P*, and in which the effects of *M* therefore occur. So *P* is causally sufficient for *M*'s effects on

an emergentist account. This also means that an emergentist can accept the weaker closure principle, Closure*. If Menzies' view is different from emergentism it must be because the causal laws in virtue of which *M* causes certain effects are non-fundamental. Contrary to the emergentist's claim, they must ultimately be grounded in laws posited by completed physics and if so, must supervene on them. But this introduces another modal fact that should be explained. Why is it that whenever we have the fundamental physical laws, we are guaranteed to have the causal laws that govern the behaviour of mental properties? On the Subset view, since there are no causal powers at the level of mental properties not already present at the level of their physical realizers, the laws that govern the latter are enough to account for the behaviour of the former. Hence, mental causal laws are simply identical to physical causal laws and trivially supervene upon them.[173] On Menzies' view, on the other hand, it appears to be somewhat of a mystery why there are laws connecting mental and physical properties (assuming that Menzies' does not want to posit these laws as fundamental).

6.8 Concluding Remarks on Menzies' View

From a physicalist perspective, the Subset view presents a more attractive position than the position endorsed by Menzies, since it gives an explanation of *why* mental properties supervene on physical properties. It is worth noting as well that the Subset view can take onboard some of the important insights of the difference-making account of causation. As I mentioned in chapter 4.10, the proportionality constraints on causation may be interpreted as locating the more relevant cause, from an explanatory perspective, among the candidates. Likewise, Menzies' account of difference-making may be seen as serving the same purpose – in ordinary explanatory contexts, citing the property that makes the difference for an effect will provide the most useful causal explanation of this effect. And this the Subset view can easily allow for. The proponent of the Subset view need not deny the importance of difference-making for explanatory purposes.

[173] Again, this is assuming that emergentism about broadly physical properties is false and that any causal laws that govern the behaviour of these properties are grounded in fundamental physical laws.

6.9 Disambiguating 'Property'

Let me now turn from Yablo and Menzies' attempts to solve the Exclusion problem by amending the closure principle, to an approach that seeks to solve the problem by disambiguating the notion of a property. The approach, developed independently by David Robb and Douglas Ehring (in the following I focus on Robb's solution), is based on a trope theory of properties.[174] First, I will sketch the trope theory and explain why Robb takes it to solve the Exclusion problem. Second, I will present an objection against his solution raised by Sophie Gibb and consider a response on behalf of Robb. Third and finally, I will extent Gibb's objection in order to show that Robb's solution is unlikely to work.

6.10 The Trope Theory

The Trope theory is a view about the nature of properties and objects.[175] As for the former, the Trope theory claims that properties are abstract particulars, rather than universals. These abstract particulars are called tropes. Being particulars, tropes have unique location. For instance, the redness of my shirt is present where my shirt is and nowhere else. Objects themselves are typically taken to be bundles of tropes. An object *o* thus has a certain trope, if and only if that trope is part of the bundle of tropes that is identical to *o*. As pointed out in chapter 3.18, the notion of parthood is given a specific interpretation by trope theorists as that of *compresent parthood*. Compresent parts occupy the same spatial region. For instance, the redness that is part of the bundle of tropes that is identical to my shirt occupies the same spatial region as each and every other trope that is a compresent part of this bundle.

The particularity of tropes raises the question of how trope theorists account for the apparent fact that distinct objects may nevertheless be identical in some respect. How is it that objects that are wholly distinct may nevertheless be of the same type? Typically this is analysed in terms of the sharing of properties – two wholly distinct objects are identical in some respect in virtue of sharing at least one property. For instance, two distinct square objects are identical with respect to their shape-property, *squareness*, which they share. If properties are universals, the identity can be taken at face value. Since universals are fully present at each of their instances,

[174] See Robb [1997] and Ehring [2003].
[175] For a comprehensive defence of the position, see Campbell [1990].

there is some property of each and every square object that is *literally* the same and which all square objects literally share. But since tropes are particulars, there is nothing about different square objects that is literally the same, nor can tropes literally be shared by distinct objects. Instead, trope theorists tend to analyse the apparent identity, or sharing of properties, in terms of resemblance between tropes.[176] Two square objects are identical in some respect in the sense that they each have two resembling square-tropes as compresent parts. Such resembling tropes form resemblance-classes and these classes play the role of type-properties in a trope ontology. For instance, the class of all exactly resembling square-tropes is identical to the type-property *square*. Different objects are of that type if and only if they have a trope as a compresent part that is a member of this class.

