

DOES THE CAUSAL EXCLUSION ARGUMENT HOLD IN A PROBABLISTIC SETTING?

ASSESSING THE EFFICACY OF MENTAL CAUSATION IN AN
INDETERMINISTIC WORLD



Lisa Grant

August 25, 2016

*Submitted in partial fulfillment to the requirements for the degree of
Master of Philosophical Studies (MPhil Stud)*

TO THE

*Department of Philosophy
University College London*

Supervised by Dr. Luke Fenton-Glynn

Declaration

I, **Lisa Grant**, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:



Date: 1/9/2016

Abstract

My thesis examines Kim's Casual Exclusion Argument (CEA) as an argument against the existence of mental causation when examined from within an indeterministic setting. One of our best current scientific theories, that is, orthodox quantum mechanics, lead us to believe our world is indeterministic, more precisely, probabilistic. However, our best current or most orthodox philosophical views, that is, physicalism and causal closure sit at odds with the idea of mental causation given an aversion to overdeterminism and identifying the mental and the physical.

If I am correct in my arguments then the CEA may only go through in deterministic worlds at least in its original form. I will therefore put forward a probabilistic analogue CEA. By examining the premises of this CEA I will argue that it doesn't hold. This is because I will argue that causal closure does not hold in probabilistic worlds. If my arguments are convincing then this means that one key argument against the existence of mental causation is overcome.

Acknowledgements

I want to start by saying a huge thank you to my supervisor Dr Luke Fenton-Glynn for his guidance and patience throughout the thesis writing process. His knowledge and expertise made writing my thesis much more enjoyable than it otherwise would have been. I'm looking forward to continuing my research with him in the future.

Furthermore, I'd like to thank all the members of the UCL Philosophy Department who I've worked with in the past few years. They all contributed to creating a great atmosphere for learning and developing ideas. Contributors to both WIP and thesis prep were invaluable for their feedback and discussion. In particular, I'd like to thank Dr Fiona Leigh for being an excellent Graduate Tutor, Thesis Prep Seminar Leader and teacher of *Aristotle*, as well as the members of my cohort. Catherine Dale, Helena Cicmil and Charles Jansen deserve particular thanks for being great philosophers and for making the department such a welcoming place to be. I likewise look forward very much to working with everyone in the future.

Thanks go also to the A. J. Ayer Scholarship fund and to the UCL Scholarship committee for providing me with financial support which allowed me to work under less financial pressure. Thanks also to the London School of Economics Library for providing me a job and being an excellent and accommodating employer for the past five years as well as to all the staff there for being excellent colleagues.

Finally, I must thank those people who, while not directly involved philosophically, were of invaluable support to me during this MPhil. The staff at Prufrock have been

very kind to me over the years while I nursed a coffee and wrote the vast majority of my work. Huge thanks go to Dr Meadhbh McIvor, Nina Lazic and Lisa Franklin for being there come what way and always understanding. I also owe thanks to Dr Nik Oikonomopoulos for his support through some stressful times. Finally, my biggest thanks must go to my family without whom none of this would have been possible. They have provided for me financially and emotionally throughout my many years of philosophy. I therefore want to dedicate this thesis to my favourite person, my sister, Phyllis.

CONTENTS

1	Introduction	11
1.1	Thesis	11
1.2	Motivation	15
1.3	Mental States	17
1.4	Free Will	18
1.5	Consciousness and Qualia	19
1.6	Plan	22
1.7	Summary	23
2	Goundwork	24
2.1	Physics and Physicalism	25
2.1.1	Determinism, Non-determinism, Probabilistic Indeterminism and Indeterminacy	26
2.1.2	Chance and Randomness	29
2.2	Events	31
2.2.1	Davidson	31

CONTENTS

2.2.2	Kim	33
2.3	Four Theories of Causation	34
2.3.1	Counterfactual Theory of Causation	36
2.3.2	Probability Raising Theory of Causation	41
2.3.3	Interventionist Theories of Causation	44
2.3.4	Process Theories of Causation	47
2.3.5	Mental Causation	50
3	Arguments for Mental Causation	51
3.1	The Manifest Image	51
3.2	The Evolutionary Argument	53
3.3	Inference to the Best Explanation	56
3.4	A Potential Problem	61
4	The Causal Exclusion Argument	63
4.1	Mental Causation and Multiple Realizability	67
4.1.1	Wrong Grain Objection	71
5	The Non-Identity	
	of the Mental and the Physical	75
5.1	The ‘Explanatory Gap’	78
5.2	Nagel and ‘What It’s Likeness’	79
5.3	The Knowledge Argument	80
5.4	Accepting the Premise	81

6	Causation and Overdetermination in a Probabilistic World	82
6.1	Denying The Problem Of Overdetermination	83
6.1.1	Sider	83
6.2	Probabilistic Overdetermination	85
6.2.1	Example of Overdetermined Probabilities	86
6.3	Overdetermination in a Probabilistic Setting	89
6.3.1	Sider Revisited	89
6.4	Chapter Conclusion	91
 7	 Causal Closure of Physics	 92
7.1	What is Causal Closure?	93
7.2	Causal Closure and Completeness of Physics	95
7.3	Arguments For Causal Closure	97
7.3.1	Physicalism	97
7.4	Arguments Against Causal Closure	98
7.4.1	Physicalist Bias	98
7.4.2	The Explanatory Gap and Knowledge Argument	101
7.5	Probabilistic Causal Closure	102
 8	 Evaluating the Probabilistic CEA	 103
8.1	Causal Closure and Indeterminism	103
8.2	Unresolved Issues	110
8.2.1	Departing from Physicalism	111

CONTENTS

8.2.2 Peculiar Conclusion? 111

8.2.3 Invalid or Unsound 112

9 Free Will 114

9.1 What is Free Will? 115

9.2 Free Will and the CEA 118

10 Conclusion 120

1

INTRODUCTION

“Begin at the beginning’, the King said gravely, ‘and go on till you come to the end: then stop’”

Alice in Wonderland - Lewis Carroll

(1856)

1.1 Thesis

My thesis argues that Jaegwon Kim’s Casual Exclusion Argument (CEA) against the existence of mental causation does not hold in an indeterministic setting and that therefore mental causation may exist in a probabilistic setting. That is the mental

may cause physical effects in a probabilistic setting and does so in virtue of its being mental¹. I will go into more detail about the relation between the mental and the physical in section 1.3 below. I will introduce the CEA very briefly now so that I may lay out the plan for the rest of my thesis more clearly, but will hold off going in to full detail until a later chapter. There are three premises to Kim's argument;

(P1) Causal Closure of Physics

Every physical effect has a sufficient physical cause.

(P2) No Systematic Overdetermination

It is not usually the case that there are multiple minimally sufficient causes of any given event which exist simultaneously.

(P3) Non-Identity of the Physical and the Mental

There are physical states and mental states and these are not identical to each other.

(C) There is no Mental Causation

Mental states cannot be causes.

In a sentence, if physics is closed then we have a sufficient causal picture of the world which means that, assuming the physical and the mental are not identical, and we reject widespread and philosophically troubling overdetermination, there is no mental causation. This leaves mental states as epiphenomenal or in some way illusionary in our mental lives because it will always be the underlying brain state which is actually bringing about effects despite feeling like our mental states are doing the causing. I

¹Whenever I write the existence or non-existence of mental causation I mean specifically mental causation of physical effects, but I will not always specify this.

will leave the argument here for now and lay out the rest of the introductory notes and the plan for the rest of my thesis.

My goal for this thesis is to show how one argument against the existence of mental causation is wrong in a probabilistic world (as I assume we live in). As the CEA does, at least on the face of it, seem to hold in deterministic worlds, I will also examine the probabilistic analogue of the CEA (which I set out below). That is, there may be a fairly straightforward parallel CEA which could be put forward in the probabilistic setting where physics fixes the probabilities of events occurring rather than the events themselves. I do not hope to argue that mental causation exists, but rather to show that the CEA is not a successful argument against it, at least as probabilistic worlds such as ours go. This far from sufficient for demonstrating the existence of mental causation.

Probabilistic Analogue CEA

(P1') Causal Closure of Physics

Every physical effect has a physical cause which is sufficient to fix its probability.

(P2') No Systematic Overdetermination of Probabilities

It is not usually the case that there are multiple sets of events that are minimally sufficient to fix the probability of a further event which exist simultaneously.

(P3') Non-Identity of the Physical and the Mental

There are physical states and mental states and these are not identical to each other.

(C') There is no Mental Causation

Mental states cannot be causes.

If mental causation exists in a non-deterministic setting could this raise questions about the truth of physicalism? Of course the mental could still supervene on the physical, but it may not be guaranteed that the relationship between the mental and the physical is exactly the same in both the deterministic and probabilistic setting. If mental states exist in probabilistic settings, but turn out to not supervene on the physical in some way, that could potentially violate sacred philosophical intuitions about physicalism. It seems unlikely that such a scenario should be the case. I will assume that there is some close relation between the mental and the physical in any setting.

The plan for the rest of this introductory chapter is to put forward my motivation for tackling these difficult issues and why we should find them philosophically interesting. I will explain what I mean by a mental state (but will leave explaining what I mean by mental causation until I have discussed some popular theories of causation). Having done this I will quickly note that I will not be interested at this time in debates about free will. I will briefly discuss consciousness, another topic I do not want to directly tackle but which is heavily connected to the work I will be doing. I will then lay out the plan for the rest of the thesis.

1.2 Motivation

The reason I think this is a question worth attempting to answer is because it cuts to the heart of our everyday experience and to the heart of philosophical debates which have raged for centuries. It relates to questions which have been asked since philosophy began; and particularly since Descartes. It would be extremely strange if our most basic experiences of the world should prove to be illusionary. This is because I take it to be uncontroversial (and indeed integral to our mental lives) that we all feel as though our mental states are causing us to act. For example, when we're hungry we eat, when our heads hurt we reach for the painkillers. However, and perhaps surprisingly, it turns out to be hard to defend our intuitive sense of the world philosophically. If the CEA is correct then it is not our mental states which cause us to act, it is purely the result of the underlying brain states. Just because this would be extremely counterintuitive does not mean that it cannot be the case, but it does lead me to think that if there is a way to preserve our everyday fundamental experiences which is also philosophically satisfying, then that picture is to be preferred.

To be philosophically satisfying a theory must be consistent with our best current scientific theories. Newtonian physics, while still utilised for some practical purposes, is agreed to be fundamentally wrong. The two front runners of our best current partial scientific theories are now quantum mechanics (though there are many interpretations of quantum mechanics each with different physical implications) at the microscopic level, and General Relativity at the macrolevel. There is currently one largely ortho-

dox philosophical way of metaphysically conceptualising the world: physicalism². Orthodox interpretations of quantum mechanics are fundamentally indeterministic theories of the world. Therefore, as a front runner in our best scientific theories, it seems philosophically astute to think of our world as being indeterministic. I want to use this background assumption to assess the existence mental causation against the CEA. Is the argument still valid in a probabilistic world? Or do one or more of the premises no longer stand up to scrutiny? I will argue that it is not sound in its original form in the probabilistic setting but that there may be an analogue which is. I will ultimately conclude that even this analogue CEA is unsound in probabilistic worlds like ours. To try to make my point clearer I have set out a table below which summarises the validity and soundness of the CEA and a potential analogous CEA in both a deterministic and probabilistic setting.

<i>Deterministic World</i>	<i>Probabilistic World</i>
CEA	CEA
√ Valid √ Sound	√ Valid X False Premises
Analogous CEA	Analogous CEA
? Valid X Unsound	? ?

The table shows that in a deterministic world the CEA is both valid and presumably sound (because causal closure does potentially hold in a deterministic world). However, when analysed in a probabilistic world, I argue it would become unsound (because

²I will define exactly what I mean by this in the next chapter.

closure would no longer hold) and therefore unsound.

1.3 Mental States

I will take the existence of mental states for granted for the purposes of my thesis. There are of course some arguments against their existence³ but I will not discuss these. What do I mean when I use the term 'mental state', as I shall be doing frequently in the course of this thesis? I won't attempt an exhaustive and exclusive definition of mental states⁴ There are many different kinds of mental state; beliefs, desires, emotions, dispositions, attitudes, and of course pain to name a few. At least some are conscious at least some of the time, or are conscious in some aspect. There are two paradigm examples of mental states: qualia and propositional attitudes. Qualia are usually defined as 'what it's like-ness', for example, "What it's Like to be a Bat" Nagel (1974). It is the part of mental life which is essentially experiential. Pain and 'seeing the colour red' are classic examples although other feelings, emotions and perceptual experiences all have a qualia aspect to them. Propositional attitudes are the mental states agents hold which have propositions or states of affairs as their content. Beliefs and desires such as the belief there is ice cream in the freezer and the desire to eat ice cream, are examples of propositional attitudes.

What is the relation between the mental and the physical? What does it mean to say a

³For example, there have been eliminativist arguments put forward by Patricia and Paul Churchland (1981).

⁴As I will go on to say, I don't want to get into discussion about which things do and do not have mental states or consciousness but I take it that, contra Descartes, non-human animals do have them.

mental state has an underlying brain state? The most popular view is that the mental supervenes on the physical. This means that the mental is determined, in a stronger than causal manner, by physical brain states. An explanation of this is put forward by Lewis (1986*b*) with his dot-matrix example.

“A dot-matrix picture has global properties - it is symmetrical, it is cluttered, and whatnot - and yet all there is to the picture is dots and non-dots at each point of the matrix. The global properties are nothing but patterns in the dots. They supervene: no two pictures could differ in their global properties without differing, somewhere, in whether there is or isn't a dot” p.14.

Sometimes supervenience is applied to the special sciences and physics. So, in much the same way that the mental is said to supervene on the physical, it could be said that biology or chemistry supervene on physics (though this is a very hotly debated issue).

1.4 Free Will

A second important note to make from the outset concerns the debate around free will and how it connects with my work. I raise this substantial issue here for two reasons. First, because the philosophical literature about determinism in particular is usually couched in talk about free will. It therefore may be expected that I would engage with this literature. However, although what I have to say will be relevant to the topic of free will, I do not have the space to do it full justice here. That said, I will devote a chapter to it after I have made my main argument. It remains an avenue I hope to

explore in the future. For now I focus exclusively on mental causation.

Second, is to mention that to have free will plausibly requires of a person that they have causally efficacious mental states. Most theories of free will take it to be necessary that our actions are caused by our beliefs, desires and decisions. So what I have to say in my thesis will bear on the free will debate in that the CEA stands as a possible counter argument to the existence of free will, even if you take a compatibilist stance. This is because the CEA argues against the existence of mental causation which is taken to be so central to agents making free choices. So, if I successfully argue that the CEA does not hold this will benefit the defender of free will (at least in probabilistic worlds).

So to recap, I mention free will here only to note its relevance but to further note that I will not mention it again until the end of my thesis. Doubtless there are interesting interconnections between the two phenomenon, but this is a topic to discuss more fully at a later time.

1.5 Consciousness and Qualia

The last important opening note I want to make relates to consciousness. What it means to be conscious, what things count as having consciousness and exactly how consciousness relates to mental causation are all interesting questions but tangential to the task at hand. I want to make it clear that my main focus is mental states. Of course, at least some mental states are conscious or can easily be brought to attention. There are some actions which are so run of the mill that you do them without thinking (take

typing as an example) which can none the less be brought to attention without difficulty. There may be other deeper unconscious mental states which are harder to bring to conscious attention. That they can be causally efficacious is unclear but is suggested by some psychological treatment which is based on the premise that sometimes what is required to correct unhelpful behaviour patterns, is to examine what may unknowingly be at the root of them. Take cognitive behavioural therapy as an example. The idea of this kind of treatment is to change mood and behaviour by changing thinking patterns. Part of changing an individual's thinking pattern is to make that person realise they may have beliefs, (often long) standing beliefs, about themselves and their abilities which they did not even realise they had.⁵ A distinction sometimes made between such conscious and unconscious mental states is termed as the difference between 'standing' or 'dispositional' and 'occurrent' mental states (see Schwitzgebel (2015)).

I'll use an example, which I will call 'my sister's birthday' example, to illustrate what the difference between a standing and an occurrent mental state is and how it relates to my work. You can know something as a 'standing' mental state (I would be said to know the date of my sister's birthday even when I am not currently thinking about it,) but it's not until it's brought to attention as an 'occurrent' mental state (that I'm currently experiencing) that you take action based on it. I have known for years that my

⁵To give a concrete example; a patient may be anxious because they are not performing well at work. When questioned why they are not performing well they may give a list of reasons such as being too tired or being distracted. If that patient was to correct for all these problems they may still find that they are not happy with their performance. If it is suggested to them that it could be because they hold a belief such as 'I can't do it' they may even disagree. However, by bringing to attention such a belief and countering *that*, they may find an improvement in their behaviour and mood. The improvement in outcome would suggest that it was in fact the unconscious belief which was the cause of the problems in the first place. This would suggest causation by an unconscious mental state, although of course my example as it stands is hypothetical and to become convincing would need empirical evidence to back it up.

younger sister's birthday is in September. However, unless I bring that fact to mind⁶ I will not act on it and will neither buy a present nor wish her happy birthday. When I remember (or more likely, on being reminded) that I've missed it, this will induce guilt and actions such as panic buying and apologising⁷

What I hope to suggest with this example and through use of such terminology, is that possibly in some cases, mental states may only be causally efficacious *because* they're mental. This is important to my argument because one obvious retort a defender of the CEA could make is that mental states could just be physical. In that case there's no violation of causal closure. Qualia on the other hand are not physical (See Chapter 3 for more on this) and consciousness necessarily involves qualia. This is why my example stresses the occurrent/standing belief distinction. However you view standing beliefs, I take it that occurrent beliefs involve qualia. All I need to show to make my case is that there are at least some cases where a mental state causes a physical effect in virtue of being mental. Therefore, if mental states are causally efficacious, but only insofar as they are conscious in some aspect, then the violation of causal closure still stands.

In conclusion, the CEA is problematic not because it denies the possibility of mental causation per se. Mental states causally effecting other mental states do not violate the CEA. What does violate the CEA is mental states bringing about physical effects which is just the kind of causation I am interested in. It doesn't matter to my argument whether there is mental-to-mental causation, or if there are such things as unconscious

⁶And assuming I also have the desire to make my sister happy, the belief that remembering her birthday will make her happy and so on.

⁷Someone could question which belief exactly is it that is causing the action. Is it the belief that her birthday is on x, or the belief that that date has passed, and so on? I'm not sure the exact belief is important to my example as long as it causes my later actions.

mental states, so long as it is clear that there are some mental states which are conscious and that these can act causally in the physical world in virtue of their being mental (such as by being conscious which is necessarily non-physical) in some aspect. This is sufficient to demonstrate that there can be mental causation which is distinct from physics and thus the CEA is vulnerable.

1.6 Plan

The broad outline of my thesis will run as follows. My first main chapter will lay the groundwork. I will introduce the physics which leads us to conclude the world is indeterministic as well as what it means to be indeterministic. I will outline four theories of causation to help spell out what it is I mean by the term ‘mental causation’. I will spend Chapter 3 laying out my reasons for thinking that mental causation exists; the argument from the Mental Manifest Image, the argument from evolution and inference to the best explanation. Having introduced all the key parts, I will move on in the next four chapters to lay out the CEA as well as its probabilistic analogue and analyse its three premises in more detail. So, in Chapter 5 I discuss the non-identity of the mental and the physical. I will not have much to say on this premise if only because I want to focus on looking at the argument through a probabilistic lens and this makes most difference the other two premises. In Chapter 6 I examine the second premise of the CEA; that there is no systematic overdetermination. I go through an argument as to why overdetermination should not be philosophically worrying and then translate this into probabilistic terms to see if it still holds good. Chapter 7

discusses the first premise of the CEA; the causal closure of physics. Lastly in Chapter 8 I will put forward my reason for thinking causal closure doesn't hold in probabilistic worlds⁸. I will conclude in Chapter 9, by arguing that while the deterministic version may hold in a deterministic world, the probabilistic version I put forward, does not hold in probabilistic worlds. Therefore, it cannot stand as a counter argument to the existence of mental causation in probabilistic worlds such as ours.

1.7 Summary

To conclude this section I have introduced my hypothesis and how I plan to go about discussing it. I have noted that I will put the topic of free will to one side and have touched on how the topic of consciousness intersects with what I want to say.

Before going any further it is now essential that I explain what exactly the CEA is and what I mean by such key terms as 'determinism', 'probabilistic' and 'mental causation'. This is what I shall now do in chapters 2 and 3 before examining the premises individually and putting forward my argument for getting around the CEA.

⁸An interesting question would be does the deterministic version of the CEA work as a special case of the probabilistic version of the CEA? If so, I would have to conclude that the CEA only goes through in this special case, but that no actual examples of the special case occur.

2

GOUNDWORK

“

”Do I dare to eat a peach?

The Love Song of J. Alfred Prufrock -

Eliot (2010)

While TS Elliot’s protagonist ponders, paralysed by indecision, whether he dare eat a peach, perhaps he should have asked himself ‘*can* I dare to eat a peach’? Do we have any power over our choices; are our mental states ever causally efficacious? My argument will be that the CEA potentially works only on a deterministic¹ picture of the world, by virtue of the fact that its first premise, relies on determinism to hold. Orthodox quantum mechanical pictures of the world are fundamentally and

¹For example a Newtonian picture of the world, although interestingly Newtonian physics may not actually be deterministic, see Norton (2008).

in principle probabilistic. Therefore, supposing that we do in fact inhabit some kind of quantum mechanical world, and at this point that is what our best theories and evidence suggest, what consequences does this have for the CEA?

Before beginning substantive philosophical work, groundwork needs to be laid by way of setting out clearly the views and terms I will be using. First, I introduce physicalism and then define key terms such as determinism, chance and randomness. I then move on to discuss orthodox interpretations of Quantum Mechanics². Lastly I will introduce four of the most commonly held philosophical theories of causation.

2.1 Physics and Physicalism

Newtonian physics, while still utilised for some practical purposes, is agreed to be fundamentally wrong when it comes to describing how we think the world really is. The two front runners of our best current partial scientific theories are now quantum mechanics at the microscopic level, and General Relativity at the macrolevel. There is currently one largely orthodox, though heavily debated, philosophical way of conceptualising the world: physicalism.

Orthodox quantum mechanics is a fundamentally indeterministic theory of the world³. I will explain further exactly what I mean by these terms after introducing what I mean by physicalism. As there are many different definitions of physicalism it will be

²I will not be discussing non-standard interpretations as they are not always probabilistic. See for example the 'Many Minds' interpretation (Albert & Loewer (1988)).

³I will briefly discuss Superposition and Heisenberg's Uncertainty Principle in order to explain why this is the case below.

important to know exactly what is meant by this term as it's going to be so central to my work. I will define it as follows:

(Phys) Everything supervenes on the physical.

In *Mind in a Physical World* (1998) Kim defines mind-body supervenience, as follows:

“Mental properties supervene on physical properties in the sense that if something instantiates any mental property M at t, there is a physical base property P such that the thing has P at t, and necessarily anything with P at a time has M at that time.” (Kim (1998) p.39)

This definition is minimal he claims, because he takes it to be the minimal commitment to which every type of physicalist must agree. In this case the relation between the mental and the physical is supervenience, so there is no possible change in the mental without some change in the physical. For Kim therefore, physicalism only holds if mind-body supervenience holds.

2.1.1 Determinism, Non-determinism, Probabilistic Indeterminism and Indeterminacy

Before going any further I want to avoid potential conceptual confusion which could arise from terminological similarity. I will set out here how I use the terms determinism, non-determinism, indeterminism and indeterminacy for the purposes of my thesis. I'll start by defining determinism (D) as;

- (D) Every actual event is necessitated by the initial conditions of the world given the complete set of the laws of physics.

Non-Determinism (ND) as I shall use the term is simply a denial of determinism. Thus;

- (ND) Determinism is false.

However, there are many different ways a world could be non-deterministic. For example it could be probabilistic, random, utterly lawless or a blend thereof, (I disambiguate the concepts of chance and randomness below). I will therefore be using ND as something of a catch all term. As noted above, orthodox interpretations of quantum mechanics lead us to believe that our world is not strictly deterministic and is therefore ND. But, it is important to note that an orthodox quantum mechanical world is ND in a particular way, namely it is probabilistic. As this is the way, to the best of our current knowledge, that we think the world actually is, this is the particular variety of ND which I will be interested in.

Probabilistic indeterminism (ID) then is the thesis that;

- (ID) The world is fundamentally probabilistic. Given an initial set of conditions and a completed set of the laws of physics a non-trivial probability distribution⁴ could be derived over possible future events.

⁴In other words one which does not just return 1s and 0s as would be the case on a deterministic picture.

This is not to say that probabilities cannot alter over time; there will be some events whose unconditional probabilities differ from their probability conditional upon certain possible initial histories of the world.

So far I've been disambiguating different shades of the same theme. Indeterminacy (IDy) on the other hand is a quite different concept. Although it's unrelated I mention it here to disambiguate quantum indeterminism from quantum indeterminacy. It is the thesis that;

(IDy) There is no fact of the matter.

An illustration of this comes in *The Uncertainty Principle* (Heisenberg (1927)). Superposition means that a particle has no definite state until it is measured. The Uncertainty Principle states that once the position of a particle is known its velocity becomes indeterminate, that is, there is no fact of the matter about the precise velocity of the particle. Rather than being a failure of our observational abilities, this has to do with the indeterminacy of the velocity of the particle in this case. However, if the velocity of the particle was then to be measured there would be a *probabilistic* fact of the matter of its velocity (and an indeterminate one of its position). So when an observer measures a particle's velocity, thereby collapsing the wave function and causing it to take a determinate value, she causes its position to fall into superposition - in which it has no determinate location. If she then measures the position of the particle, causing it to have a determinate location, she causes the particle's speed to fall into superposition, where recall, there is no fact of the matter. The result of the collapse of the wave

function is probabilistic. Because of this, and given that there is no way out of this cycle, the quantum mechanical world is in principle probabilistic and indeterminate.

2.1.2 Chance and Randomness

Often in everyday speech the terms ‘chance’ and ‘randomness’ are used interchangeably. These are philosophically distinct concepts however. Following Anthony Eagle’s *Chance versus Randomness* (Eagle 2016) I will make the distinction between process and product. Chance lies in process whereas randomness applies to products of processes. Further, predictability will be a factor in distinguishing the two. Chances are probabilistic, whereas truly random events are strictly unpredictable⁵.

Compare the process of flipping an unbiased coin to the product of a random sampling. The outcome of the coin toss is 50/50 coming up heads or tails. As I’ve been stressing, processes at the quantum level are intrinsically probabilistic or chancy. Random samples on the other hand need not be the result of chancy processes. Eagle gives the example of using a simple heuristic to collect demographic details; that is, take the personal details of babies born at any time ending in seven. Because (we assume) there is no connection between the exact time of peoples’ birth and their other pertinent demographic details, this simple process can produce random results. There will not be any pattern within the data collected by which you could predict the results. Let’s return once more to the coin toss. While you would know for each toss the probability

⁵Indeed one definition of a random sequence is that you would need an algorithm at least as long as the list itself in order to specify the sequence (Eagle (2016)). This is because there is no pattern whatsoever by which the algorithm could compress the sequence. Therefore it must just give a brute list.

of it coming up heads and tails (i.e. 0.5), the actual sequence which results will be random. This is extremely brief but should be enough to highlight the differences between chance and randomness.

So why is this distinction important for my work? The kind of probabilistic world I am interested in for the purposes of this thesis are chancy but not merely random. That is, I'm not interested in worlds where merely the outcomes are chancy, but ones in which the causal processes that produce them are themselves chancy. In a chancy world it could still be that the probabilities of future events may be fixed by the set of initial conditions and previous events. In this case the CEA as set out by Kim would still potentially hold. To be a probabilistic world in the sense I am interested in is to be chancy in this way.

Why am I not interested in truly random worlds? These are worlds which are either completely lawless or worlds in which causal processes are indeterministic but not probabilistic. A world could be deterministic and merely look by its product to be random; what Carl Hoefer (2016) calls "deterministic chaos". There could also be truly random worlds where both the outcome and the process is random. As I've said I don't want to discuss random worlds principally because current orthodox theories of quantum theory suggest our world is not like this and I want to focus on the closest picture we can get to of our world. However, more than this, in a truly random world it might be hard to see how our mental states could be bringing about effects in any meaningful way given that no leading theory of causation accommodates indeterministic but not probabilistic causation.

2.2 Events

Before going into any theories of causation there are two important questions which need answering⁶. What are the relata of causation and what are events? These questions are of course linked if you take the relata of causation to be events as I do. It's important to be clear, because what an event is will bear on which events we can genuinely take to be the relata of mental causation. A Davidsonian event for example, may be too coarse grained to meet the criteria for true mental causation. To be a case of genuine mental causation it is not enough for an event with a mental property to do the causing. It must be that the event causes the effect in virtue of its mental property⁷. I will now lay out a Davidsonian and a Kimean view of events.

2.2.1 Davidson

Davidson's current theory of events is coarse grained. While he used to identify events by their cause and effects he later individuated events by their spatio-temporal location⁸. So two events are identical if they share the same space-time. For Davidson 'the stabbing of Caesar' and 'the killing of Caesar' are one and the same event under different descriptions.

In *Mental Events* Davidson (2001*b*) combines three principles which he takes to be true.

⁶I follow Casati & Varzi (2015) and Schneider (2017) in this section

⁷Furthermore, as I've mentioned, to be the type of mental causation that violates the CEA, it must be a mental cause of a physical effect.

⁸See Davidson (2001*b*)

They are (1) The Principle of Causal Interaction, (2) The Principle of the Nomological Character of Causality and (3) The Anomalism of the Mental. The first principle states that “at least some mental events interact causally with physical events” Davidson (2001*b*) p.208. The second principle states that “where there is causality, there is a law” Davidson (2001*b*) p.208. He takes the first two principles to be assumptions. The third principle states that “there are no strict deterministic laws on the basis of which mental events can be predicated and explained” Davidson (2001*b*) p.208. On the face of it these principles are inconsistent. If its assumed that mental and physical events causally interact, and that with causality comes laws, then how can there be no strict psycho-physical laws.

He dissolves this inconsistency by arguing that the physical and mental events are identical to each other and as such share spatio-temporal location. So for example, pain is identical to the firing of C-fibres. Therefore, every mental event is also a physical event which does have laws associated. To be the kind of cause I am interested in then, it must be in virtue of the instantiation of a mental property that any event (individuated by its spatio-temporal location) brings about an effect.

There is a counterexample to Davidson’s view provided by Davidson himself in response to Lemmon (1996)⁹. The example is designed to show that there can plausibly be two events which share the same spacial-temporal location. Imagine a metal ball which is being heated up at the same time as rotating. Stipulating that the warming is happening over and above that caused by the rotating, intuition suggests that there are two separate events occurring. That is, the event of heating and the event of rotating,

⁹The counterexample can be found in Davidson (2001*a*) p.178

rather than one event of heating-rotating. But, given the molecules which are being heated and the molecules which are being rotated share the same space-time location, it seems unclear why this is not just one Davidsonian event.

If you hold a Davidsonian view of events then, a mental event will always also be a physical event. This will not suffice for my purposes because any occurrence of mental causation will also be an occurrence of physical causation and to be genuine mental causation the cause must be mental.

2.2.2 Kim

Kim (1993) has a much more fine grained conception of events. He views them as a triple $\langle P, o, t \rangle$. That is, the instantiation of a property P by an object o at a time t. I will be following Kim's conception of events¹⁰. It is more fine grained as it individuates events by the property they instantiate and therefore the property of 'stabbing Caesar' and the property of 'killing Caesar' produce different events. This is because the property of being a stabbing and the property of being a killing are different properties. Not all stabbings are fatal and there are many more ways of killing people than by stabbing. Kim argues that despite the fact that in Caesar's case the stabbing did amount to a killing, they are still not the same event. He presses this by making the point that "to explain Brutus' killing Caesar (why Brutus killed Caesar) is not the same as to explain Brutus' stabbing Caesar (why Brutus stabbed Caesar)" Kim (1966) p.232 (footnote).

¹⁰Some people prefer to call such triples facts or states of affairs (Casati and Varzi discuss this in Casati & Varzi (2015)) in which case my view can be modified to take these as the relata of causation.

A Kimean view of events allows mental events and physical events to be different because it can instantiate either a mental or a physical property. This allows the kind of mental causation required to violate the CEA.

So in summary, when I talk about events bringing about effects I specifically mean a mental event bringing about a physical event in virtue of its mental property or because it is a genuine mental event. It doesn't make any difference for my purposes which theory of events is endorsed. As long as the cause is either a mental event in the Kimean sense, or is a Davidsonian event which is only efficacious in virtue of its mental property, then this suffices for my argument. I will use both of these locutions throughout but anytime I do, the other is substitutable. With these clarifications made I will move onto introducing the four most popular theories of causation.

2.3 Four Theories of Causation

Other than 'not correlation', what is causation? As this is a thesis on mental causation, I must first make it clear what it is I mean by mental causation. To this end I will also discuss a few theories of causation to make clear which theories I have in mind for the purposes of this work. Of course, there are far too many theories of causation to cover here¹¹, in what is really a quick overview rather than a detailed examination. Therefore, I shall only cover four sets of theories; counterfactual, probability raising, interventionist, and process theories. I picked these based on their popularity within

¹¹For example, I will not be talking at any point about regularity theories of causation and Hume will only be mentioned here. See (Hume (2009) and Hume (1988)).

the field and their applicability to chancy or probabilistic causation. Probability raising accounts for example, seem particularly apt for discussing probabilistic causation. These theories are often interrelated to each other in various ways but distinguishing them clearly will not be necessary for my purposes. I wish to remain as neutral as possible in regards to causal theories so that what I have to say can have the widest possible appeal.

The thread which holds these (except perhaps process theories), seems to be the idea of difference making. Often the causal relata in question are events, although they can be other things, for example, facts¹². So at a first approximation here is what I mean by mental causation. An instance of mental causation occurs when, for any event M which is mental in nature, and any event E (which can be either mental or physical), then M is the cause of E. It is important to note however, that not only must M bring about the effect, it must do so in virtue of its mental quality to be true mental causation.

Although the CEA only targets mental causation of physical effects so this will be my main focus. In the most general terms, it is the mental event (or property of the event) which makes the difference (as opposed to some underlying physical event). It was my thirst which made me reach for my ginger beer and not only the underlying neural firing. The mental and not only the brain state was the thing which brought about my drinking.

¹²Although of course this is still a hotly debated issue. See for example Mellor (1995)

2.3.1 Counterfactual Theory of Causation

Perhaps the first kind of theory which springs to mind when causation is mentioned is the counterfactual theory of causation, and particularly that of David Lewis (1973) and Lewis (1986a)¹³. Counterfactual theories including Lewis' make use of possible world semantics in order to assess the truth or falsity of any given counterfactual statement. Lewis does this through the idea of cross world similarity. So, for any counterfactual statement, if A had not occurred then B would not have occurred, if and only if some non-A world in which B does not occur is closer to the actual world than any non-A world in which B occurs.

The gist of these theories is that for any two events A and B, A causes B if and only if had A not occurred then B would not have occurred. More fully it would be correct to say, A causes B if and only if, had A not occurred then B would not have occurred, and, if A occurs, then B occurs¹⁴. For Lewis though, A and B need only to be connected by a chain of counterfactual dependence¹⁵. However, for the purposes of my thesis it would perhaps be better to put translate this into talk of physical effects counterfactually depending on mental events (or events in virtue of their mental property). So event e is brought about in virtue of mental property M if and only if had e not instantiated M (or there had been no chain of counterfactual dependence from M to e) then e wouldn't have occurred.

¹³I follow Menzies (2014) in my exposition.

¹⁴In the instance that A and B are both actual events the second of these counterfactuals is automatically true. This is because the closest A world is the actual world and it is also a B world

¹⁵Lewis (1973)

For example, for the mental event T (thirst) and the event G (of reaching for my drink), had T not occurred then G would not have occurred and if T occurred then G occurred. Looking to the worlds in which T did not occur, the closest world to the actual world is one in which G does not occur. Therefore, it is true to say that had T not occurred then G would not occur and that therefore T is the cause of G. This would be a case of counterfactual causation.

There are some problems with counterfactual theories however, that is the problems of early and late preemption in which some cases fail to count causes as causes. Early preemption cases involve scenarios where a potential cause is preempted by an earlier cause. Assassin cases are typically used to illustrate this. Lewis can get around this problem by invoking his idea of casual chains¹⁶.