6.11 The Trope Solution to the Exclusion Problem

The presentation of the trope theory reveals an ambiguity in the notion of a property. It can either refer to the qualities or features of a particular object, for instance the redness of my shirt, in which case 'property' refers to a trope. Or, it can refer to types of properties, for instance the quality or feature that all red objects have in common, in which case 'property' refers to a class of resembling tropes.[177] Robb proposes to exploit this ambiguity in order to solve the Exclusion problem. Consider first the claim that mental and physical properties are distinct (Distinctness). What does the word 'properties' refer to in this claim – tropes or classes of resembling tropes? Recall, Distinctness is motivated by the argument from multiple realization which tells us that different individuals may share a mental property *M* without sharing any relevant physical properties. On a trope ontology, this is interpreted as the claim that, one, different individuals each have distinct *M*-tropes as compresent parts and that each of these tropes are members of *M* (the class of all resembling *M*-tropes), and two, that these individuals have no relevant physical tropes that are members of the same class of resembling tropes. Assuming a trope ontology, the argument from multiple realization thus makes claims about the resemblance relations between pairs of tropes had by different

[176] Resemblance must be taken a primitive here rather than as a relational type, to avoid an infinite regress. If resemblance was a relational type, the question arises why certain tropes are members of this type. To answer that it is in virtue of resembling would be to introduce a new type with members that resemble, and so on.

[177] On an account of properties as universals, there is no such distinction to be made. The redness of my shirt is the very thing that each red object have in common.

individuals. There are pairs of *M*-tropes, each had by some individual, that resemble, while there is no pair of relevant physical tropes, each had by one of these individuals, that resemble. Note that this blocks reduction at the type-level only. Mental type *M* (i.e. the set of resembling *M*-tropes) cannot be identical to physical type *P* (i.e. the set of resembling *P*-tropes), since two individuals can have each have a trope that are members of the M-set, without having relevant tropes that are members of the same P-set (i.e. a set of resembling physical tropes).

So much for Distinctness. The word 'properties' also figure in the claim that mental properties are sometimes causes (MC). What does the word refer to here – tropes or classes of resembling tropes? According to Robb, it's the former. As he says: "[…] it is hard to see how a class could be causally relevant to producing an effect. Types as classes just do not seem like the right sort of entities to do this work."[178] This means that MC is the claim that mental tropes are sometimes causes. If tropes, rather than classes of tropes, are causally efficacious, this also means that the Exclusion principle and the Closure principle should also be formulated in terms of tropes. The Exclusion principle would then be the claim that if trope *C* is a sufficient cause of effect E, then all other tropes distinct from *C* is excluded from causing E, and the Closure principle would (roughly) be the claim that for every physical effect, there is a physical trope that is a sufficient cause of that effect. Now, in the original formulation of the Exclusion problem, MC, Exclusion, and Closure, together entail the denial of Distinctness. This is a problem since it seems we have good reasons for thinking that Distinctness is true. However, given a trope reading of these claims, they entail only that any mental trope is identical to some physical trope. And as we just saw, on a trope ontology, Distinctness does not make a claim about tropes, but rather about classes of tropes. This dissolves the Exclusion problem. The trope theorist can accept that any *class* of mental tropes is distinct from any *class* of physical tropes, while still insisting that any causally efficacious mental *trope* is identical to some physical *trope*, thereby securing mental causation without compromising Exclusion or Closure.

[178] Robb [1997], p. 192.

6.12 The Trope Solution and the Subset View

Prima facie, the trope solution to the Exclusion problem is attractive since it does not require us to reject or amend any of the four claims constituting the problem. But is it compatible with the Subset view? First, note that the proponent of the Subset view can accept the kind of trope ontology sketched above. The Subset view would then be that the subset relation holds between pairs of mental and physical tropes – that is, the causal profile of any mental trope will form a proper non-empty subset of the profile of the physical trope that realizes it. However, according to the trope solution, the relation between these two tropes is not realization but identity. And this is incompatible with the Subset view since the causal profiles of these two tropes will then be identical as well. Of course, accepting a trope ontology does not necessarily commit one to accept the trope solution. Nevertheless, if a trope ontology is accepted, this makes it less obvious why one should accept the Subset view. Recall, in chapter 4 I motivated the Subset view via the four statements of the Exclusion problem, one of them being the amended version of the Exclusion principle. To briefly recap, the motivation had four steps, one for each of these statements:

i) Mental properties have powers to cause physical effects (from MC).
ii) These powers are also had by physical properties (from Closure).
iii) These physical properties have additional causal powers, not had by mental properties (from Distinctness).
iv) The powers shared by mental and physical properties are identical at both the type and token level. That is, the token-powers had by an instance of a mental property are numerically identical to token-powers had by some instance of a physical property (from the amended Exclusion principle which prohibits strong systematic causal overdetermination).