On the other hand late preemption cannot be solved by using the idea of causal chains of dependence. Late preemption involves cases where a potential cause cannot be the actual cause because an alternative has already brought the event about. This is different from early preemption because in late preemption the preempted cause is cut short after the event has been brought about rather than before. The classic case from Ned Hall (2004) is that of Suzy, Billy and a bottle. Billy and Suzy both throw a rock at an innocent by-standing bottle. Because Suzy throws first, her rock hits the bottle and smashes it before Billy's can. However, as Billy had also thrown his rock, his throw would have been the cause of the bottle's smashing had Suzy thought better of throwing hers. Thus there is no causal dependence from Suzy's throw to the bottle's smashing (because the bottle's smashing does not counterfactually depend on Suzy's

¹⁶See Lewis (1973).

throwing) despite the fact that her throw is the actual cause of the bottle's smashing. This is a case of causation without counterfactual dependence.

A possible reply to the problematic nature of late preemption cases is that events are modally fragile and therefore the smashing which would have occurred had Billy's throw been successful would have been a different smashing from the one which actually occurred after Suzy's throw. Lewis does not think this is the correct way to analyse the situation (See Postscript E *Redundant Causation* in Lewis (1986a))¹⁷, however it is the one which intuitively makes sense to me. In his *Postscript to Causation* Lewis (1986a) gives the example of death by poison. The poison in question kills its victim much more quickly if taken on an empty stomach. When taken after food the poison is much more slow acting and painful. So, is it fair to say that the two deaths are different versions of the same event (in which case whether or not the victim ate beforehand is largely irrelevant) or different events? If the latter, then it would be true to say that the victim's eating dinner before ingesting the poison was actually part of the cause of the death, as it lead to this specific death as opposed to a much quicker and more painless one. It seems counterintuitive to say that the victims eating dinner was (at least part of) the cause of their death.

However, I think this is problematic only insofar as one's intuition allows. I personally do not find this way of analysing events to be that counterintuitive. There is perhaps a reason why someone may have the opposite intuition to me; that is, the intuition that the dinner is not a cause of the death. Hitchcock & Knobe (2009) analysed some

¹⁷Although in (Lewis (2000)) he introduced the idea of alterations of events into his theory which allows for alterations of events to be more modally fragile.

experimental philosophy¹⁸ which suggests that subjects tend to conflate cause and attribution of moral responsibility. So when someone is deemed to have acted in a morally unacceptable way¹⁹, their action is more likely to be deemed the cause and is likely to be attributed more causal weight. This makes sense given that, often, violation of a norm (particularly moral but also statistical) will lead to a negative outcome. To return to the poisoning case, this could help explain the widely held intuition that the poison and not the dinner is that cause. The poisoning is a violation of a norm (which does indeed lead to a negative outcome) whereas eating dinner is not. Therefore it intuitively seems that it is the poisoning which is the cause and the eating of the dinner is not.

Perhaps, on an everyday bases, a conflation between token and type causation occurs. Again this would make sense given that on a day to day basis, we have to make generalisations about causation. This links to Hitchcock and Knobe's claim that we attribute cause to that which we can most easily intervene on (See Hitchcock & Knobe (2009) pgs 606-607). So, in the dinner case, although the dinner may be a token cause, generally dinner is not the type of thing which brings about death. Poison on the other hand is both a token cause in this case and a type of thing which causes death. We must therefore be careful around poison in a way we don't generally have to be careful around dinner, so we attribute the cause to the poison. The poison is the thing we

¹⁸Subjects in the Knobe & Fraser 2008 experiment were given different scenarios where two people's actions were required to bring about the effect but only one of whom's actions violated a norm in some way. There was a statistical difference in the attribution of causation with the agent violating the norm being considered the cause much more frequently than the morally 'neutral' agent.

¹⁹Or, more strictly speaking, when they violate either a "prescriptive norm" or "statistical norm" Hitchcock & Knobe (2009) p.597.

could most easily intervene on and thereby prevent the death from occurring²⁰.

It is therefore my opinion that being able to handle problematic counterexamples outweighs any counterintuitive pull there may be. In a sentence, I do not think it's fair to say more causation must mean spurious causation. If thinking of events as modally fragile goes against our everyday way of treating events, but allows us to deal with the problem of late preemption, then so much the worse for our everyday ways of talking. Much more needs to be said about this before my argument can become convincing, but, as this section is designed to be merely introductory sadly I must leave this topic here and move on.

How does counterfactual causation work on a probabilistic interpretation? In his *Postscript to 'Causation'* Lewis (1986) discusses just this. Roughly, for any events A and B, A caused B if and only if, had A not occurred, the probability of B occurring would be lower, and if A did occur then the probability of B occurring is higher. This formulation is a combination of counterfactual and probability raising accounts which is why I shall go on to introduce probability raising accounts next.

There is a class of counterexample which is problematic for probabilistic counterfactuals provided by Menzies (1989). Menzies' example involves two systems which produce the same effect. He talks about two systems of neurons which I will label system A and

²⁰There is a similar issue with hastener/delayer intuitions. There is an asymmetry between which are attributed as causes whereby hasteners generally are and delayers generally aren't. Bennett (1987) gives the example of heavy rains delaying the forest fire from May till June. The heavy rains, like the dinner, is a token cause of this particular fire, but not in general a cause of fires, and similarly is not usually attributed as a cause in peoples intuitions despite the fact that it is a cause in this case. If we should, as I argue, look past the intuition that non-violations of norms should be considered a cause, just as violations should, then delayers should also be considered causes as hasteners are.

system B. System A is more reliable than system B however. On one occasion, system A switches off and turns system B, the less reliable system, on. On this occasion system B cooperates and produces the effect. So, despite lowering the chances of bringing about the effect, system B is a cause of the events coming about. It is important to note that, although I can't go into them here, there are many arguments put forward²¹ in response to these problems with counterfactual theories.

So mental causation on this interpretation amounts to a counterfactual causal dependence from a mental event to another event of the form 'had M not occurred, E would not have occurred (or the probability of E's occurring would have been lower) and if M had occurred then E would have occurred'²². To put this into 'mental terms' again; had event e not instantiated mental property M then further event e* would not have occurred (or the probability of its occurring would have been lower) and if e had instantiated M then e* would have occurred.

2.3.2 Probability Raising Theory of Causation

As this thesis is not only about causation, but specifically probabilistic causation, it would make sense to discuss probability raising theories of causation²³. The idea is that some event a is the cause of further event b if and only if a's occurring raises the probability of b's occurring. To put this more formally a is a cause of b iff $P(b | a) > P(b | \neg a)$. In prose, a causes a if and only if the probability of b is higher given a occurs

²¹See for examples, Menzies (1989), Hitchcock (2001) and Fenton Glynn (2016).

²²Note that if M and E are both actual events, then on Lewis' semantics the counterfactual "if M had occurred, E would have occurred" is automatically true.

²³The following exposition is influenced by Hitchcock (2012).

than the probability that b occurs given a not occurring.

More accurately for my purposes, event e which instantiates mental property M brings about further event e^* if and only if event e raises the probability of e^* 's occurring in virtue of instantiating M . The inequality in this case is as follows; e is a cause of e^* in virtue of instantiating M iff $P(e^* | e \ \& \ M) > P(e^* | e \ \& \ \neg M)$.

There is a problem for accounts of causation which involves two events being correlated due to sharing a common cause. This can then produce the illusion of causation where none actually exists. The famous example of this, originally given by Hans Reichenbach (1956), involves atmospheric pressure, a barometer and a storm. The atmospheric pressure is a common cause both of the barometer reading changing and of the storm occurring. Because the barometer reading changes just before the storm hits, it may be thought that the reading on the barometer changing is a cause of the storm, although obviously this is not the case. Say A is the proposition that the atmospheric pressure is changing, B is the proposition that the barometer reading changes and S is the proposition that the storm occurs. Then $P(B | A) > P(B | \neg A)$ and $P(S | A) > P(S | \neg A)$. But the inequality $P(S | B) > P(S | \neg B)$ also holds. That is, the probability of a storm occurring, given the change in barometer reading, is higher than it would be given no change in the reading. This is obviously because there is no change in barometer reading without the atmospheric change which also actually causes the storm, but there's nothing in this inequality to tell us that.

Reichenbach (1956) however, came up with a solution to this problem; screening off. When two events are screened off by an earlier event, they are not causally related. The

idea is to hold other conditions fixed when assessing the inequalities in order to try to isolate the specific thing which is actually doing the causing. What does it mean to hold something fixed? It means including the fixed variable among the things being conditioned on. So, say we hold atmospheric pressure fixed. Does the inequality $P(S | A \& B) > P(S | A \& \neg B)$ still hold? No, it will not because the atmospheric pressure will cause the storm whether or not the barometer reading changes. However, holding the barometer reading fixed, the inequality $P(S | A \& B) > P(S | \neg A \& B)$ will hold. So the storm would be conditioned not just on the atmospheric pressure but also on the barometer reading. Using this method (and so long as the cause comes earlier than the two screened off events), it becomes obvious that the atmospheric pressure is what is actually causing the storm. It also puts us in a position to be able to tell that the storm and the barometer reading share a common cause. That probabilistic accounts of causation can deal with these kinds of cases is one reason why these kind of accounts are to be preferred over other kinds of accounts, such as regularity theories.

How do probability raising accounts of causation work when considering probabilistic causation? As they deal in probabilities, the application is obvious. This makes them a natural way to talk about causation for the purposes of this thesis. In my opinion probability raising accounts also have the advantage of being very intuitively satisfying.

There are some counterexamples to probability raising accounts. For example there is a class of counterexample in which a cause lowers the probability of the event's coming about (See for example Hitchcock (2004)). The two systems case given in the probabilistic counterfactual cause section above would be an example of such a case. There are also cases of probability raising non-causation. Due to space restrictions I

will not go into detail, but it is important to note that there are a variety of arguments which have been put forward in response to these problems²⁴.

So mental causation on this interpretation would be a case of an event raising the probability of a physical event occurring in virtue of its instantiating a mental property. Formally; event e causes event e^* in virtue of instantiating M iff $P(e^* | e \& M) > P(e^* | e \& \neg M)$. The probability of my drinking the ginger beer was higher given that I was in a certain physical state and thirsty than it would have been given I was in a certain physical state and not thirsty.

2.3.3 Interventionist Theories of Causation

The third set of theories I will discuss here are interventionist accounts²⁵, particularly as I will make reference to these later in my thesis when I discuss the placebo effect²⁶. Interventionist accounts relate to counterfactual and probability raising accounts of causation. For example, you could think of interventionist accounts as being counterfactual accounts which promote a certain type of semantics. The counterfactuals we are to evaluate involve worlds where the antecedents are realised by interventions. Further, probabilistic interventionist theories could be considered a version of probability raising accounts. So, when an intervention raises the probability of an event occurring then it can be considered a cause. This paragraph shows that it can sometimes be

²⁴See for examples Hitchcock (2001), Fenton-Glynn (2009) and Kvart (2004)

²⁵I followed Woodward (2013) in my exposition of this section.

²⁶Indeed one strength of such theories is that they are made wide use of outside philosophical circles, for example in medicine. This would lead us to believe that, like Newtonian Physics, whether they are correct or not, they are at least prima facie effective at getting correct enough answers. Such putative empirical success seems like a good reason to me to consider such theories carefully.

difficult to clearly delineate one set of theories of causation from another. I don't think is problematic from my standpoint as I do not need to tie myself to any one particular theory of causation for my argument to stand. Indeed, I'd like to remain as neutral as possible on this issue.

Interventionist accounts hold that A is a cause of B if and only if the correlation between A and B holds even after manipulation or intervention. I will be sticking to interventionist accounts and not referring to manipulability accounts which make reference to human agency in their explanation of causation²⁷. This is not least due to the entanglements which would no doubt occur when using human agency to talk about the causation involved in human agency. However, I believe a more philosophically important reason exists to not make use of manipulability theories of causation which is their reduction of causation to human agency. Without spending too much time on a tangent, I believe this is fundamentally mistaken and anthropocentric. Surely there was plenty of causation happening before there were even humans around to have agency. Indeed, the theory of evolution, which I believe gives us part of our reason for thinking that mental states are causally efficacious must rely on the notion of causation while explaining how humans came into being²⁸. Therefore I think to reduce causation to human agency is a mistake it is best to avoid.

Interventions on the other hand do not depend on human agency. Woodward 2003 defines an intervention as;

²⁷See for example G. H. von Wright (Georg Henrik) (1971).

²⁸It could be of course that the evolution of human agency brought about a new kind of causation, distinct from the causation which had been in play up until that point. This new type of causation could be what manipulability theories are tapping into. Against this I will only say that I see no reason which this should be the case and parsimony warns against just this kind of proliferation of kinds.

“(I’s assuming some value $I = z_i$ is an intervention with respect to Y if and only if I is an intervention variable for X with respect to Y and $I = z_i$ is an actual cause of the value taken by X ” p.98

where the aforementioned intervention variables make no reference to human agency or manipulation. So while a human manipulation can be an intervention for Woodward, not all interventions are human manipulations.

As they avoid any such anthropocentric pitfalls I think they are worth examining. A is a cause of B if and only if the correlation between A and B holds even after intervention. Thus if B and C both share a common cause A, then there will be correlation between B and C. However, this correlation will not hold under intervention. Take the barometer example. Although the storm and the barometer reading are correlated, this correlation breaks down upon intervention. Say we intervene (though again, interventions need not be human manipulation) fix the barometer reading to storm. Because of this intervention, the barometer reading becomes independent of whether the storm occurs or not. However, despite our fixing the barometer to read ‘storm’ the probability of the storm coming about remains the same. This demonstrates that the barometer reading is not the cause of the storm.

As James Woodward (2013) points out, interventionist accounts can deal with some preemption cases which counterfactual theories can’t handle. Take for example the problematic gunman case outlined below. These type of cases were tricky for counterfactual theories because they seem to be cases of causation without counterfactual dependence. The first gunman who shoots and kills their victim is the cause of the

death, but there is no counterfactual dependence because the second gunman would have shot and killed the target if the first gunman had thought better of it. On the interventionist accounts, holding the actual action (non-shooting) of the second gunman fixed, the intervening on the first gunman's actions, that is, preventing the first gunman from shooting, brings about a change in the effect, that is the target is not shot dead. So, in this case, the first gunman's shooting would qualify as a cause despite the lack of counterfactual dependence.

Again, space prevents me from going into more detail but there have been many defences put forward for interventionist theories. Rather than explore these however, I will now move onto the the last set of theories I want to discuss; Process Theories.

2.3.4 Process Theories of Causation

There are many variations on Process Theories so I will focus on Dowe's Conserved Quantity Theory of Causation as the most popular version of such theories²⁹.

The idea behind Conserved Quantity theories of causation is the exchange of some conserved quantity from one object to another. Usually the conserved quantity in question is something like energy or momentum. For example Fair (1979) characterises it as "energy-momentum transference in the technical sense of physics" p219. What counts as a conserved quantity is taken from physics which therefore makes such theories quite empirically based. If a non-conserved quantity is exchanged then this is pseudo-causation. Dowe (1995) p324 gives shadows as an example of a pseudo-object

²⁹I followed Dowe (2008) in my discussion of this topic.

which is not capable of acting causally. That is because it only possess non-conserved-quantities such as shape and size. They can change their size and shape but as they possess no conserved quantities, they are incapable of causing things. If two shadows cross, they will both leave with the same conserved quantities as they arrived; which is to say none. Rather it is the surface they are cast upon which is the genuine object.

So what is a causal process on such a view? It is the world line of an object (a causal interaction then being an interaction of two or more of these world lines). A world line of an object is the set of all of the space-time points in its history. An object is anything specified in the ontology of the best current scientific theories. A conserved quantity is a property which is universally conserved such as energy although other versions of such theories put forward other quantities which are also conserved for example momentum or charge.

I want to question the causal power of pseudo-objects however. Shadows are, on Dowe's view, unable to have causal power. But, I have reacted to shadows in the past, usually by jumping at them, which means they have caused me to do things. This would suggest that either shadows do possess a conserved quantity of some kind (which defies empirical evidence), or that transfer of conserved quantity is not required for causation. This is similar to the problem of causation by omission or absence which conserved quantity theorist have trouble explaining. Dowe (2001) gives the example of a father's inattention causing a child's accident. Which conserved quantity was transferred by 'not paying attention to the child'? Schaffer (2000) gives the famous example, first given by Hart & Honoré (1985) p.38, of "the gardener's failure to water the flowers (absence) caused them to die" p.295.

Dowe deals with this problem by positing another, secondary, type of causation he names causation* or quasi-causation which relies on counterfactual dependence. How does this work with my shadow problem? Dowe could say one of two things. He could say this is a case of causation* and that had the light been there I would not have been scared. Alternatively, Dowe could reply that, although it's a convenient way to speak as though the shadow did the causing in this case, it was actually the ground the shadow was cast on which made me jump. I do not find either of these replies particularly satisfying however due to space I will have to put this issue to one side.

Mental causation could potentially be problematic for conserved quantity theories because it is unclear how mental states can possess conserved properties and which conservation laws would govern such cases. This wouldn't be a problem if mental states are identical to brain states, however I will rule this out (see Chapter 5). Assuming mental states could possess conserved quantities and assuming they're not identical, the process theorist would likely need to find a new conserved quantity for mental states. It's further unclear how this new conserved quantity would interact with the kinds of conserved quantities which have already been mentioned; energy, momentum and so on.

Again, there are various replies to these issues but I will not go into them here but will rather move on now to the CEA itself.

2.3.5 Mental Causation

So, what is it that I mean when I talk about mental causation? This will be important to set down as it will be key to understanding what my question is let alone what the answer is. I mean that a mental event, such as a feeling, for example thirst, is the cause of another event either physical or mental. It is important that the mental event is either a true mental event (in a Kimean sense) or is an event which is causally efficacious in virtue of its mental property (in a Davidsonian sense). For the purposes of this thesis I will confine myself to talking about mental events causing physical events because this specifically is the kind of causation the CEA rules out. Furthermore, the above theories of causation are often interrelated and may be difficult to distinguish clearly from each other. However, I don't find this too problematic for me as I do not want to tie myself to any particular theory of causation or events, so as to remain as open as possible when analysing the CEA.