Note, however, that the third step is invalid assuming a trope ontology. On a trope ontology, Distinctness between mental and physical types will *not* show up as difference in the causal profiles for these types, for the simple reason that types are not the kind of entities to have causal profiles at all. Rather, a difference between mental and physical types will show up in a difference between the members of the classes that are identical to these types. Thus, the most that can be inferred from the

argument from multiple realization is that any mental type *M* (i.e. the class of resembling *M*-tropes) must have at least one member that is not a member of physical type *P* (i.e. the class of resembling *P*-tropes). This, however, tells us nothing about the relation between the causal profiles of mental and physical tropes. All we know is that mental tropes have causal powers (from MC), that these powers are also had by some physical trope (from Closure), and that these powers are identical at both the type and token level (from the amended Exclusion principle). Now, all this would be true if mental and physical tropes stand in the subset relation. However, it would also be true if mental tropes are identical to physical tropes. The problem is that the four statements constituting the Exclusion problem gives us no reason to prefer one over the other.

Of course, the proponent of the Subset view may argue that there are independent reasons for preferring the Subset view over the reductionist view entailed by the trope solution (reductionist at the level of tropes, not at types, naturally). For one thing, it may be that there are reasons to think that a trope ontology is false, in which case the trope solution would never even get off the ground. Alternatively, it may be possible to extend the original argument from multiple realization to mental tropes and argue that a mental trope can be realized by different physical tropes - the very same mental trope that Jones has when he is in pain *could* have been realized by a different physical trope than it actually is. It might even be the case that mental tropes are multiply realized – the mental trope Jones has now and which is realized by some physical trope, is the very same mental trope he had a few minutes ago where it was realized by a different physical trope.[179] Such an argument would have to say more about the persistence conditions of tropes, though in any case, I think most will agree that the soundness of this version of the argument from multiple realization is a lot less intuitive than is the case with the original version. Fortunately, there is a more straightforward route to the denial of the kind of trope reduction entailed by the trope solution. As I will argue, this kind of reduction either won't work or it will work but probably not in the case of mental tropes.

[179] This would not even require that the different realizers of a mental trope are members of different resemblance classes – they could all be physical tropes of the same type.

6.13 Gibb: Trope-identity Entails Type-identity

Let me first present an argument by Sophie Gibb that purports to show that identity between tropes entails identity between the resemblance classes of which these tropes are members.[180] Later I will consider a reply on behalf of Robb that I argue is unlikely to work in the case of psychophysical trope reduction.

Gibb's argument rests on the assumption that all and only classes of exactly resembling tropes play the role played by universals and are genuine types. In other words, two tropes T1 and T2 exactly resemble if and only if they are of the same type. Hence, if two mental tropes M1 and M2 are of the same type, they must exactly resemble each other. According to the trope solution, each of these mental tropes is identical to some physical trope. Say that M1 is identical to P1 and M2 is identical to P2. But if two tropes are identical, it should be the case that they enter into the same resemblance relations. That is, if M1 exactly resembles M2, then P1 must exactly resemble M2 (since P1 = M1). And if M2 exactly resembles P1, then P2 must exactly resemble P1 as well (since M2 = P2). Given Gibb's assumption that all and only classes of exactly resembling tropes are types, this means that P1 and P2 are of the same type. And likewise for all the physical tropes that form the reduction-base for mental tropes of type *M*. This, however, is in direct contradiction to the argument from multiple realization according to which two individuals can share a mental type (i.e. each have a mental trope that is a member of the same resemblance class) without sharing any relevant physical type (i.e. among the physical tropes had by these two individuals, there are no relevant pair that are members of the same resemblance class). What Gibb's argument shows is that if two individuals share a mental type, there is guaranteed to be a relevant physical type they both share as well.

6.14 From Exact to Inexact Resemblance between Tropes of the Same Type

Gibb herself considers and rejects a possible rejoinder from the proponent of the trope solution. If tropes have aspects, then two tropes may resemble exactly with respect to one aspect and not another. In particular, a class of mental tropes may

[180] Gibb [2003].

resemble each other exactly with respect to their mental aspect and fail to resemble exactly with respect to their physical aspect. This would then allow for multiple realization since the class of physical tropes that form the reduction base for a mental type would have members that fail to resemble exactly with respect to their *physical* aspect and so, fail to form a genuine *physical* type. Gibb offers two reasons for thinking that this presents an unattractive patch on the trope solution. Firstly, Gibb does not believe properties are the kind of entities that have aspects. As she says: "The claim that there are property aspects seems to be based upon a misguided ontology and hence should be rejected". [181] Secondly, even if properties do have aspects, it would not help the proponent of the trope solution since the Exclusion problem would then simply reappear at the level of aspects. For any case of mental causation, we could ask whether the causally efficacious trope was efficacious in virtue of its mental aspect or in virtue of its physical aspect. The closure principle pushes us towards going for the latter option, which in turn makes the claim that there are *mental* causes questionable (at least, a mental trope would not be causally efficacious in virtue of its mental aspect). So the trope solution is untenable whether or not properties have aspects. If they do not, then following Gibb's argument, it seems there can be no multiple realization, which is the prime motivation for accepting the distinctness between mental and physical types. If properties do have aspects, then the trope solution can allow for multiple realization, however the Exclusion problem is simply moved to the level of aspects rather than tropes. In either case, it is bad news for the trope solution.

I think more can be said here on behalf of the trope solution, though ultimately I agree with Gibb that it does not solve the Exclusion problem. First, we might question Gibb's assumption that only classes of *exactly* resembling tropes can form genuine types. Consider trope-reduction applied to determinable tropes. This would be the view that any determinable trope is identical to a maximally determinate trope. For instance, a certain red-trope is identical to a scarlet-trope, another red-trope identical to a crimson-trope, another to a pink-trope, etc. The type *red* would then be the class of all scarlet-tropes, plus all crimson-tropes, plus all pink-tropes, and so on for all other shades of red. Do the members of this class exactly resemble

[181] Gibb [2003], p. 475.

each other? No. Pink, for instance, does not exactly resemble scarlet. If it did, pink and scarlet would not be *different* shades of red. But to infer from this that this precludes *red* from being genuine type seems too quick. Even if different shades of red do not exactly resemble each other, there is still an intuition that they form a unity – 'red' really does designate a genuine type in nature and scarlet-trope has a nature that is closer to that of crimson-trope than, say, a mass-trope.[182] The best way to make sense of these intuitions is to allow that the tropes forming the type *red* resemble each other to a high, though not exact, degree. Of course, some of them *will* exactly resemble each other, namely those tropes that form a maximally determinate type, say *scarlet*. But this is not a necessary requirement for tropes to be parts of the *class* red.

Now, the question is, is this point applicable to mental and physical tropes? That is, could it be that a mental type is a class of physical tropes that resemble to a high, though not exact, degree? If so, Gibb's argument for why trope-reduction rules out multiple realization fails to go through. This argument relies on the assumption that two mental tropes of the same type must exactly resemble each other, which then, contra the argument from realization, means that the physical tropes to which they reduce must exactly resemble each other as well. But if mental tropes do not need to exactly resemble each other to be of the same type, then we can allow that different mental tropes of the *same* type reduce to physical tropes that are *different* types, just like two tropes both of the same type *red* can reduce to tropes that are of different types, *scarlet* and *crimson*. To answer this question, consider what it takes for one thing to resemble another. Resemblance is always with respect to something. For instance, two persons may resemble each other with respect to their height or mass or hair-colour. Of course, height, mass, and hair-colour are here properties of these persons. Two individuals are similar or resemble each other in virtue of sharing at least one property. But what if the resemblance is between properties or tropes themselves? Of course, different tropes can resemble, in some sense of the word at least, with respect to extrinsic relational properties. For instance, the mass-trope and the height-trope that are compresent parts of the bundle of tropes that is identical to

[182] Those with qualms about the reality of colours can substitute with some other determinable property.

my body are both in my office as I write this. However, this is not the sense of resemblance that is relevant here. As Gibb says:

> […] resemblance is an internal relation. A trope is a member of a certain resemblance class because of what it is. Whether or not two tropes resemble each other wholly depends on their particular natures.[183]

In other words, to the extent that there is a notion of resemblance with respect to extrinsic properties, it is not the notion of resemblance the trope theorist has in mind when she talks of tropes resembling each other.

6.15 Determination Dimensions and Resemblance between Tropes

To flesh how resemblance depends on the nature of tropes, it is useful introduce a framework for properties developed by Eric Funkhouser.[184] Properties, Funkhouser claims, have *determination dimensions*. These are variables that individuate the property in question. Some properties have just a single determination dimension. For instance, *mass* is determined by a mass unit and nothing else – if two mass-tropes have the same value of mass unit (say 1 kilogram), then they are exactly the same mass. Or given a trope ontology, they resemble exactly. There is simply no further question of whether or not these two mass-tropes are the same since there is no further dimension along which they can differ. Other properties have more than one determination dimension. For instance, colours are plausibly individuated along three dimensions – hue, saturation, and brightness. If two tropes have exactly the same values of hue, saturation, and brightness, then they are exactly the same colour, or they resemble exactly. Funkhouser illustrates this mathematically as points in an *n*-dimensional space.[185] In the case of *mass*, the space is 1-dimensional. There is but a single axis and every point on that axis corresponds to a class of exactly resembling mass-tropes, for instance the class of all one kilogram-tropes. In the case of *colour*, the space is 3-dimensional and every point in that 3-dimensional space corresponds to a class of exactly resembling colour-tropes, for instance the class of all scarlet-tropes. Are all types represented as a point in the relevant property-space? If the answer is yes, Gibb would be right that only classes of exactly resembling tropes are genuine types. But we can now see how this leaves no