3

ARGUMENTS FOR MENTAL CAUSATION

Here I will lay out my three reasons for thinking Mental Causation exists and that it is a phenomenon worth arguing for. The first is the Mental Manifest Image, the second is an argument from evolution and the third is an inference to the best explanation. I will end by introducing and replying to a potential problem with my view.

3.1 The Manifest Image

Having introduced the scientific side of the picture I shall now introduce my gloss on the Manifest Image (MI) (Sellars (1963)) which is ostensibly in tension with the straightforward physicalist picture.

(MI) The world is how it appears to be to us and we should try to accommodate this in our philosophical theories in so far as is possible given our best current scientific theories.

Thus we take it to be good practice to prefer theories which, all other things being equal, accommodate our everyday view of the world over those which are inconsistent with it.

More specifically I will be working with the Mental Manifest Image (MMI);

(MMI) Our introspections and mental phenomenology are how they appear to us to be and we should try to accommodate this in our philosophical theories in so far as is possible given our best current scientific theories.

To reiterate, MMI means that, all else being equal we should prefer scientific theories which maintain and explain our introspective lives as they appear to be, over ones which deny or are inconsistent with our introspective evidence. Part of our everyday, introspective and phenomenological experience is that mental causation exists and brings about physical effects. You feel pain so you move, you feel thirsty so you drink and it appears to us that these pains and thirsts are the causes.

3.2 The Evolutionary Argument

Aside from the pull of everyday introspection, an evolutionary argument could be given for, at least *prima facie*, thinking that mental states can be causally efficacious. The staggering number of species, with varying degrees of biological relatedness to each other, must have evolved consciousness¹, and this convergent evolution gives us reason to believe that there must be something at work behind this. That so many creatures would evolve to become conscious is then either a huge coincidence, perhaps a by-product or free rider, or it is useful for survival and reproduction and was thus selected for which means it must have causal efficacy. Given the variation and genetic distance in the many examples of this phenomenon, it would seem unlikely to be the former². Therefore, we have two reasons to take mental causation seriously; introspection and evolutionary considerations.

One example of a mental state which has evolved because it helps survival and reproduction is disgust³. It seems to have clear evolutionary benefit; those who avoid items which provoke a disgusted response are less likely to get infections or illnesses which could potentially harm them. This makes sense when considering mental states as causally efficacious but less so when you think of the random patterns of neural firings which would bring about such effects.

¹Naturally to different degrees, but almost indisputably to some greater or lesser extent. Many animals do not possess self-consciousness, for instance, but it would be more of a surprise to me if not all, or at least most, animals have pain qualia, for example, than if they don't.

²The sceptic could still press the point that consciousness is a necessary result of the kind of central nervous system which confers evolutionary advantage. However the burden of proof would be on them to explain why it's the neural and not the mental states doing the causal work.

³See Curtis et al. (2004).

Furthermore, many mammals can be said to have conscious mental states. Animals as diverse as cats, elephants, humans and whales, all shared an ancestor but evolved along divergent lines from that ancestor on. That they all are conscious then is perhaps less surprising if we think that their last common ancestor was itself conscious. What remains surprising though, is that all these mammals would develop the level of complex consciousness and mental abilities unless we think that our last common ancestor also has this level of consciousness. With no direct proof I can not say that it didn't, but it does seem highly unlikely.

Even non-mammals can have some complex intelligences. Take octopuses for example. Katherine Courage (2013) explains why octopuses are a good example for my purposes;

“chimpanzees are, like humans, primates. Dolphins are mammals. Even clever crows and ravens are vertebrates. But our last common ancestor with the octopus was probably some kind of wormlike creature with eye spots that lived as many as 75million years ago, the octopus has a sophisticated intelligence that emerged from an almost entirely different genetic foundation.”

She quotes Peter Godfrey-Smith; "octopuses are the closest thing we have" (ibid) to alien intelligence. This makes octopuses a useful example for pressing my point that consciousness likely evolved because it has a causal impact rather than merely free riding on some other adaptation. The argument runs thus; both humans and octopuses have complex intelligences and both possess consciousness. Our last common ancestor, while perhaps conscious, almost certainly was neither as intelligent nor as

sophisticatedly conscious as either humans or octopuses today. Humans and octopuses diverged evolutionarily so long ago that each has since evolved in very different ways and are now only distantly related. Given this genealogical distance⁴ it suggests that consciousness emerged in both cases because it was useful in survival and reproduction and not merely as a free rider or coincidence. This suggests it does make some kind of causal difference. This makes it likely that this is a case of convergent evolution and not a simple shared inheritance.

Take the New Caledonian Crow as another example. This is one of only a few species of bird we have so far discovered which use tools⁵. In the New Caledonian Crow's case they trim branches carefully down to hooks to pick insects out of trees. As a bird, humans and crows last common ancestor lies many generations back⁶. Continuing the assumption that such an early organism can only have had basic consciousness, it seems again we have another case of not merely shared inheritance, but genuinely convergent evolution.

As a final example, take the Leafcutter Ant⁷. This extraordinary ant is known to use agriculture to harvest the fungus it eats. It does this by collecting leaves, carrying it back to the nest and waiting for the fungus to grow. Ant societies are commonly known to be complex, but this is the only example we know of where ants use agriculture. I

⁴There will also possibly be energy costs associated with having conscious mental states which would add weight to my argument. However, there is an obvious reply in that it could actually be the underlying brain states which require the energy.

⁵See Hunt & Gray (2003) and Weir, Chappell & Kacelnik for more on this fascinating crow.

⁶"The last common ancestor of birds and mammals lived some 300 million years ago, at a time when the six-layered neocortex, which gives rise to sophisticated cognition in primates, had not yet developed" Veit & Nieder (2013).

⁷See Hölldobler (2011)

think what this example shows most clearly is how little we know about non-human animals and their behaviour. This leads me to believe that the more we learn, the more consciousness we will uncover, all of which lends support to my argument from convergent evolution. There are many more examples I could give here of putative cases of genuine convergent evolution which would support my case, however I will now move on to my third reason for thinking mental causation exists.

3.3 Inference to the Best Explanation

My third argument for the existence of mental causation comes from empirical evidence which leads to inference to the best explanation. These come from the medical sciences, particularly from psychology. It seems a much better explanation that mental properties are causally efficacious in the case of placebos and talking therapies than the sets of underlying neuron firings. In fact, I take the causal efficacy of mental events to be the best explanation of these phenomena (and of the MMI more generally), at least given our current understanding of the brain.

I will start by examining the placebo effect. A placebo is a medical treatment which has casually efficacy based not on some chemical mechanism, but rather one based on expectation. Thus the same drug could be administered to the same patient under different names and be efficacious in one case while failing in the latter (or can be more effective in the former than the latter). The difference in the two cases is the mental state of belief in the first case that it will be efficacious and the belief in the second case

that it will not be⁸. Alternatively, a sugar pill can be substituted in place of a drug with a known effective mechanism. If the patient believes that the sugar pill is medicine, then by the placebo effect, they can derive curative benefit⁹. In this case, a treatment which is known to have no medically or chemically relevant properties actually proves effective. In one study a placebo marked as an active drug was statistically as effective as the active drug marked as a placebo (Kam-Hansen et al. (2014))¹⁰.

Let's press on the distinctness of the causal mechanisms some more. Say you have two drugs. Effectron (E) is a drug which blocks pain receptors and therefore reduces pain by a chemical mechanism. Pretendtron (P) is a totally chemically inert drug which, by definition, does not interact chemically with the brain in any relevant way. However, when P is represented as being effective its use has been shown to reduce pain. It therefore works by a mechanism which must be something other than a chemical mechanism. As the difference maker in the P case is how the drug is represented to the patient, the putative mechanism at work here is a psychological one based on belief and expectation. This at least places a burden of proof onto someone who wants to deny mental causation to explain why their purely physical mechanism is to be preferred.

What's more, the placebo effect has been shown to be robust. It's been shown there are

⁸The two sets of underlying brain states will also be different but as I will go on to argue it's hard to see what the difference in the two sets could be causing an effect whereas the difference in mental states does make explanatory sense.

⁹Although it's also important to note that traditional medicines also carry their own placebo effect as well. It's almost impossible to avoid.

¹⁰For another example of a study which found a positive placebo effect see Kaptchuk et al. (2010). Ben Goldacre (2009) also summarises the results of many placebo studies in his very helpful chapter *The Placebo Effect*. However, it is important to bear in mind this is obviously a very small sample of a much larger literature and I have not directly searched for papers which provide evidence against the placebo effect.

predictable patterns in response to different types of placebo. For example, studies have shown that a placebo injection of water is more effective than a sugar pill (Goldacre (2009) p. 70). The packaging (and price) which medicines are sold in (for) has also been shown to affect their perceived efficacy which goes some way to explaining how you can buy the same paracetamol for 30p or £3 (Goldacre (2009) p.71). All this sums to suggest that there is some systematic phenomenon in play with the placebo effect, not mere coincidence.

Furthermore, and in some ways most persuasively, there is evidence to suggest that in some cases, intervening directly on a person's mental states is at least as effective as intervening directly on their physical brain states¹¹. This is the case with certain mental health conditions such as depression and anxiety. There are both drug and talking therapies available to treat these conditions. These can both be effective independently, however, it is believed that a combination of treatments is the most effective. In some patients though, only talking therapy proves effective, no amount of drugs will bring about the desired result.

“CT [Cognitive Therapy] is the best-known and most widely tested of a family of cognitive behavioural interventions. Like ADM [antidepressant medications], it is a safe and efficacious treatment for acute episodes of major depressive disorder. CT is based on the premise that inaccurate beliefs and maladaptive information processing (forming the basis for repetitive negative thinking) have a causal role in depression. This ‘cognitive model’ posits that when maladaptive thinking is corrected, both acute distress and

¹¹Which would make the mental state the cause on an interventionist account of causation.

the risk for subsequent symptom return will be reduced. Contrasting with the lack of evidence of enduring effects of ADMs is the substantiation of claims that CT provides protection against relapse and, possibly, recurrence” (DeRubeis et al. (2008) p.790)

Indeed, the success of counselling as psychological treatment, and the increasing demands for counselling services, speaks (excuse the pun) to how effective intervening directly on mental states to achieve a desired goal is. Of course, its not entirely impossible that this is not a function of our relative lack of knowledge about the mind, brain chemistry and how the two interact. Future medicine may indeed have the drugs or equipment required to most effectively manipulate brain and mental states. That said, even in a future where psychiatric treatment has entirely dispensed with all talking therapy, it works now, and that’s reason enough to examine it more closely. It’s also reason to think that even if in the future, another, purely physical, explanation could be put forward, that this may be compatible with mental causation. Again, this places a burden of proof on the denier of mental causation to explain away this potential compatibility.

To put my point in a different way, say that C is the effect of being cured, D is the known mechanism of drug efficacy and D^* is the mental state of expecting the drug to work. If it’s the case that the mental state is doing some causal work then we this inequality would hold; $P(C|D) < P(C|D \& D^*)$. This is exactly the kind of result placebo studies indicate. Given that the mental state here seems to raise the probability of the event coming about, it would count as a cause on a picture of probabilistic causation. So it seems we have empirical evidence of mental causation bringing about physical

effects.

But, is this a case where the only difference is the mental state? It certainly seems so in an everyday sense of talking about the placebo effect. Even the medical literature appears to take it for granted that it is the feeling of expectation, or belief in efficacy, which is making the difference in the two cases. Of course, in both the placebo and the specific pain case it could be the brain state underlying the mental state which is making the medical difference. The placebo effect in and of itself then, cannot be used as an argument against the CEA. It does not prove that mental states can have non-overdetermining effects over and above their underlying physical state.

There may be more reason to think that the mental state is what is actually at work here than just the feeling that a brain state making the difference here is weird. And that comes from thinking of the placebo effect in the context of scientific laws. In order to feature in a scientific account, the phenomenon in question needs to be a natural kind¹². If I'm right and there is no multiple realisation then Yablo's argument (more on this below) does not go through because pain would be a disjunctive property unavailable to be playing a causal role. Rather it would be the more specific type of pain (say 'throbbing pain') which would play the role. 'Throbbing pain' seems to me to be a better candidate for a natural kind than whatever set of neural firings underlies it.

Take the counselling case again. What would the neural correlate of a conversation be? A mass of neural firings. Take the placebo case again. What would the neural correlate of an expectation be? Again, a mass of neural firings. Not plausibly a natural

¹²See for example Goodman's debate on 'grue' and 'bleen' in (Goodman (1983) and Goodman (2000)). For more, related discussion see Kim (1992) on the relationship between jade, jadeite and nephrite.

kind. Maybe, the mental state could play the required natural kind role here? And if the mental states provide a better candidate for natural kinds, and therefore are more suited to feature in laws, then is this not reason to think that they may be playing a causal role?

In summary, there is a burden of proof argument on those who argue against the placebo effect and mental causation more generally. The medical community treat the placebo effect as though it works based on expectation. There is undeniable intuitive appeal to the idea that the expectation brings about the result.

3.4 A Potential Problem

There is a potential problem with my two arguments for mental causation that I would like to address before continuing. It might be thought strange that the difference between mental causation existing and not existing is just the difference between whether determinism holds or we live in a probabilistic world. After all, why would the MMI not hold in a deterministic world. It would be an illusion but it could still hold. Similarly, consciousness could have evolved in a deterministic world, perhaps by free riding off its physical base.

However, what reason is there to think that in a deterministic world, we would have the same MMI? There would be a different scientific, physical picture, so why not a corresponding different MMI? In fact, given that the physical pictures would be different, should it not be expected the mental one would be too? Perhaps the deterministic

MMI would not give us phenomenal experience of mental causation. I have made the assumption that we live in a probabilistic world, therefore I have to assume that we only have access to the probabilistic MMI. In fact, the CEA may give us reason to think that, at least as deterministic worlds go, there is no mental causation. Therefore, we have no reason to think that beings in a deterministic world would feel as though they do. The MMI would be an illusion in such a world. So the best explanation of why the MMI holds true, is that ours is a probabilistic world in which mental causation does in fact exist.

If it is the case that the deterministic MMI does not include the phenomenal experience of mental causation, then this would fit with the CEA holding in deterministic worlds. There would be no tension between the two and no philosophical problem.

To summarise in a sentence: Our MMI depends on mental causation and in turn my argument for why mental causation may exist depends on the world being probabilistic, so why think that our MMI would be the same in a deterministic world? I will be coming back to this issue in a later section when I have covered more ground.

4

THE CAUSAL EXCLUSION ARGUMENT

To deny the MMI and our introspective experiences of our own mental states is to deny some very direct and compelling evidence, and would be a huge bullet to bite. It is my aim to argue that the CEA is therefore wrong. Let me now set out the CEA a little more clearly:

- (P1) Causal Closure of Physics
- (P2) No Systematic Overdetermination
- (P3) The Non-Identity of the Mental and the Physical
- (C1) There is no Mental Causation of Physical Effects

(P1) is the principle that for any physical effect, there is a sufficient physical cause. Kim states “if you pick any physical event and trace out its causal ancestry or posterity, that will never take you outside the physical domain” (Kim (1998) p.40). While no conclusive argument can be given for holding causal closure, it seems natural that a physicalist would be loathe to give it up. As Kim notes, giving up on causal closure means giving up on the completeness of physics.

(P2) states that there is not any systematic overdetermination in the world. This means that there should not be more than one sufficient cause at any given time for any given event. If physical states and mental states are both causally efficacious (and not identical) then this could potentially result in widespread systematic overdetermination. This is because, a wide range of events would have more than one non-identical sufficient cause, the physical brain state and the mental state. Kim argues that while occasional cases of overdetermination happens (think of firing squad cases¹), this would be strange at a systematic level.

Lastly, in (P3), Kim makes the antireductionist stipulation that there is no identity between the mental and the physical. This is the premise that Kim rejects in order to dissolve the CEA. In support of his reductionism, Kim offers the analogy with properties from the special sciences. Like mental properties, special science properties supervene on basic physical properties. Kim argues that the reduction of special science properties to basic physical properties seems unproblematic². Why not then, Kim suggests, hold

¹Although perhaps a more fine grained view of such events would overcome any overdetermination worries.

²Although this is not a stance agreed upon by everyone, for example see Fodor (1974) and Fodor (1997).

the same for mental properties? I will briefly examine this premise in the next chapter.

If the above three premises hold then the CEA goes through and (C1) follows. If mental causation is adding nothing to our theories about the world, then there is no need to posit it. Having laid out the three premises I will not explain in more detail why, when combined, they lead to this conclusion.

So why is (P1) necessary for the CEA to hold? It is required because if mental states are causally efficacious, they are efficacious precisely because of their mental aspect (either in a Kimean or Davidsonian sense). And if causal closure didn't hold then the causal efficacy of these non-physical states would be create widespread systematic overdeterminism. Causal closure says nothing about mental states causing mental events (and thus is not the target of the CEA). But mental to mental causation is not the phenomenon at hand. It appears therefore, that if a premise is to be eliminated or modified, this will be it. Indeed, I will modify it, as I have noted above to (P1'): every physical effect has a physical cause which is sufficient to fix its probability. I will examine this premise further in Chapter 7.

Say, while stipulating non-identity, that you do believe that mental states are causally efficacious in this way then P(1) must be false. If you take an event, say my drinking some water, and trace the causal ancestry of my drinking (a physical event), in order to find a sufficient cause, you will have to look outside of the physical realm to my thirst (a mental event). This violates causal closure. However, if causal closure did hold, then the physical basis would be sufficient and the mental cause, if it exists, would be superfluous. The event of my drinking would therefore be overdetermined. This would

of course apply to all examples of mental causation. The causal exclusion argument would then, if correct, lead us to believe that the mental has no causal power on the physical. This is because, if every event has a sufficient physical cause, any and all mental causes would be superfluous (and is indeed ruled out by P(2)).

But why think widespread and systematic overdetermination doesn't happen? Firstly, intuitively it seems odd (Kim calls it "implausible" (Kim (1998) p.44)). Secondly, parsimony would suggest that, all else being equal, we should prefer theories which do not over cause effects rather than those which do. Or more precisely, we should prefer theories which do not posit more than one sufficient cause for a given effect if one will do. Neither of these are knock down arguments however. Intuition and parsimony are both good guides, but not final words. I will examine this further in Chapter 6.