[183] Gibb [2003], pp. 470-471.
[184] In Funkhouser [2006].
[185] Funkhouser [2006], pp. 554-555.

room for less than maximally determinate types. For instance, the determinable type *red* cannot be represented as a point in property-space for *colour*, since the determination dimensions for the members of this class do not take the same values – a scarlet-trope and a crimson-trope will both be members of the type *red*, yet their determination dimensions take slightly different values, and so, each of these tropes will be represented as different points in the property-space. Instead, determinable types will be extended over an area of the relevant property-space. Within this area will be the points that correspond to all determinates of the determinable type. A trope T is thus a member of all the types that are represented as areas of the relevant property-space that includes the point that represents the class of exactly resembling T-tropes.

Given Funkhouser's framework for properties, we can see more clearly how resemblance between tropes depends on their nature. The nature of a trope is the value of its determination dimensions. Two tropes exactly resemble each other if they have the same determination dimensions and those dimensions have the exact same value. They resemble to a relatively high, though not exact, degree if they have the same determination dimensions and those dimensions take very similar, though not exactly the same, values. They resemble to a relatively low degree if they have the same determination dimensions and those dimensions take very different values. What about tropes that have no common determination dimensions? It seems to me the best thing to say is that their degree of resemble is zero. For instance, a scarlet-trope resembles a mass-trope as much as it resembles a height-trope or a pitch-trope – i.e. not at all. As I mentioned earlier, resemblance is always with respect to something. Resemblance between objects, for instance, is with respect to properties. In the case of tropes, resembles is with respect to their determination dimensions, with the degree of resemblance determined by the value of these. So for tropes with no common determination dimensions, there is nothing internal with respect to which they can resemble each other. And this seems right. It is difficult to make sense of the claim that a scarlet-trope could resemble a mass-trope or a height-trope.

6.16 Determination Dimensions and Psychophysical Trope-Reduction

So what does this have to do with the proposal that mental types are classes of inexactly resembling tropes? Well, consider again the claim that mental trope M1 is identical to physical trope P1, that mental trope M2 is identical to physical trope P2, and that M1 and M2 are of the same mental type. Gibb's argument was that since M1 and M2 must exactly resemble, this means that, necessarily, P1 and P2 must exactly resemble as well, and this is what the argument from multiple realization rules out. But this seems too strong since M1 and M2 might inexactly resemble while still being of the same type. What the discussion above shows, however, is that if M1 and M2 resemble at all, it is with respect to their determination dimensions. And since P1 and P2 are identical to M1 and M2, respectively, this means that P1 and P2 must resemble with respect to those same determination dimensions. In other words, trope-reduction is possible only if the reduced trope and the trope it reduces to have the same determination dimensions. This is the case for determinable tropes and their corresponding determinates. For instance, scarlet-tropes and red-tropes are all individuated with respect to hue, saturation, and brightness. This is what allows for the fact that a particular scarlet-trope can be a member of several resemblance classes – it is a member of the resemblance class represented by a point in the property-space for colour (this is the resemblance class for which the members exactly resemble and which is identical to the type *scarlet*) and it is also a member of the resemblance class represented by an extended area of the property-space which includes this point (this is a resemblance class for which the members inexactly resemble, for instance the type *red*). But is this also the case for mental and physical tropes?

Unfortunately, there are good reasons for thinking that the answer to this question is negative: Mental and physical tropes do arguably not have the same determination dimension. For instance, the determination dimensions for *pain* are plausibly sharpness and intensity and Funkhouser proposes that the determination dimensions for the intentional property *believing that P* are the content of P and the degree of confidence with which one holds it.[186] But whatever the determination dimensions are for the physical tropes that could plausibly function as a reduction base for such

[186] Funkhouser [2006], p. 563.

mental tropes, they will almost certainly not include *those* determination dimensions. This is revealed by the plausibility of even maximally determinate mental types being multiply realizable. Say that the determination dimensions for pain are indeed sharpness and intensity and that two individuals each have a pain-trope with the exact same values of these dimensions. This means they exactly resemble and now we can run Gibb's original argument against the trope solution. Since the two mental tropes exactly resemble each other, the physical tropes to which they reduce must necessarily exactly resemble each other as well. But this seems unlikely. Surely it is possible for two individuals to have a pain with the exact same sharpness and intensity and nevertheless fail to share any relevant physical properties?[187]

6.17 Concluding Remarks on the Trope Solution

If mental and physical tropes do not have the same determination dimensions, psychophysical trope-reduction is blocked and the trope solution to the Exclusion problem fails. Note how this once again tips the scale in favour of the Subset view. If Distinctness applies not only to mental and physical types but also to mental and physical tropes, then the difference between mental and physical tropes must reveal itself somehow. Assuming a causal theory of properties, this must be in the form of a difference in the causal profiles of mental and physical tropes. And now we can get from the statements of the Exclusion problem to the Subset view: All powers contained in the profile of any mental trope must be identical to powers contained in the profile of some physical property (from Closure). This means that the profile of this physical property must contain additional causal powers (from Distinctness). And the powers which the mental and the physical property share must be identical at both the type and token level (from the amended Exclusion principle). This is the Subset view.

[187] Funkhouser [2006], p. 565. A similar point is made by Ehring in his [1996], pp. 472-473.

Conclusion

In this thesis I have defended and motivated a non-reductive view of mental properties which allows for their causal efficacy in world that is fundamentally physical. Mental causation fits in because it *is* physical causation. Every time a mental property manifests a power to cause an action, the very same power is manifested by a physical property. This has two advantages. Not only does it allow for the causal closure of the physical domain, it also explains why there is such an intimate connection between mental and physical properties – a mental property and its physical realizer are so intimately connected because the latter share all the causal powers of the former. This is fully compatible with mental properties having a multitude of distinct physical realizers. While any realizer of a mental property *M* must have all the powers of *M*, different physical realizers can be distinguished by whatever *additional* powers they have.

It is worth stressing again, however, that the Subset view does *not* present an *ad hoc* solution to the Exclusion problem. The claim is not that the Subset view should be accepted because it happens to solve the Exclusion problem. Such a claim gets the order of explanation the wrong way around. Rather, the Subset view *follows* from the statements that make up the Exclusion problem – statements that are individually motivated, independently of the Subset view. The amended Exclusion principle that allows for weak overdetermination plays a crucial role here. If the original Exclusion principle was amended to allow for weak overdetermination on the sole basis that this would open the door to the Subset view, then I think we should worry that the view presents an *ad hoc* solution to the Exclusion problem. However, as I argued, this is not the case. We have reasons to think that some effects are systematically and weakly overdetermined that are entirely independent of any motivation for the Subset view. This I consider a strong point in favour of accepting the Subset view.

In a certain sense, the Subset view is a reductive view. It holds that mental causal powers are identical to physical causal powers. For philosophers attracted to physicalism, this should be embraced. If genuine non-reductive physicalism is committed to the existence of irreducible mental causal powers, then I have doubts

that the position is a coherent one. Rather, the closest we get to non-reductive physicalism is emergentism. So the fact that the Subset identifies instances of mental causation with instances of physical causation is no strike against it. If we take the truth of physicalism seriously, this is to be expected. What matters is that the Subset view does not identify mental *causes* with physical causes. Mental properties are properties in their own right with distinctive causal profiles not found at physical level. This, I think, is all a non-reductive physicalist can hope and wish for.

Bibliography

Alexander, Samuel [1920]: *Space, Time, and Deity*, 2 vols., Macmillan.

Armstrong, David [1986]: 'In Defence of Structural Universals', *Australasian Journal of Philosophy* 64:1, 85-88.

Bailey, Andrew [1998]: 'Supervenience and Physicalism', *Synthese* 117, 53-73.

Baker, Lynne Rudder [1993]: 'Metaphysics and Mental Causation', in Heil & Mele [1993].

Beckermann, A., Flohr, H., Kim, J. (eds.) [1992]: *Emergence or Reduction*, De Gruyter.

Bennett, Karen [2003]: 'Why the Exclusion Problem Seems Intractable and How, Just Maybe, to Tract It', Nous 37(3), 471-497.

- [2004]: 'Global Supervenience and Dependence', *Philosophy and Phenomenological Research* 68(3), 501-529.

Bird, Alexander (2007): *Nature's Metaphysics: Laws and Properties*, OUP.

Block, Ned [1978]: 'Troubles with Functionalism', in Block [1980].

- [1980] (ed.): *Readings in Philosophy of Psychology Vol. 1*, Harvard University Press.

- [2003]: 'Do Causal Powers Drain Away?', *Philosophy and Phenomenological Research* 67, 133–150.