The CEA appears to require our world be deterministic. That is because causal closure (in the sense of physical effects having *sufficient* physical cause rather than in a probability fixing sense) only seems to hold in a deterministic world if it holds at all. I will explain in Chapter 8 why it's implausible that causal closure would hold in a probabilistic world. As I've stated though, it seems unlikely, or at the very least questionable, that the world is indeed fundamentally deterministic. If the world is indeterministic what consequences would that have for the CEA? Would it merely render a probabilistic version of the same argument, or would such a shift cause issues for the argument going through? If it were the case that such a shift did cause a problem for the argument going through in an indeterministic world, that does not preclude the possibility that the would still hold. Before going on to discuss this line of thought I first must reply to an influential counter argument against the CEA made by

Yablo, because if the CEA can be defeated in the way Yablo wants to, there would not be any reason to continue to argue against it.

4.1 Mental Causation and Multiple Realizability

Stephen Yablo (1992) disagrees with Kim that the physical can never leave causal room for the mental and thinks that there is a case to be made for the mental having genuine causal power. I take it that Yablo thinks the CEA is unsound because he doesn't think that mental causation would cause widespread systematic overdetermination, but I will explain this more fully below. His argument rests heavily on the concept of multiple realizability and the idea of proportionate causal explanation. It is important that I discuss Yablo because if he is right then the CEA does not go through even on its own terms.

The thesis that the mental is multiple realisable by the physical is the idea that, contra Identity Theory, there's no one-to-one mapping from the physical to the mental. One mental state can be realised by more than one physical 'realiser' or base. For example, pain can be realised by more than one underlying neural brain state. Analogously the colour red is multiply realised by its different shades. The base necessitates the realised property but not vice versa exactly because more than one base can realise the same property. In other words, there is "asymmetric necessitation" Yablo (1992) p.250. Holding supervenience along with this asymmetry leads Yablo to characterise the mental/physical relation as one of determinate/determinable. For example, crimson is the determinate of the determinable, red. Yablo defines this Determination Relation as;

“(Δ): P determines Q iff: for a thing to be a P is for it to be a Q, not *simpliciter*, but in some specific way” (1992 p.252).

On a metaphysical level this translates to;

“(Δ) P determines Q ($P < Q$) only if: (i) necessarily, for all x, if x has P then x has Q; and (ii) possibly, for some x, x has Q but lacks P” (1992 p.252)

Yablo constructs an analogy between mental properties and events claiming that “we find that the relation between mental and physical events effectively duplicates that of mental to physical properties” (1992 p.270) which yields this principle;

“(d) A mental event m occurs iff some physical determination p of m occurs” (1992 p.271)

where the determination relation for events is;

“(δ): p determines q iff: for p to occur is for q to occur, not *simpliciter*, but in a certain way” (1992 p.260).

Again this translates to;

“(δ): $p > q$ iff: (i) necessarily if p exists, then q exists and is coincident with p; (ii) possibly, q exists and p does not exist” (1992 p.265).

In other words, an event p is determined by q if for p to occur then q must also occur in a certain way. For example “Icarus’s flying too near the sun determines his flying *per se*.” Yablo (1992) p.261. And there exists a world in which Icarus flies but does not do so too close to the sun.

Yablo's idea is to try to hone in on the most proportionate causal explanation for any given effect. "It seems clear" Yablo (1992) p.277 to him that a more proportionate cause will be preferred as an explanation for an event than a less proportionate cause. The most proportionate cause is one which doesn't leave out causally relevant information but at the same time doesn't include too much. This requires that the cause be both contingent³ and adequate⁴ which guarantees that the effect would not occur without the cause and counterfactually, would occur if it did⁵. He gives the death of Socrates by guzzling hemlock as an example (1992 p. 275). To say that his guzzling the hemlock caused his death is to be overly specific. Presumably a drinking event could have occurred even if a guzzling event had not and this would still lead to death by poisoning⁶. This is therefore a violation of contingency. As a violation of adequacy, Yablo gives the example of a safety valve which, due to a freak malfunction, stops opening at the correct speed. This allows pressure to build which eventually causes the boiler to explode. The opening in itself is not adequate for the effect; the slowing of the door opening is needed to cause the explosion. To see this, think of the scenario in which the malfunction had not happened but the opening had. This is just the scenario in which the valve functions as normal. In that case it is to be hoped that the boiler would not have exploded.

Yablo introduces Sophie the pigeon who has been trained to peck at red objects. One

³Defined as "(C) If x had not occurred, then y would not have occurred either" (1992 p.274).

⁴Defined as "(A) If x had not occurred, then if it had, y would have occurred as well" (1992 p.274).

⁵There are two further requirements for proportionality; the cause must be required and enough for the effect, which is to say that the cause shouldn't contain any extraneous causal factors, but that it should include enough to ensure the event.

⁶Unless you think that this would not actually be the same death.

day, her guardians show her a scarlet object so she pecks at it.

“Assuming that the scarlet was causally sufficient for the pecking, we can conclude by the exclusion principle that every other property was irrelevant. Apparently then the redness, although it looked to be precisely what Sophie was responding to, makes in reality no causal contribution whatever” (1992 p.257).

Yablo thinks this result is nonsensical. Of course it is the redness which is causing the pecking, it's the very thing she's been trained to peck at. Just as in the hemlock example, the pecking was not contingent on the scarlet, because if scarlet had not been presented, but some other shade of red had, then the pecking still would have happened. Further, the redness seems to be an adequate cause of the effect because, on the presumption that the object was in fact green (and thus remained un-pecked) had the object been red then it would have been pecked. In this case then, the more proportionate cause of the pecking was the object's redness, not its scarletness. If this holds for mental and physical properties, as Yablo claims it does, then the CEA fails to go through. He is thus questioning the validity of the CEA. This is because if there is a single, more proportionate mental cause, then there will be no widespread and systematic overdetermination.

Will there be times when it is more proportionate to explain a given action by mental states rather than the underlying physical states? Yablo thinks so, in cases where the effect in question does not depend too sensitively on its physical implementation. He gives as a final example, his ringing a doorbell (1992 p.278). He hypothesises that there

are many physical implementations that can instantiate the decision to ring a doorbell. This means that despite having a specific physical determination, that particular brain state is akin to scarletness in this case. The effect is not contingent on the brain state. Other brain states are available which would have instantiated the decision and would have caused the doorbell ringing. Without the decision altogether though, it seems as though there would be no bell ringing. So the decision is adequate for the doorbell ringing. So, here we have an example of a mental cause being more proportionate to an effect than a physical one and therefore should be the preferred explanation of the cause.

To return to why Yablo thinks the CEA is invalid, I take it that he thinks because there will be only one most proportionate explanation for any given event. If this is the case then there will not be any widespread systematic overdetermination. I'll now consider an objection to this argument.

4.1.1 Wrong Grain Objection

My objection to Yablo is that he is comparing different grains in the physical and mental cases. The mental state 'pain' is a much more coarse grained mental phenomena than 'specific set of neural firings' is a physical state. There are many ways of being in pain, compare a paper cut to a headache (or even compare a throbbing headache to a searing headache) whereas there's only one way to exemplify a certain set of neural firings. I think it would be more accurate to compare 'pains' to 'sets of brain states which correspond to pains' and 'specific searing headaches' to 'specific sets of neural firings'.

If we were to individuate these pains ever more finely, the idea of multiple realizability and asymmetric determination become more questionable. Therefore while Yablo may be right that the most proportionate explanation is the preferable explanation, he is wrong in which explanations he picks out as being proportionate in that he thinks there's a mental explanation which will be more proportionate than the physical one.

But what about propositional attitudes? I think the same can be said for them. So rather than talking about 'belief' and 'specific set of neural firings' but rather either 'belief' and 'sets of neural firings' in the general case, or 'specific belief' and 'specific set of neural firings' in the particular case. However, as I have stated before, I do not intend to focus on propositional attitudes. I don't need to argue that all mental states are causally efficacious as long as I can show that *some* mental states are efficacious. So, if you don't find my argument against Yablo strong in the case of propositional attitudes, then I still think my argument is plausible against phenomenal states involving qualia.

If mental states are not multiply realisable by their physical bases then pain, rather than being one natural kind is a disjunctive kind including many types of pain. Sets of pain states are multiply realisable by different sets of brain states, but whichever grain you chose to examine, there will always be a corresponding physical state or set of physical states. So, the mental states will never be the most proportionate cause or at least there will be an equally proportionate mental state.

An interesting analogous debate is had between Fodor and Kim with regard to natural kinds in the special sciences and whether these can be related one-to-one with physical

natural kinds⁷. If the different pain kinds do correspond one-to-one with neurological kinds then multiple realizability looks less plausible, at least in this case.

Given all these various considerations, I do not find Yablo's argument convincing. While I agree that the most proportionate explanation should be preferred, I do not think he picks out the correct proportionate response in token cases. The redness is the most proportionate cause of a pecking but the scarletness was the most proportionate cause of *that* pecking. Likewise, pain may be the most proportionate cause of reaching for the paracetamol in general (although that claim could also be had by the appropriately selected group of brain states), but the brain state will be an equally proportionate cause for this reaching for the paracetamol. Or at least, there will be a brain state or set of brain states which is *equally* proportionate to the mental state. I don't want to argue that it is always the case that there is a *more* proportionate physical explanation, or I will have ruled out the possibility of mental causation.

Could it be the case that some mental properties could be the most proportionate cause in the more coarse grained physical cases? So thirst would be the most proportionate cause of drinking (although this particular thirst, determined by this particular brain state, would be the most proportionate cause of this drinking). Whether or not this is the case, I am interested in actual causation⁸ which deals in token cases. As Baumgartner and Glynn state, "actual causation is also to be distinguished from *type causation*: actual causation holds between token events in a particular, concrete scenario; type causation by contrast, holds between event kinds in scenario kinds" Baumgartner &

⁷See (Fodor (1974), Fodor (1997) and Kim (1992)).

⁸Defined as follows; "An *actual cause* of some token effect is itself a (distinct) token event (or fact, or state of affairs, ...) that helped to bring about that effect" Baumgartner & Glynn (2013) p.1.

Glynn (2013) p.1.

How does this sit with what I've said above about the placebo effect and the effectiveness of talking therapies? I've presented the mental states in these cases to be the most proportionate response even in token cases⁹. That is, not only would counselling¹⁰ be the most proportionate response to depression in general, one particular session of counselling (or set of sessions) would be the most proportionate response to one particular case of depression. Yablo's argument would therefore go through. However, I have also argued that the MMI and further mental causation itself, would likely not hold in a deterministic world. I posit that with no mental causation there would also be no placebo effect and no effective counselling. This is precisely because the phenomenal aspect of these therapies seem to so integral to their effectiveness. It is only because the medicine *feels* like it will work that it works. I'm not sure that without mental causation talking therapies would be efficacious because phenomenal experience without mental causation would be passive. Therefore the CEA, in deterministic worlds, still stands. Having defended the CEA from Yablo's criticism I will now put forward my reasons for thinking mental and brain states are not identical.

⁹I rule out the idea that there could be physical bases which are causally equally proportional based on the idea that such bundles of neural firings are not plausibly natural kinds.

¹⁰It's important to note that I take counselling to be a case of mental to physical causation because the aim is to intervene directly on mental states to bring about a change in behaviour.

5

THE NON-IDENTITY OF THE MENTAL AND THE PHYSICAL

“The first observation I make at this point is that there is a great difference between the mind and the body, inasmuch as the body is by its very nature always divisible, while the mind is utterly indivisible.”

Meditations on First Philosophy -

Descartes (2013) p.119

I will introduce the premise that there is no identity between the mental and the physical. I will spell out what exactly this premise means and why Kim rejects it. I will then give three reasons to accept the premise which I do.

Kim rejects this premise allowing him to get round his own argument and avoid the unpalatable conclusion that no mental state or property can bring about physical events. How does denying this premise avoid the unwanted conclusion? It works because if the mental and the physical are identical then they can both be playing a causal role without any overdetermination taking place and without violating the principle of causal closure.

Kim is therefore a reductive physicalist. However, many philosophers find this conclusion to be just as unpalatable as the conclusion Kim manages to avoid. Hence, the many different types and kinds of physicalisms and the many and varied attempts to work around the CEA by different methods, including mine. Why might someone find the idea of the reduction of the mental to the physical unpalatable? In order to answer this question I will now spell out what it means for the mental to be identical with the physical or for the mental to be reducible to the physical.

An early and well known proponent of Identity Theory was J.J.C. Smart (1959). His was a *type* identity theory (as opposed to a *token* identity theory). The idea behind identity theory is that there is nothing over and above the mental than the physical state underlying it. In token identity theory each particular mental state will be identical to some brain state. Whereas type identity theory says that each type of mental state or property has its type of identical brain state. Analogies have been put forward to water being identical to H₂O (Kripke (1980)) and lightening being identical to electrical discharge (Smart (1959) p.145). Whether such paradigmatic cases of physical reduction apply to the mind is of course a further and contested issue. According to identity theories, it may be possible (although in practice this may never be the case) to

reduce every mental state to its underlying physical brain state. The famous empirical example given is ‘pain’ and ‘C-fibers firing’¹. So every case of pain is a C-fiber firing and every case of C-fiber firing is a case of pain².

How does this premise work when placed into a probabilistic setting? In my opinion, there is little difference between the two settings for this premise. Whether or not the world, and causation is deterministic or probabilistic, all this premise states is that the mental and the physical are not identical. Or, in other words, the former is not reducible to the latter. This could be the case, or not the case, in either a probabilistic or a deterministic world. However, I do think there are reasons to think it is not the case that the mental and physical are identical despite the fact that I don’t think the mental is multiply realisable by the physical; the ‘Explanatory Gap’, ‘What it’s Likeness’ and Jackson’s Knowledge Argument. I will only briefly summarise each as I want to focus my discussion on the other two premises.

¹See Rorty (1965) for example.

²Empirically this is now known to be extremely simplistic both because C-fibers are responsible for other sensations than pain and because there are other fibers associated with pain sensations. See Puccetti (1977) p.303. Scientists now believe they have narrowed down at least two kinds of fibres the firings of which produce difference kinds of pain sensation. They call these C-fiber and A δ -fiber firing respectively. A study “identified ‘pricking’, ‘dull’ and ‘pressing’ as distinguishing best between A δ mediated (punctate pressure) and C fiber mediated (blunt pressure) pain sensations” (Beissner et al. (2010) p.3). This in itself is not important to the identity theorist as it does nothing to show that their theory is wrong. However, I do think it lends more weight to my argument against multiple realisability I made above.

5.1 The ‘Explanatory Gap’

One reason to think that Identity Theory may be wrong is the ‘explanatory gap’ (Levine (1983)). Levine’s argument is that reducing phenomenal mental states to mere physical brain states misses something about the mental states in question. The thing this reduction misses is precisely the phenomenal quality of the mental states. Pain is a great example of such a mental state. There is nothing in identity theory, or indeed in *any* reductive physicalist theory which explains why pain feels the way it does. There is nothing in identity theory, or again, in *any* physicalist theory, which explains why this mental state is associated with this brain state.

Levine contrasts pain with heat. For the type identity theorist, heat is nothing over and above the motion of molecules. Trying to imagine heat without the movement of molecules is not possible in the way imagining pain without c-fibres firing is. And this is because;

“the experience of pain, the sensation of pain, counts as pain itself. We cannot make the distinction here, as we can with heat, between the way it appears to us and the phenomenon itself” (Levine (1983) p.355)

Stronger versions of the argument claim that this issue is not merely practical or technical, but that there is *in principle* no systematic way of explaining the gap³. To take the view that pain is reducible to mere C-fiber firing is therefore to sell the

³There are two options to sidestep the problem however. You could deny identity between the mental and the physical and explain phenomenal states as basic or you could take an eliminative stance as Levine suggests.

phenomenon short and to miss something crucial about it.

Why does a non-reductionist view not suffer from this problem? This is because it is open to the non-reductionist to explain the phenomenal character of pain as brute or basic⁴. If they are basic or fundamental, then we should not expect a further explanation to be forthcoming and the gap is no longer problematic. Another way to look at this is that while there seems to be nothing mysterious about heat being nothing more than the movement of molecules, or c-fibres firing, there does seem to be something mysterious about the phenomenon of pain. This mystery implies that something has been left unexplained by the physical explanation alone.

In summary, if there is, as Levine argues, a gap between the mental and the physical, then they cannot be identical.

5.2 Nagel and ‘What It’s Likeness’

The second argument I put forward to motivate the belief that the mental and the physical are not identical is from Nagel (1974). Nagel argues that, because of consciousness, the mind-body problem is uniquely different from other kinds of reductionist questions. Indeed, reductionist accounts fail to explain consciousness adequately. Although there may be more to the story “fundamentally an organism has conscious mental states if and only if there is something it is like to *be* that organism - something it is like *for* the organism” Nagel (1974) p.166. Nagel claims that an objective reductionist

⁴Although it should be noted this isn’t the route Levine takes; he takes an eliminative stance.

or physicalist account will in principle be unable to capture the “single point of view” p.167 which is essentially connected to the subjective conscious experience. Nagel uses bats to illustrate his point. As they echo-locate, and as we assume that they have phenomenal experience, there must be something it is like to sense by echo-location. But, what it must be like, will be so different from anything humans can experience based on our sense modalities, that we would be unable to imagine it. Therefore, there is a fact and that fact is inaccessible (and possibly even inexpressible) for us. As physicalism and reductionism are both objective theories (moving further from individual perspectives) they are both unsuitable for capturing the essentially subjective nature of phenomenal experience. So, if there is something about the mental that the physical story misses, then the two cannot be identical.

5.3 The Knowledge Argument

Frank Jackson’s influential paper *What Mary Didn’t Know* was designed to question our overwhelming physicalist preoccupation. The argument goes that Mary is a future ‘superscientist’ in that she knows all there is to know about a by now complete physical science. However, sadly for poor Mary, she has been trapped in one room for her entire life, and even more sadly, that room is entirely black and white. Putting to one side the possibility of this (let alone the moral implications) let’s assume she has never before observed colour. One day, Mary is released from her room and on that day observes colours for the first time. When she sees a rose for the first time she learns something that she, in principle, could never know in her room; what red actually looks like.

The argument is that her physical education, complete though it was stipulated to be, actually left something out about the world, that thing being qualia, or the phenomenal qualities to the physical structures. As qualia are essential to certain mental states, and further, as qualia are non-physical, there are certain types of (phenomenal) mental states which are non-physical. Therefore the mental and the physical can not be identical.