Bontly, Thomas [2002]: 'The Supervenience Argument Generalizes', *Philosophical Studies* 109, 75-96.

Braddon-Mitchell, David & Jackson, Frank [1996]: *Philosophy of Mind and Cognition: An Introduction*, Blackwell.

Broad, C.D. (1925): *Mind and Its Place in Nature*, Kegan Paul, Trench, Trubner & Co., Ltd.

Campbell, Keith [1990]: *Abstract Particulars*, Basil Blackwell.

Chalmers, David [1996]: *The Conscious Mind*, OUP.

Chalmer, David & Manley, David & Wasserman, Ryan [2009]: *Metametaphysics*, OUP.

Crane, Tim [2007]: 'Cosmic Hermeneutics vs. Emergence: The Challenge of the Explanatory Gap', in Penco, Beaney & Vignolo [2007]

Davidson, Donald [1987]: 'Knowing One's Own Mind', *Proceedings and Addresses of the American Philosophical Association* 60(3), 441-458.

Dovell, J.L. [2006]: 'Physicalism: Empirical, Not Metaphysical', *Philosophical Studies* 131, 25-60.

Ehring, Douglas [1996]: 'Mental Causation, Determinables and Property Instances', *Nous* 30:4, 461-480.

- [2003]: 'Part-Whole Physicalism and Mental Causation', *Synthese* 136, 359-388.

Fodor, Jerry [1974]: 'Special Sciences, of The Disunity of the Science as a Working Hypothesis', in Block [1980].

Funkhouser, Eric [2006]: 'The Determinable-Determinate Relation', *Nous* 40(3), 548-569.

Gibb, Sophie [2004]: 'Mental Causation and the Nature of Properties', Australasian *Journal of Philosophy* 82(3), 464-476.

Gillett, Carl & Loewer, Barry [2001]: *Physicalism and Its Discontents*, Cambridge University Press.

Gillett, Carl and Rives, Bradley [2005] 'The Non-existence of Determinable Properties: Or, a World of Absolute Determinates as a Default Hypothesis', *Nous* 39:3, 483-504.

Hawthorne, John [2002]: 'Blocking Definitions of Physicalism', *Philosophical Studies* 110, 103-113.

Heil, John & Mele, Alfred [1993]: *Mental Casuation*, Clarendon.

Horgan, Terence [1993]: 'From Supervenience to Superdupervenience', *Mind*, Vol. 102, No. 408, 555-586.

Hohwy, Jacob & Kallestrup, Jesper [2008]: *Being Reduced*, Oxford University Press

Howell, Robert [2009]: 'Emergentism and Supervenience Physicalism', *Australasian Journal of Philosophy* 87:1, 83-98

Hüttemann, Andreas [2004]: *What's Wrong with Microphysicalism?*, Routledge.

Jack, Anthony [1994]: 'Materialism and Supervenience', *Australasian Journal of Philosophy* 72, 426-443.

.Jackson, Frank [1998]: *From Metaphysics to Ethics*, Clarendon Press.

Kallestrup, Jesper [2006]: 'The Causal Exclusion Argument', Philosophical Studies 131(2), 459-485.

Kim, Jaegwon [1992a]: 'Multiple Realization and the Metaphysics of Reduction', *Philosophy and Phenomenological Research*, Vol. 52, No.1, 1-26.

- [1992b]: '"Downward Causation" in Emergentism and Non-Reductive Physicalism', in Beckermann, Flohr & Kim (eds.) (1992).
- [1993a]: 'Supervenience as a Philosophical Concept', in Kim (1993b).
- [1993b]: Supervenience and Mind, Cambridge University Press.
- [1997]: 'Does the Problem of Mental Causation Generalize?', *Proceedings of the Aristotelian Society*, 87, 281–297.
- [1998]: *Mind in a Physical World*, MIT Pres
- [2005]: *Physicalism, or Something Near Enough*, Princeton.
- [2010]: 'Thoughts on Sydney Shoemaker's *Physical Realization*', *Philosophical Studies* 148(1), 101-112.

Kim, Jaegwon & Sosa, Ernest (eds.) [1999]: *Metaphysics – An Anthology*, Blackwell.

Levine, Joseph [2001]: *The Purple Haze*, OUP.

Lewis, David [1972]: 'Psychophysical and Theoretical Identifications', in Block [1980].

- [1973]: 'Causation', *Journal of Philosophy* 70, 556–567.
- [1980]: 'Mad Pain and Martian Pain', in Block [1980].
- [1986a]: 'Against Structural Universals', *Australasian Journal of Philosophy*, 64:1, 25-46.
- [1986b]: *On the Plurality of Worlds*, Blackwell.