5.4 Accepting the Premise

So to summarise, my reason for accepting this premise is that I believe it is the case and therefore that to reject it as a premise would be to make a mistake about how the world is. Why do I think it is the case that the mental and the physical are not identical? I am convinced by the arguments put forward by Levine, Jackson and Nagel that there is something that reductive accounts of the mental miss about the phenomenon. Of course, there is some connection between the mental and the physical however, I agree with a majority of contemporary philosophers of mind that the relationship is not one of identity.

6

CAUSATION AND OVERDETERMINATION IN A PROBABILISTIC WORLD

“It all depends if you’ve smeddum or not”

Smeddum - Gibbon & Bold (2001)

An event can be said to be overdetermined if it has more than one distinct, sufficient cause all obtaining at the same time¹. More specifically “say that a set of events A *overdetermines* event b if and only if (i) b would still have occurred if any member of A had not occurred while all the others had, (ii) b would not have occurred if none of the members of A had occurred, and (iii) all members of A have an equally good (or bad)

¹A quick note on the epigraph for this chapter. Smeddum is a Scots word for grit or determination. I hope the reader can forgive this degree of tangential pun in my epigraph.

claim to be a cause of b” (Kroedel (2008) p.128-129) For the rest of this chapter I will examine Sider’s argument that overdetermination is not a bad thing. As I’m interested in probabilistic worlds, I will then discuss whether overdetermination occurs, and whether the aforementioned arguments will still hold, in such worlds.

6.1 Denying The Problem Of Overdetermination

6.1.1 Sider

Sider (2003) doesn’t understand why overdetermination should be seen to be a bad thing. He thinks it’s only natural to want to say a given event can be caused by both a macro-object as well as the micro-parts of the object as well as the fact of the objects causing the effect and so on². He argues that Kim thinks such views are “at best extremely odd” (Kim (1993) p247). Sider considers three different possible objections to overdetermination; (1) metaphysical incoherence, (2) coincidence and (3) epistemic doubts, and then goes onto to explain why these are not actually problematic.

Regarding (1), he argues that no currently popular account of causation (he lists counterfactual, covering law, probabilistic and primitivist accounts) rules out overdetermination³ and to merely state that they *shouldn’t* allow it would be to discard all the aforementioned theories. There may be legitimate worries about epiphenomenalism

²This kind of relationship is different from but analogous to the relationship between the mental and the physical. I will discuss whether this difference is relevant later.

³And shouldn’t given that at least occasional examples of overdetermination, think firing squad cases, do happen. Furthermore it is important to note that Kim is only objecting to widespread and systematic overdetermination.

and its distinction from causation, but if anything the burden of proof falls on Sider's opponent.

Regarding (2), it's important to note that this only works against systematic overdeterminism given that it's unobjectionable that there should be occasional coincidences. Except that even though the overdetermination here does seem to be systematic it does not seem to be coincidental given the necessary connections at play between an object and its parts or in this case the connection between a mental state and its underlying brain state. There is no baseball without its atoms. Further, there is no metaphysical possibility of the baseball existing without its parts. Therefore, it should not be surprising, nor objectionable, that both the baseball and its atoms cause the smashing. To extend this to the mental case, there is a nomological relationship between a mental state and its underlying brain state, therefore it is not problematic that both should bring about an event.

Regarding (3), it can lead to doubt over the existence of some of the entities in question. That is, if both the baseball and its constituent parts both cause the window smashing, then we may have parsimony reasons to doubt the existence of the baseball; it's "epistemically redundant" Sider (2003) p.5. This epistemological point can also be made metaphysically. If the baseball is redundant for the smashing (because the atoms are sufficient) then we have a parsimony argument against the baseball's existence. Sider admits this is a more reasonable objection than the previous two, but it's limited in its power. Rather than these parsimony considerations being an argument against mental causation/macroscopic entities, they're an argument against an argument for mental states and macro-entities. In other words, epistemic worries don't show any internal

incoherence within theories allowing for overdetermination. Rather they only show that we cannot rely on simple causal arguments to argue for the existence of macro-entities/mental causation. That is, we don't need macro-entities/mental causation to account for our experience of the world, therefore we cannot use our experiences to argue for their existence. We need other reasons. Given that few people rest their arguments for mental states/ macro-entities on the kind of simple causal argument which the epistemic argument has force against, (3) is not problematic. Interestingly, as I've outlined in Chapter 3, we do in fact have compelling reason for positing mental causation.

In conclusion, the best that can be said for the three arguments given as to why overdetermination is bad, is that those positing the existence of mental causation/macro-entities need to give more than a simple casual argument for their existence. It just so happens that we have some such arguments in favour of the efficacy of mental causation and moreover these reasons work in probabilistic worlds. First, what I have termed the Mental Manifest Image, secondly, arguments from evolution and lastly the argument that mental states may play natural kind roles in certain scientific theories more plausibly than physical ones.

6.2 Probabilistic Overdetermination

So far I have been looking at overdetermination in deterministic settings. However, does the phenomenon occur in probabilistic settings? An event with overdetermined probability would be the probabilistic analogue of the deterministic overdetermined

event. So, the analogue premise from the probabilistic CEA would rule out widespread, systematic overdetermination of probabilities of events.

On the face of it overdeterminism and indeterminism are uneasy bedfellows. How can one event be both over- and in- determined? Of course, it is not impossible for one event to be both overdetermined and indetermined. I will give an example of such an event below taken from Fenton-Glynn (2009). But, can the same arguments given above in defence of overdetermination be ‘translated’ into the probabilistic setting, and if so, do they still work? If they do in fact succeed in arguing against overdetermination in the probabilistic setting then the CEA would fall through.

6.2.1 Example of Overdetermined Probabilities

Fenton-Glynn (2009) gives the following example of the probability of an event being overdetermined:

“ Example 4.6: Symmetry

Barbara and Claire are armed with rifles and have a small spatio-temporal window for killing Ernst. If they want to kill him, they must shoot at t_1 and must shoot through the same small aperture (perhaps a chink in his armour). Both want him dead and at t_0 each is disposed to fire when the chance comes. Each is an excellent shot and, if she shoots, has a good chance of accuracy (and this chance is independent of whether the other shoots). If either fires alone at t_1 and shoots accurately, then her bullet will travel through the aperture, pierce Ernst’s heart and kill him. If, however, both

shoot accurately at t1 there is a high chance of a collision between their bullets that will deflect each other off course. Both in fact shoot at t1 and are on target. Their bullets miss each other by a whisker and pierce Ernst's heart simultaneously. Ernst dies." (p.284)

In this case, because both women are such good shots and have such a high chance of shooting it seems that probability of one shooting and thereby deflecting the other is so high as to negate the probability that Ernst will in fact be hit by her bullet. So, whether each shoots alone or whether both shoot together does not effect the probability of Ernst's death. As this is equally the case for both shooters, each of their shots seems to have a probability canceling effect. Barbara's shot has such a high probability of deflecting Claire's shot that Claire's shot now makes no probabilistic difference to Ernst's death and vice versa. Further because the two shooters are exactly equally matched in skill, if their bullets miss each other and both strike Ernst, their bullets will both strike Ernst at exactly the same time, it is impossible to say which has the greater claim to have caused Ernst's death. This seems to be a case of overdetermination in a probabilistic setting. It looks as though there are cases where the probability of an event is overdetermined.

Is it that problematic that examples of probabilities of events being overdetermined can be found though? As long as the overdetermination is not systematic it may not be problematic. We can after all give (at least putative) cases of actual overdetermination in the world such as the one given above. How does this impact the CEA? Could (P2) simply be replaced with (P2') that there be No Systematic Overdetermination of Probabilities? Questions could (hopefully) be asked about how realistic such a scenario

is in real life, however, the example goes to show that overdeterminism of probabilities is at least on the face of it possible in the probabilistic world. To recap the probabilistic analogue of the CEA is as follows;

(P1') Causal Closure of Physics

Every physical effect has a physical cause which is sufficient to fix its probability.

(P2') No Systematic Overdetermination of Probabilities

It is not usually the case that there are multiple sets of events that are minimally sufficient to fix the probability of a further event which exist simultaneously.

(P3') Non-Identity of the Physical and the Mental

There are physical states and mental states and these are not identical to each other.

(C') There is no Mental Causation

Mental states cannot be causes.

If the other premises of the CEA hold in probabilistic worlds, then it looks as though arguments against the existence or problematic nature of systematic overdetermination in the probabilistic world are required. This is because, if causal closure holds, then every probability will have a physical cause sufficient to fix its probability. Therefore, any mental cause which fixes a probability will be an overdetermining one. Therefore, assuming for now that causal closure holds, an argument that overdetermination is not problematic in the probabilistic case in order to prevent the CEA holding. I will now go through the arguments given above to see if they can work.

6.3 Overdetermination in a Probabilistic Setting

Do the above defences of overdetermination as being unproblematic still hold good in a probabilistic world? This is what I shall now discuss.

6.3.1 Sider Revisited

To recap, Sider's argument that overdetermination is no bad thing. He puts forward three different possible objections to overdetermination; (1) metaphysical incoherence, (2) coincidence and (3) epistemic doubts.

It seems to me that these arguments, particularly (2), can easily be transferred to the probabilistic setting. In fact, given that Sider explicitly mentions probabilistic theories of causation in (1), there is little need to 'translate'. However, I don't think it is strong enough to show that widespread overdetermination of probabilities is unproblematic. Just because no current theory *rules it out* in occasional cases isn't an argument in its favour in a widespread or systematic case. It is merely an argument against one argument against it.

As I said, (2) seems to be a strong argument to me in any deterministic or probabilistic world, as the systematicity which is required in order for overdetermination to be problematic on the CEA will never be un-law-governed whether we're considering overdetermining events or overdetermining the fixing of probabilities. An argument like (2) may have force in an utterly random world, but as I've stated this is not the kind

of indeterministic world I am interested in. In that world, given the utter randomness of causation, it is possible there would be no law-like relations between mental states and their underlying brain states. In that case systematic overdetermination would be a miraculous coincidence of the kind that Hume schooled us against. However, I stress again, that my focus is more narrow, and that on a probabilistic picture, there is no reason to think the connections can not be robust enough to counter coincidence.

However, although there is an intimate connection between the mental and the physical (supervenience), is it actually the correct kind of relationship to nullify worries about overdetermination? Sider is talking about a mereological relationship between a macro-entity (baseball) and its parts (atoms). This is not the relationship between the mental and the physical. So, in conclusion, I'm not sure that Sider's second defence of overdetermination applies to the case I am interested in.

Regarding (3), it can lead to doubt over the existence of some of the entities in question. That is, if both the baseball and its constituent parts both cause the window smashing, then we may have reasons to doubt the existence of the baseball. There are parsimony reasons to doubt their existence; they're epistemically redundant and by extension, metaphysically redundant. This is the 'mereological version' of the CEA. In the probabilistic case though we are not talking about entities causing things over and above their atomic parts, we are talking about probability fixing. However, it looks as though Sider's arguments still hold. The baseball and its atoms both fix the probability of the window smashing and Sider's arguments as to why this is unobjectionable still apply. So, the mental state and the physical state both fix the probability of the event coming about. The problem is, as I've argued above, the relationship between the mental and

the physical, unlike the relationship between the baseball and its atoms, may not be suitable to counter worries about overdetermination.

In conclusion, it is not clear that Sider's arguments hold. If they don't then the CEA is still an argument against mental causation. However, even if they do hold, it is well worth the time to examine the first premise of the CEA.

6.4 Chapter Conclusion

In conclusion, I have examined Sider's argument for why overdetermination need not be problematic. I then examined the possibility of overdetermination in probabilistic worlds and conclude that it is possible as the fixing of probabilities can be overdetermined. Lastly, I 'translated' the opening arguments defending overdetermination into probabilistic terms to see if they hold in that setting. I brought this chapter to an end by arguing that Sider's argument at least is perhaps not as strong in the probabilistic rather than a deterministic setting at that therefore overdetermination is still problematic in a probabilistic setting.

I will now move onto the first of the premises of the CEA: the causal closure of physics.

7

CAUSAL CLOSURE OF PHYSICS

I have stipulated that the mental and the physical are not identical. I have further come to the conclusion that overdetermination may still be problematic in a probabilistic world. I will assume for now then that overdetermination *is* still problematic in order to assess causal closure 'on its own merits'. This then, is the premise I will have to reject in order to overcome the CEA. If physics were not closed then it would be uncontroversial that non-physical causes could bring about physical effects without causing widespread and systematic overdetermination. There would be no overdetermination to object to. The CEA then, would not hold and there would be no objection to mental causation effecting the physical. So the question boils down to, does causal closure hold in a deterministic world?

To recap, the original premise is (P1) Causal closure of Physics which states that every

physical effect has a sufficient physical cause. The Probabilistic analogue (P1') Causal Closure of the Physics states that every physical effect has a physical cause which is sufficient to fix its probability. Combing this with non-identity and the probabilistic version of the no overdetermination premise (it is not usually the case that there are multiple sets of events that are minimally sufficient to fix the probability of a further event which exist simultaneously) would mean there is no mental causation even in probabilistic worlds.

7.1 What is Causal Closure?

In a sentence, the causal closure of physics (or of the physical) states that any physical event which has a cause, has a sufficient physical cause. However perhaps more than any other premise of the CEA the devil is in the detail, particularly when it comes to what you mean by 'physics' and 'physical'. Furthermore, care must be taken over exactly how to interpret this sentence as it can be read more strongly or more weakly (I will discuss this further in the next section).

To settle the first issue, I will take 'physics' to be anything under the purview of the physical sciences in the broadest possible sense. In other words I will take to be 'physical' not only all those things such as quarks, forces and fields which are posited by our best current theories of quantum mechanics but also all those things posited by our other best current theories such as General Relativity.

If the answer is so comparatively easy, then why it is so vital to pin down exactly

what is meant by 'physics'. The reason is to avoid circularity when applying causal closure to the CEA. If 'physics' were to be interpreted too broadly then potentially any phenomenon could count as falling under it and therefore could count as a cause. The causal closure would become trivial under these circumstances. In that case, it would be a simple matter to stipulate a physicalist world view whilst also maintaining the existence of mental causation. In that case the premise would not allow the CEA to go through as mental causation's not being an overdetermining cause would be compatible with completeness.

Of course, if one day cognitive science can explain the mental in physical terms¹, or if one day quantum science can give us the tools to explain mental phenomenon then this is what would happen. Until such a day though this strategy would be a bit of a cheat. Therefore, to stay true to the spirit of the world view the principle is supposed to represent we must restrict what we mean by physical.

This restriction cuts both ways however. It is important that, though we restrict what we mean by physics that this limits the phenomenon which causal closure covers, namely physical effects. If, on the other hand the phrasing was "any effect has a sufficient physical cause" this would be an overly stringent principle. It would also be to go beyond the available evidence precisely because the only evidence we have for the principle comes from the physical sciences. However, it is important to note this because it allows for mental causation even if only in the (scientifically redundant) case of a mental cause bringing about a mental effect². This is important because it

¹Say by demonstrating that mental states are composed of physical elements in a way we haven't yet discovered, or by positing new physical properties which we don't know about yet.

²It also remains silent on the issue of a physical cause bringing about a mental effect. Supervenience

means mental causation is not completely written off.

Before moving on to give arguments for against holding causal closure, I need to take a brief sidestep to discuss the related notion of the completeness of physics.

7.2 Causal Closure and Completeness of Physics

What is the relationship between causal closure and the thesis of the completeness of physics? The two are sometimes used interchangeably. Others consider causal closure to be the stronger formulation of completeness, see for example Marcus (2005) pgs 28-29. In such cases the completeness thesis states that physics doesn't need to look outside of itself in order to find sufficient causes for physical effects. Contrast to causal closure which is then interpreted as claiming that there can be no sufficient non-physical causes to physical effects. The kind of causal closure used in the CEA is the weaker sense, so this is the sense I will use from now on. Therefore, I will take causal closure and completeness to be interchangeable.

In the appendix to Papineau (2002) (pgs 232-256) he gives a brief history of the completeness of physics which I shall now even more briefly summarise. Although this serves as somewhat of a sidenote to the topic at hand, I will refer back to some of this in a later section.

Papineau gives Leibniz as the first example of someone who's conservation laws were such as to give completeness of physics. This is because his conservation of linear is not causation so that's not problematic, but we do think that physical causes can have mental effects.

momentum and kinetic energy together (plus the assumption of no action at a distance) are enough to close physics to any kind of mental ‘interference’. Newtonian physics took a different tack to Leibnizian physics in taking neither contact nor impact as his basic notion, but rather ‘impressed force’. Such impressed forces are much more permissive in their origins, thus opening up the possibility that mental forces could be among them. This is because, at least initially, while Newtonian physics, like Leibnizian physics conserved momentum, it did not conserve energy. Physics does not look so complete or closed anymore.

Latterly however, the conservation of energy did come to be considered a basic physical tenet. Experiments done by scientists such as James Joule led people to think that something was in fact conserved in some physical processes. In Joule’s case, it was heat and mechanical energy which he found to be equivalent. Such work in fact led to the creation of the universal theory of the conservation of energy. Herman von Helmholtz was the one to bring all the loose ends together. Luckily for history (if you will) Helmholtz had a reductionist project of his own, attempting to reduce biological phenomena to underlying non-biological laws. This goal led him to the assertion that energy must be conserved by all forces, even those, such as friction, which had traditionally not been considered conserved.

For this point on in history conservation was taken as given and the question shifted to what implications this had for the completeness of physics. Papineau cites what he calls “the argument from fundamental forces” (Papineau (2002) p.250) for leading scientists such as Helmholtz to hold their view that there were no animate forces meaning that conservation applied to only physical forces. Advances in the 1950s into

biochemical and neurophysiological forces made it more and more difficult to argue for extra-physical, animate forces. This addition of empirical evidence to the position of the conservation of energy left little room for those who did not want to hold the completeness of physics.

7.3 Arguments For Causal Closure

There is no ‘knock-down’ argument in favour of holding causal closure. However, the strongest argument which can be made for causal closure comes from physicalism and our general current scientific world view.

7.3.1 Physicalism

Perhaps the most obvious and compelling reason to think the causal closure must hold is that it so naturally fits with our current scientific and wider philosophical world view; namely physicalism and adherence to scientific practice. Science, and in particular, the natural sciences, has had a hugely successful track record. Take for example the massive advances in medical sciences³ or the achievement that was unifying electromagnetic theory⁴. The progress of science is unparalleled which can lead thinkers to place all their eggs in its basket.