Lowe, E.J. [2000]: 'Causal Closure Principles and Emergentism', Philosophy 75(4), 571-585.

McLaughlin, Brian [2006]: 'Is Role-Functionalism Committed to Epiphenomenalism?', *Journal*

of Consciousness Studies 13, 39-66.

- [2007]: 'Mental Causation and Shoemaker-Realization', *Erkenntnis* 67, 149-172.

Melnyk, Andrew [1997]: 'How to keep the 'Physical' in Physicalism', *Journal of Philosophy* 94:12, 622-637.

- [2003]: *A Physicalist Manifesto,* Cambridge University Press.

Menzies, Peter [2008]: 'The Exclusion Problem, the Determination Relation, and Contrastive Casuation', in Hohwy and Kallestrup (eds.) [2008].

Menzies, Peter & List, Christian [unpublished]: 'Non-Reductive Physicalism and the Limits of the Exclusion Principle'.

Montero, Barbara [2006]: 'Physicalism in an Infinitely Decomposable World', *Erkenntnis* 64, 177-191.

Moore, Dwayne [2010]: The Generalization Problem and the Identity Solution, *Erkenntnis*, Vol. 75, 57-72.

Morgan, C.L. [1923]: *Emergent Evolution*, Williams and Norgate.

Moser, Paul & Trout, J.D. [1995]: 'Physicalism, Supervenience, and Dependence', in Savellos & Yalcin [1995].

Nagel, Thomas [1974]: 'What Is Is Like To Be a Bat?', *Philosophical Review* 83, 435-450.

Ney, Alyssa [2008]: 'Physicalism as an Attitude', *Philosophical Studies*, Vol. 138, 1-15.

Noordhof, Paul [1999]: 'Micro-Based Properties and the Supervenience Argument', *Proceedings of the Aristotelian Society* 99, 109-114.

Papineau, David [2001]: 'The Rise of Physicalism', in Gillett & Loewer (eds.) [2001].

- [2002]: *Thinking about Consciousness*, OUP.b

- [2008]: 'Must a Physicalist be a Microphysicalist?', in Hohwy & Kallestrup [2008].

Penco, Carlo, Beaney, Michael & Vignolo, Massimiliano (eds.) [2007]: *Explaining the Mental: Naturalist and Non-Naturalist Approaches to Mental Acts and Processes*, Cambridge Scholars Publishing.

Putnam, Hilary [1967]: 'The Nature of Mental States', in Block (1980).

Raatikkainen, Panu [forthcoming]: 'Casuation, Exclusion and the Special Sciences', *Erkenntnis*.

Robb, David [1997]: 'The Properties of Mental Causation', *Philosophical Quarterly* 47(187), 178-194.

Savellos, Elias & Yalcin, Ümit (eds.) [1995]: *Supervenience: New Essays*. Needham Heights: Cambridge.

Schaffer, Jonathan (2003): 'Is There a Fundamental Level?', *Noûs* 37:3, 498-517.

- [2005]: 'Contrastive Causation', *Philosophical Review* 114(3), 297-328.

- [2009]: 'On What Grounds What', in Chalmers & Manley & Wasserman (eds.)

Shoemaker, Sydney [1980]: 'Causality and Properties', in Kim & Sosa [1999]

- [2001]: 'Realization and Mental Causation', in Gillett & Loewer [2001].

- [2007]: *Physical Realization,* OUP.

Smart, J.J.C. [1959]: 'Sensations and Brain Processes', *Philosophical Review* 68, 141-156.

Tye, Michael [2009]: *Consciousness Revisited*, MIT Press.

Van Inwagen, Peter [1994]: 'Composition as Identity', *Philosophical Perspectives* 8, 207-220.

Williams, J.R.G [2006]: 'Illusions of Gunk', *Philosophical Perspectives* 20, 493-513.

Wilson, Jessica [2005]: ‘Supervenience-based Formulations of Physicalism’, Nous 39:3, 426-459.

- [2006]: ‘On Characterising the Physical’, Philosophical Studies 131, 61-99.

- [2009]. ‘Determination, Realization, and Mental Causation’, *Philosophical Studies* 145 (1), 149-169.

Woodward, James [2003]: *Making Things Happen*, OUP.

- [2008]: ‘Mental Causation and Neural Mechanisms’, in Hohwy & Kallestrup [2008].

Yablo, Stephen [1992] ‘Mental Causation’, *Philosophical Review* Vol. 101, No. 2, 245-280.

Yates, David [2009]: ‘Emergence, Downwards Causation and the Completeness of Physics’, *Philosophical Quarterly* 59 (234), 110-131.