The argument goes, because physics has operated well without reference to non-

³Perhaps interestingly these advances, in my own anecdotal experience, seem to have occurred more in physical health than mental health.

⁴See for example Maxwell (1863) as one step in this journey.

physical causation, that we should extrapolate from past experience to the logical conclusion that there is no non-physical causation. It is essentially an inductive argument from our best scientific experience. How strong is this as a defence of causal closure though?

One note should be made here however. Usually, rather than physicalism being used as a reason for holding causal closure, the opposite argument is made. That is, that causal closure and the success of physics are used as reasons to be physicalist. However, I believe both physicalism and causal closure both rest on potentially biased foundations. Therefore, I argue both are vulnerable in the same way in that both views could potentially stand a little scrutiny as I shall go on to argue in the next section.

7.4 Arguments Against Causal Closure

As I've discussed there is not any fully convincing argument in favour of physicalism and causal closure. Now I will discuss reasons for thinking causal closure may not hold; physicalist bias, the explanatory gap and the Knowledge Argument.

7.4.1 Physicalist Bias

Jones (2008) refers to causal closure as a “sort of ‘philosophical glue’ that binds a theory together” (p.181) rather than a straightforward summary of physical, scientific observations. By this he means this is that philosophers and scientists use causal closure as a kind of heuristic (although not a word he uses) to bind their observations

and theories into a cohesive system. There is no *direct* observation of causal closure, rather it is an inductive conclusion we have come to from our physical observations to date. Furthermore Vicente (2006)⁵ says:

“However, it [causal closure] is not a law that appears in physics textbooks. Where does it come from? Two answers spring to mind. First it can be said that it is not a physical law, but rather a methodological norm or principle that guides physicists in their research. Moreover, it can be defended that it is a norm well supported by inductive evidence... Second it may be said, although causal closure P [causal closure of physics] is not strictly a truth of physics, it is supported by, or depends on, actual laws of physics.” (pgs 150-151)

If the best that can be said for causal closure is that it coheres with a wider world view, or is a useful heuristic norm, then it might not be on the sturdiest ground. As genealogist theorists such as Michael Foucault⁶ have argued for decades, if you can track the history of an idea and find it originates in bias then that theory, at best, should be thoroughly examined. While I would not go so far as some genealogist thinkers in saying because we can trace the origin of the idea to human bias we should abandon the theory altogether, I think it definitely shows that more argumentation needs to be put forward as to why we should hold to this principle. As I have shown through tracing the history of the completeness of physics from Papineau’s appendix, this principle can

⁵Vicente’s paper is a defence of causal closure and therefore should put forward the best case for it holding.

⁶See *The History of Madness* (Foucault & Khalfa (2006)), *The Birth of the Clinic* (Foucault (2010)) and *Discipline and Punish* (Foucault (1991)) to name but a few examples.

indeed be traced back to biases and what may be colloquially termed as ‘physics envy’. Take Helmholtz as an example. Papineau notes that Helmholtz’s “physiological context undoubtedly played a fundamental role in Helmholtz’s articulation of a universal principle of the conservation of energy” (Papineau (2002) p.246). Further Papineau says it’s “likely that it was Helmholtz’s specific combination of physiological interests and sophisticated physical understanding that precipitated his crucial synthesis of the different strands of research feeding into the conservation of energy” (Papineau (2002) p.247). All this is to make just one example of how historical coincidence can lead the course of intellectual history in a particular direction. Had Helmholtz had a different particular history then the course of causal closure may have been derailed and may not have taken its predominant spot in our philosophical world view. Of course, this line of reasoning is speculative, but it does lead me to question the extent to which contingencies lead to philosophical positions, causal closure in particular, that we hold dear.

At the very least, raising awareness of these contingencies leads me to think we must be very careful in examining why we hold the views that we do to make sure we don’t place more faith in them than the evidence would allow. In Helmholtz’s case Papineau himself asks “how far was this almost immediate agreement on the conservation of energy dictated by the strength of evidence rather than by intellectual fashion” (Papineau (2002) p.250). In Helmholtz’s case the evidence was strong, is this the case for causal closure also? Ultimately every position boils down to intuition and assumptions such as ‘the argument from fundamental forces’. Uncovering these biases to see why and how they could be effecting our views leading us to accept some conclusions over

others can surely never be wasted work.

So far the argument has been all negative, reasons to knock our belief in causal closure. Now I will offer a more positive argument for thinking causal closure may be lacking in the form of Levine's *Explanatory Gap* and Frank Jackson's *Knowledge Argument*.

7.4.2 The Explanatory Gap and Knowledge Argument

Levine's *Explanatory Gap* should give pause for thought to those who think physics can ever be complete. All that closure states is that there is no non-physical cause to physical effects. It is not incompatible with non-physical causation of non-physical events nor with physical causation of non-physical events. However, as I've argued, we have reasons such as the MMI to think that there are mental causes of physical effects. The explanatory gap implies that these mental causes are non-physical. Therefore, we have reason to think that there are non-physical causes of physical effects and causal closure is wrong. Furthermore, these considerations make it seem all the more implausible that the physical world is entirely causally isolated from the mental world.

As I've already discussed this argument earlier I will not go into more detail again except insofar as to say that the argument goes that physicalism not only does, but must, leave something out of the picture of our world. And thus, in so far as we take the coherence of the overall physicalist picture to be evidence for causal closure, causal closure is therefore wrong.

One last argument which, again, I have already mentioned, may give us reason to

question causal closure; the *Knowledge Argument*. Again, given we have reason to think that there are mental states (which the knowledge argument suggests are not identical to physical states) which bring about physical events, we have reason to think causal closure is wrong.

If, as the argument claims, physicalism is not the final word in explaining the world, then what consequences does this have for causal closure? Arguments against the knowledge argument, such as the argument that what Mary gains on leaving the room is not propositional knowledge ('knowledge that'), rather it is another kind of knowledge such as 'knowledge how'⁷ are therefore not problematic for physicalists. Further, appeal to the potential of future physics to explain consciousness and its phenomenal aspect would also put paid to the knowledge argument. If qualia could be explained by physical theory then she would know it inside the room and would not learn anything upon leaving the room.

7.5 Probabilistic Causal Closure

I have spent this chapter discussing the case for and against causal closure. In my next and final substantive chapter I will put forward my most important argument; my reason for thinking causal closure, and therefore the CEA, does not hold in a probabilistic world.

⁷See for a few examples, Ryle (1988), Lewis (1983), Lewis (1990) and Conee (1994)

8

EVALUATING THE PROBABILISTIC CEA

“

”If it’s a story I’m telling, then I have control over the ending

A Handmaid’s Tale - Atwood (1995)

8.1 Causal Closure and Indeterminism

The original version of the CEA may still hold in deterministic worlds. Although, as I have just argued, there may be some reasons for thinking causal closure, doesn’t hold (meaning the CEA would be unsound) even in deterministic worlds. However, whether or not causal closure holds in deterministic worlds, we have got good reasons

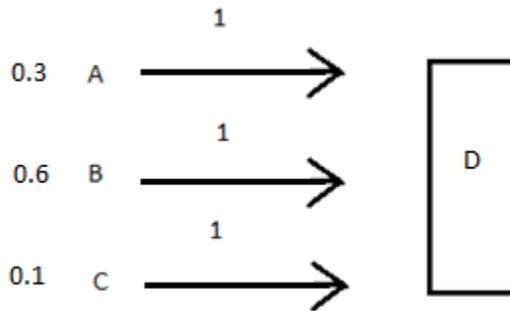
for thinking that the premises of *this* original version of the CEA are not true in our probabilistic world. Therefore, I turned to analysing the probabilistic analogue version of the CEA, whose premises we have better reason to be true in our world.

However, I think my argument against causal closure is much stronger in probabilistic worlds. This is the vital distinction my argument rests on. If I am right and causal closure is false in probabilistic worlds like ours, then the analogue CEA will be unsound. Given that I assume we live in a probabilistic, rather than deterministic world, we must now look to the probabilistic analogue to the CEA. In such a world causal closure means that every physical effect has a physical cause which is sufficient to fix its probability. This means that causal closure is no longer about bringing an event about or not, but rather about bringing a probability about or not.

How then could the CEA could be overcome on an indeterministic picture? One way would be to claim (P1') is false and deny causal closure. Therefore because physical causes do not always guarantee their effects, nor fix their probabilities, there may be room for the mental to be doing some work. However, it is possible that even in probabilistic worlds the mental doesn't do any causal work. In that case, while it is merely probabilistic which exact physical cause actually brings about the event, it is nonetheless guaranteed that only a physical one will.

But the picture is complicated. There are different places along the causal chain where mental causation could enter. Examine figure 1.

Figure 1.

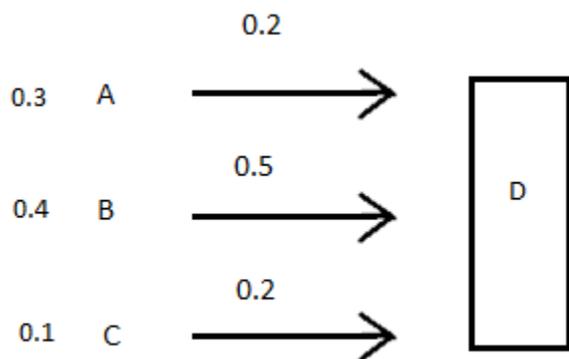


Say that A, B and C are all potential physical causes of an event¹ represented by the rectangle labeled D which is conditional upon either A, B or C (each of which is mutually exclusive and jointly exhaustive) occurring and each of which is sufficient for D's occurring. However, it is unclear in this probabilistically indeterministic case which of the three physical causes will be the one that does the causing. I have assumed in this diagram that each will on its own be sufficient². That is what the 1s above the arrows represent, D is certain to occur on the condition that either A, B or C occurs. Because A, B and C are mutually exclusive, should any of these physical cases occur they are guaranteed to bring about the event D; it's just not certain in a deterministic sense which one will occur. A has a 0.3 chance of occurring, B a 0.6 chance and C a 0.1 chance. Now take the slightly adapted figure 2.

Figure 2.

¹Or rather the causes of a kind of event represented by D for simplicity.

²It is strange to show a probabilistic diagram with local determinism in it. The reason I have done this is purely to highlight the differences with Figure 2 in order to emphasise my point. I will drop my assumption of local determinism in Figure 2.



In this case further indeterminism has been added³. It is now not certain that if any of the physical causes were to come about that they would guarantee that D happens⁴. A now has a 0.3 chance of occurring and a there is a 0.2 chance of D occurring given A. B still has the best chances of actually occurring (0.4). and D now has a 0.5 chance of occurring given B. C is still the underdog with only the very small 0.1 chance that it occurs and D has a 0.2 chance of occurring given C.

In Figure 2. there is not sufficiency of the physical cause. If the world is no longer running strictly according to deterministic rules then it looks as though events could potentially be underdetermined by their physical causes; in other words, causal closure does not hold. It is no longer the case that any cause A, B or C will definitely cause D. And, perhaps, what can be said about bringing the event about can also be said about the fixing of the probabilities.

³In other words, I have dropped the artificial simplification of Figure 1.

⁴It's important to note that this seems to be a much better representation of how the world actually is than the scenario represented in Figure 1.

Take Figure 2. There's only a 0.3 chance that physical cause A will occur⁵ and given A occurs, D then has a 0.2 chance of occurring ($P(D|A)=0.2$). Say event D does come to pass, then it seems as though there is room for mental state e to contribute to raising the probability. Or, it could be that some physical cause confers a probability onto an effect but either also requires a mental cause to fully fix the probability, or at least allows room for mental states to effect the probability of the event occurring. This would mean that the probability of D occurring is not actually fixed purely by the physical cause A which means that causal closure does not hold. It is also the case that the mental state is not an overdetermining cause nor an overdeterminer of the probability of the physical effect. Contrast this with the deterministic case in which all physical events have probability 1 conditional on their physical causes so there is no room for mental states to 'top up' the probability (at least not without being an overdeterminer).

To reiterate; typically in probabilistic worlds, the probability of a physical effect given its physical cause will be less than 1. This is represented in the following inequality where E is the physical event and P is the putative physical cause; ($P(E|P) > 1$). This leaves room for the mental to 'top up' the probability. Say M is the putative mental cause. Then the inequality $P(E|P\&M) > P(E|P)$ could now hold. However, in deterministic worlds, the probability of a physical effect given its putative physical cause will typically be 1; ($P(E|P) = 1$). There is now no room for the mental to do any non-overdetermining work because the probability cannot be greater than 1.

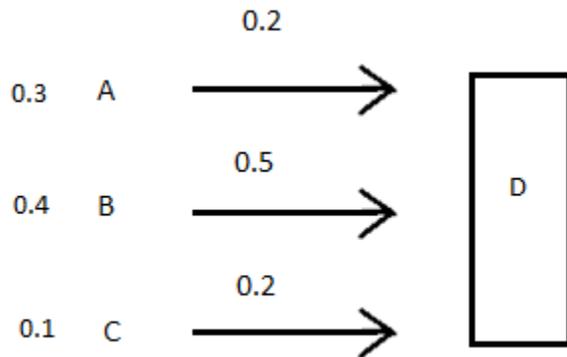
⁵As I have assumed there is a 0.3 probability of A occurring in purely physical terms, perhaps there is also room for mental states to be causally contributing to causing A to occur rather than B or C. In other words, perhaps mental causation could also be occurring further up the causal chain.

There is an important distinction to be made between events and the probabilities of events coming about. The third premise of the CEA states that there be no systematic overdetermination of events. Does that same premise apply to overdetermination of probabilities? If mental states help to fix the probability of the physical event occurring then are they in fact causally efficacious. For example, the mental state of desiring a coffee may raise the probability of my getting a coffee as opposed to the mere physical neural firings. Or, it may help shape the exact manner in which I go about getting the coffee. My mental states may be effecting the exact way I get my coffee, for example by raising the probability that I rush to the coffee shop as opposed to walking slowly. Actual me may have a different probability distribution over collecting coffee in any particular way to 'zombie me', as the latter has only physical causes to fix the probabilities.

If it is the case that mental states can contribute to fixing the probability of a physical state, over and above the contribution made by the brain states, causing a physical effect then it looks as though we have a way mental causation can exist contra the causal exclusion argument.

So, how would models of mental causation work in a probabilistic setting? Let's examine figure 2 again:

Figure 2.



The diagram indicates that there is no one sufficient physical cause of the event D's occurring. Let's say that the event occurs and is caused by physical event A. If causation was deterministic then the probability of D occurring given A would be 1. But in this probabilistic setting A has a 0.3 chance of occurring and, given it occurs, D now has a 0.2 chance of occurring. Perhaps if we add the mental into the picture we can see how it can be causally efficacious by raising the probability that the event will occur. I will now set out the inequalities to make my point clearer. Call the event E and the physical cause A. We can demonstrate that physical cause A is a cause by showing that the following holds:

(i) $P(E | A) > P(E | \neg A)$

All this inequality says is just the probability of the event occurring is higher given the physical cause than the absence of the physical cause. Now call the mental state M. If the following inequality holds then this would be a sign of the causal efficacy of

mental state M. If the inequality holds when we hold A fixed, then this implies that it is mental state M that is making the difference.

(ii) $P(E | A \& M) > P(E | A \& \neg M)$

In prose, the probability of the event occurring is higher given the presence of the physical and mental state is higher than given only the physical state. I believe the same could be said of other special science properties which would mean that special science properties would count as causes of physical events. Take as an empirical example the collapse of the wave function on certain orthodox interpretations of quantum mechanics. On such views the wave function collapses on measurement. The wave collapse is a physical micro-level event. Measurement is a macro-level event. This would appear to be an example of a higher-order property causing a fundamental physical event. If the measurement was an observation made a person as in Schrödinger's (1935) thought experiment this would be a case of mental causation. So, although, much more work must be done before this conclusion can become any more substantial, at least it is a good step in the right direction.

8.2 Unresolved Issues

There are of course many questions my work will raise which will not be answered in this thesis. I will comment on three of them now; how far my argument pushes me from physicalism, the potentially strange conclusion my argument has for mental

causation in deterministic worlds, and another potential avenue for refuting the CEA.

8.2.1 Departing from Physicalism

Given that I argue closure (and completeness) do not hold in probabilistic worlds, what consequences does this have for physicalism? Or, in other words, how far do we have to depart from physicalism given my arguments? This is an important question because it cuts to the heart of how the mental and the physical interact and therefore how the mechanism behind mental states raising the probability of physical events coming about actually works.

I don't want to focus too much on this issue due to space reasons. Again, it is possible that in future physics will be able to 'fill in the gaps' and fully explain the mental in physical terms. However, the explanatory gap, knowledge argument and so on, all lead me to think it is in principle not going to be the case. It appears on the face of it, to suggest that some kinds of property dualism is the case. But, where exactly my argument leaves physicalism is an interesting question for future work.

8.2.2 Peculiar Conclusion?

I have already mentioned a problem which could be raised against my argument. That is, it seems odd and counterintuitive to say that the existence of mental causation depends on whether the world is deterministic or probabilistic. As I have argued, there is more reason to think that the deterministic CEA holds in deterministic worlds than

we have to think the probabilistic analogue holds in probabilistic worlds. This would seem to suggest that mental causation exists in probabilistic worlds but might not in deterministic worlds. As I have assumed we live in a probabilistic world (on the basis of certain scientific evidence) our experience cannot contradict the notion that there is no mental causation in deterministic worlds.

What would mental life be like in a deterministic world then? Perhaps there would be no experience of free will such as we have if we examine our inner mental lives. Or perhaps there would be but this would be strictly illusory.

Although this conclusion does on the face of it seem counterintuitive, I do not think this is the strongest argument against it. This is a bullet I am willing to bite as I do not think we have any direct evidence against it.

8.2.3 Invalid or Unsound

There are two ways I could try to refute the CEA. I could, as I have been doing, deny that causal closure holds in probabilistic worlds, and therefore that the CEA is valid but unsound. However, I wonder if there could be a potential second route to denying the CEA. This would be to claim that the argument is invalid. This could be done by suggesting that causal closure holds even in probabilistic worlds, but that there is somehow still a way for mental causation to be doing some work without overdetermining effects (assuming of course that (P2') still holds). In other words, that physical causes are always sufficient for fixing the probability of physical effects but somehow mental causes also make a difference in a non-overdetermining way. I have

not looked into this possibility, as I think it is much less promising than the route I have taken, but I note it here as potentially interesting thought.

9

FREE WILL

An important note to make now concerns the debate around free will and how it connects with my work. The philosophical literature about determinism in particular is usually couched in talk about free will. It might therefore be expected that I would feature free will more heavily in my thesis. However, my main focus has been mental causation. There are two reasons I have decided to include a chapter on free will.

My first reason is because the free will debate is both interesting and important. Both in its own right and as an area of philosophy which has direct applications to real life, for example, I think it is a debate which can bear directly on law and ethics. I touch on this below. To reiterate however, free will is not the main focus of my current work and I will not be directly interested in it for these reasons. Moral responsibility has never been my area of expertise.

My second reason is that I see it as a natural extension to my work. It is plausible to think that for an individual to have free will, mental causation must exist. I explain my reasoning for this further below.

9.1 What is Free Will?

Free will is often defined as "the ability to do otherwise". The argument (from a strictly non-compatibilist stance) goes that there cannot be free will if the world is deterministic as there would be no ability to do otherwise given that everything is fixed by an initial set of conditions and the laws of nature. Kadri Vivelin summarizes the situation;

“If determinism is true, we are never able to do otherwise. If we are never able to do otherwise, we have no free will. If we are never able to do otherwise, we are never morally responsible. Therefore, if determinism true, we have no free will and are never morally responsible.” (Vihvelin (2013) p.1)

As I have stated I am not directly interested in the debate on free will, at least not in so far as it is concerned with moral philosophy. The arguments in moral philosophy go along the lines that you cannot be held morally responsible for an action which you could not have done otherwise. In order to be held morally responsible for your actions therefore requires you to have free will. And to reiterate, free will requires determinism to be false (again from a non-compatibilist stance). So, moral responsibility requires, at a minimum, determinism to be false. This debate has spawned various types of

compatibilist arguments which, through various different means, argue that free will or moral responsibility is compatible with determinism¹.

As stated however, I am not so interested in the moral implications of the free will debate although there will no doubt be moral implications from my work which will be interesting. I will be more interested in free will insofar as it relates to mental causation. In so doing I will be guilty of a phenomenon Sara Bernstein and Jessica Wilson describe; “discussions of mental causation have neglected the efficacy of events of free choosing, focusing almost exclusively on the efficacy of qualitative and intentional mental events (pains and colour experiences, beliefs and desires) for which freedom is not an issue” (Bernstein & Wilson (2016) p.1).

I think the definition of free will given above is actually inadequate to capture what we mean when we talk about having free will. It is not sufficient to have more than one option available to you, at least on a weaker reading of the word ‘available’. Free will as we talk about it also requires that you be able to choose which of the options you take. If you read ‘available’ to include this element of the ability to actively choose, then the above definition is adequate. If you read ‘available’ to merely mean exist then

¹See for example accounts which claim as long as you can endorse the decision, it does not matter that you cannot have done otherwise. Dennett’s example is one of my favourites;

“‘Here I stand,’ Luther said. ‘I can do no other.’ Luther claimed that he could do no other, that his conscience made it *impossible* for him to recant. He might, of course, have been wrong, or have been deliberately overstating the truth. But even if he was - perhaps especially if he was - his declaration is testimony to the fact that we simply do not exempt someone from blame or praise for an act because we think he could do no other. Whatever Luther was doing, he was not trying to duck responsibility.” (Dennett (2015) p.145)

This seems to merely push the problem back a step further, as it would seem that you would have no control over whether you endorsed your action or not. Of course, you could then endorse your endorsement but this would lead to a vicious infinite regress of endorsements.

it is not. For the sake of clarity, and because I think it is the more natural reading, I will use the second reading and keep the 'choosing requirement' separate. To have free will then requires both options to choose from and the ability to make an active choice as to which to go for.

To clarify, I will illustrate with an example. Say I have two books, *Mind in a Physical World*² and *Physicalism, Or Something Near Enough*³. In a determined world I may as well not have two books as I will have no choice over which I read first. The laws of nature and all previous states of the world will leave me no choice but to read *Mind in a Physical World*. In an indeterminated world however, I still have no free will if I only have access to one book. Unless the library has a copy of *Physicalism or Something Near Enough* then there is no way I can choose to read it. Luckily though I have managed to obtain a copy of each. It is not enough that I have two books available to me if I cannot actually pick which one I read. There has to be some mental event (or physical event with a mental property in virtue of which the choice is made) of mine that makes the difference in which book I read. Mental causation is therefore necessary (although not sufficient) for free will. This is where I really become interested in the free will debate.

²Kim (1998)

³Kim (2008)

9.2 Free Will and the CEA

There is an obvious relationship between free will and mental causation and an obvious parallel between the debates on free will and determinism and the CEA. If determinism is the case then, putting compatibilist arguments to one side⁴, free will cannot be true. If the CEA is right, then mental causation does not exist. But the scenarios are, in my opinion, even more linked than this. It is my argument that the CEA does not go through in an indeterministic world. So, if the CEA holds *and determinism* is true, then mental causation of physical effects does not exist. Therefore, free will and mental causation stand or fall together to a large extent. While not every instance of mental causation will be related to an instance of free will, every instance of free will will involve mental causation. What I have to say on the issue of the efficacy of mental causation will therefore bear on the debates surrounding free will which may make my work of interest to a wider group of people. Bernstein and Wilson describe what they call the *The General Problem of Mental Causation*⁵.

“*The General Problem of Mental Causation*. How can a mental event *M* of a given type be efficacious vis-a-vis an event *E* in virtue of being the type of mental event it is, given that there is reason to think that events of *M*'s type are causally irrelevant to the production of events of *E*'s type?”
(Bernstein & Wilson (2016) p.5).

⁴I will do these partially because I have never found any compatibilist arguments particularly convincing but mainly for reasons of space.

⁵Following (Horgan (1989))

Rather than put forward a positive argument for the existence of mental causation (or the reality of free will) I have defended it against an argument for its non-existence. If I am successful then an alternative argument would need to be put forward why probabilistic mental causation does not exist. Therefore, I have indirectly also defended free will from one argument against its non-existence.

10

CONCLUSION

Given my examination of the premises of the CEA and my translating them into probabilistic terms, does the CEA still go through, thus ruling out that the mental can have any physical effect? To reiterate, if it no longer holds, then this does not necessarily mean that mental causation exists, but it does mean that one counter argument to it does not hold. So while more would remain to be done, it would be a step in the right direction.

To conclude I will recap the two versions of the CEA; the deterministic version and the probabilistic analogue. I will then briefly recap my main argument and end by drawing together final conclusions.

To recap the argument goes as follows; Deterministic CEA:

(P1) Causal Closure of Physics

Every physical effect has a sufficient physical cause.

(P2) No Systematic Overdetermination

It is not usually the case that there are multiple minimally sufficient causes of any given event which exist simultaneously.

(P3) Non-Identity of the Physical and the Mental

There are physical states and mental states and these are not identical to each other.

(C) There is no Mental Causation

Mental states cannot be causes.

I posit that this is the version of the CEA which is most discussed. As it is, while I have raised some cause for concern, it may well go through and leave no room for the efficacy of mental causation. However, the second version of the CEA on the other hand may be a different story as I think we have even more reason to doubt it holds.

Probabilistic Analogue CEA:

(P1') Causal Closure of Physics

Every physical effect has a physical cause which is sufficient to fix its probability.

(P2') No Systematic Overdetermination of Probabilities

It is not usually the case that there are multiple sets of events that are minimally sufficient to fix the probability of a further event which exist simultaneously.

(P3') Non-Identity of the Physical and the Mental

There are physical states and mental states and these are not identical to each other.

(C') There is no Mental Causation

Mental states cannot be causes.

There are good reasons for believing that this is the version of the CEA we should be focused on, for the simple reason that this is the version whose premises seem more likely to be true in the world we live in. It is this version therefore, that must be analysed.

In conclusion then, I argue that I have shown the CEA to be ineffective at proving mental causation can not cause physical effects in so far as probabilistic worlds like ours go. I have argued that the mental and the physical are not identical and that there may still be reasons to think overdetermination is problematic. So, I have concluded that the first premise of the analogue version is wrong. This is because, I argue, causal closure does not hold in probabilistic worlds. Whether or not the original CEA is valid, it is fairly clear that it doesn't apply in our world. In other words the CEA may be valid but unsound in probabilistic worlds. However, I think there is good reason to think that the probabilistic analogue of the CEA does not go through. Where does this leave the causal efficacy of mental states? It seems to suggest that mental causation exists. At the very least, one argument against it has been refuted, which means there may still be reason to think that mental states can bring about physical effects at least in probabilistic worlds.

BIBLIOGRAPHY

- Albert, D. & Loewer, B. (1988), 'Interpreting the many worlds interpretation', *Synthese* 77(2), 195–213.
- Atwood, M. E. (1995), *The handmaid's tale*, reprint edn, Virago Press, London. OCLC: 246673445.
- Baumgartner, M. & Glynn, L. (2013), 'Introduction to Special Issue on 'Actual Causation'', *Erkenntnis* (1975-) 78, 1–8.
URL: <http://www.jstor.org/stable/24010948>
- Beissner, F., Brandau, A., Henke, C., Felden, L., Baumgärtner, U., Treede, R.-D., Oertel, B. G. & Lötsch, J. (2010), 'Quick Discrimination of Adelta and C Fiber Mediated Pain Based on Three Verbal Descriptors', *PLoS ONE* 5(9).
URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2944851/>
- Bennett, J. (1987), 'Event Causation: The Counterfactual Analysis', *Philosophical Perspectives* 1, 367–386.
- Bernstein, S. & Wilson, J. (2016), 'Free Will and Mental Quausation', *Journal of the*

BIBLIOGRAPHY

American Philosophical Association 2, 310–331.

Carroll, L. (1856), *Alice's Adventures in Wonderland*, Macmillan. Google-Books-ID: Y7sOAAAAIAAJ.

Casati, R. & Varzi, A. (2015), Events, in E. N. Zalta, ed., '*The Stanford Encyclopedia of Philosophy*', winter 2015 edn, Metaphysics Research Lab, Stanford University.

URL: <https://plato.stanford.edu/archives/win2015/entries/events/>

Churchland, P. M. (1981), 'Eliminative Materialism and the Propositional Attitudes', *The Journal of Philosophy* 78(2), 67–90.

URL: <http://www.jstor.org/stable/2025900>

Conee, E. (1994), 'Phenomenal Knowledge', *Australasian Journal of Philosophy* 72(2), 136–150.

Courage, K. H. (2013), 'How the Freaky Octopus Can Help Us Understand the Human Brain'.

URL: <https://www.wired.com/2013/10/how-the-freaky-octopus-can-help-us-understand-the-human-brain/>

Curtis, V., Aunger, R. & Rabie, T. (2004), 'Evidence that disgust evolved to protect from risk of disease.', *Proceedings of the Royal Society B: Biological Sciences* 271(Suppl 4), S131–S133.

URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1810028/>

Davidson, D. (2001a), *Essays on Actions and Events*, Oxford University Press.

BIBLIOGRAPHY

- Davidson, D. (2001*b*), *Mental Events*, in *'Essays on Actions and Events'*, Oxford University Press, pp. 207–228.
- Dennett, D. C. (2015), *Elbow room: the varieties of free will worth wanting*, new edition edn, MIT Press, Cambridge, Massachusetts ; London, England.
- DeRubeis, R. J., Siegle, G. J. & Hollon, S. D. (2008), 'Cognitive therapy vs. medications for depression: Treatment outcomes and neural mechanisms', *Nature reviews. Neuroscience* **9**(10), 788–796.
URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2748674/>
- Descartes, R. (2013), *Meditations on first philosophy: with selections from the Objections and Replies : a Latin-English edition / René Descartes ; edited and translated with textual and philosophical introductions by John Cottingham.*, Cambridge University Press, Cambridge.
- Dowe, P. (1995), 'Causality and Conserved Quantities: A Reply to Salmon', *Philosophy of Science* **62**(2), 321–333.
URL: <http://www.jstor.org/stable/188437>
- Dowe, P. (2001), 'A Counterfactual Theory of Prevention and 'Causation' by Omission', *Australasian Journal of Philosophy* **79**(2), 216–226.
URL: <http://www-tandfonline-com.libproxy.ucl.ac.uk/doi/abs/10.1080/713659223>
- Dowe, P. (2008), Causal Processes, in E. N. Zalta, ed., *'The Stanford Encyclopedia of Philosophy'*, fall 2008 edn, Metaphysics Research Lab, Stanford University.
URL: <https://plato.stanford.edu/archives/fall2008/entries/causation-process/>

BIBLIOGRAPHY

- Eagle, A. (2016), Chance versus Randomness, in E. N. Zalta, ed., '*The Stanford Encyclopedia of Philosophy*', fall 2016 edn, Metaphysics Research Lab, Stanford University.
URL: <http://plato.stanford.edu/archives/fall2016/entries/chance-randomness/>
- Eliot, T. S. (2010), *The Waste Land and Other Poems*, Broadview Press. Google-Books-ID: tUVrhzxW3xIC.
- Fair, D. (1979), 'Causation and the Flow of Energy', *Erkenntnis* (1975-) **14**(3), 219–250.
URL: <http://www.jstor.org/stable/20010665>
- Fenton-Glynn, L. (2009), A Probabilistic Analysis of Causation, DPhil, University of Oxford, Oxford.
- Fenton Glynn, L. (2016), 'A Proposed Probabilistic Extension of the Halpern and Pearl Definition of 'Actual Cause'', *The British Journal for the Philosophy of Science* (2016) (In press). .
- Fodor, J. (1974), 'Special sciences (or: The disunity of science as a working hypothesis)', *Synthese* **28**(2), 97–115.
- Fodor, J. (1997), 'Special Sciences: Still Autonomous after All these Years', *Noûs* **31**, 149–163.
- Foucault, M. (1991), *Discipline and punish: the birth of the prison*, Penguin social sciences, reprint edn, Penguin Books, London. OCLC: 45870209.
- Foucault, M. (2010), *The birth of the clinic: an archaeology of medical perception*, Routledge classics, 1. publ., reprinted edn, Routledge, London. OCLC: 845118786.

BIBLIOGRAPHY

- Foucault, M. & Khalfa, J. (2006), *History of madness*, Routledge, New York.
- G. H. von Wright (Georg Henrik) (1971), *Explanation and understanding.*, Contemporary philosophy, Cornell University Press, Ithaca, NY.
- Gibbon, L. G. & Bold, V. (2001), *Smeddum: a Lewis Grassic Gibbon anthology*, number 97 in 'Canongate classics', Canongate Books, Edinburgh. OCLC: ocm42274394.
- Goldacre, B. (2009), *Bad science*, fourth estate paperback edition. edn, Fourth Estate, London.
- Goodman, N. (1983), *Fact, Fiction, and Forecast*, Harvard University Press.
- Goodman, N. (2000), The New Riddle of Induction, in S. Bernecker & F. I. Dretske, eds, 'Knowledge: Readings in Contemporary Epistemology', OUP Oxford.
- Hall, N. (2004), Two Concepts of Causation, in J. Collins, N. Hall & L. Paul, eds, 'Causation and Counterfactuals', The MIT Press, pp. 225–276.
- Hart, H. L. A. & Honoré, T. (1985), *Causation in the law / by H.L.A. Hart and Tony Honoré.*, Oxford scholarship online Y, 2nd ed. edn, Clarendon, Oxford.
URL: <http://dx.doi.org/10.1093/acprof:oso/9780198254744.001.0001>
- Heisenberg, W. (1927), 'Äøeber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik', *Zeitschrift fur Physik* **43**, 172–198.
URL: <http://adsabs.harvard.edu/abs/1927ZPhy...43..172H>
- Hitchcock, C. (2001), 'A Tale of Two Effects', *The Philosophical Review* **110**(3), 361–396.
URL: <http://www.jstor.org/stable/2693649>

BIBLIOGRAPHY

- Hitchcock, C. (2004), Do All and Only Causes Raise the Probabilities of Effects?, in J. Collins, E. J. Hall & L. A. Paul, eds, *'Causation and Counterfactuals'*, MIT Press.
- Hitchcock, C. (2012), Probabilistic Causation, in E. N. Zalta, ed., *'The Stanford Encyclopedia of Philosophy'*, winter 2012 edn, Metaphysics Research Lab, Stanford University.
URL: <http://plato.stanford.edu/archives/win2012/entries/causation-probabilistic/>
- Hitchcock, C. & Knobe, J. (2009), 'Cause and Norm', *The Journal of Philosophy* **106**(11), 587–612.
- Hölldobler, B. (2011), *The leafcutter ants: civilization by instinct / Bert Hölldobler and Edward O. Wilson.*, Norton, New York.
- Hoefer, C. (2016), Causal Determinism, in E. N. Zalta, ed., *'The Stanford Encyclopedia of Philosophy'*, spring 2016 edn, Metaphysics Research Lab, Stanford University.
URL: <http://plato.stanford.edu/archives/spr2016/entries/determinism-causal/>
- Horgan, T. E. (1989), 'Mental Quausation', *Philosophical Perspectives* **3**(n/a), 47–74.
- Hume, D. (1988), *An enquiry concerning human understanding*, Prometheus Books, Buffalo, N.Y. OCLC: 19868862.
- Hume, D. (2009), *A Treatise of Human Nature*, Oxford philosophical texts, Oxford Univ. Press, Oxford. OCLC: 934530118.
- Hunt, G. R. & Gray, R. D. (2003), 'Diversification and cumulative evolution in New Caledonian crow tool manufacture', *Proceedings of the Royal Society of London B: Biological Sciences* **270**(1517), 867–874.
URL: <http://rspb.royalsocietypublishing.org/content/270/1517/867>

BIBLIOGRAPHY

- Jones, K. (2008), 'The Causal Closure of Physics: An Explanation and Critique', *World Futures* 64(3), 179–186.
- Kam-Hansen, S., Jakubowski, M., Kelley, J. M., Kirsch, I., Hoaglin, D. C., Kaptchuk, T. J. & Burstein, R. (2014), 'Altered placebo and drug labeling changes the outcome of episodic migraine attacks', *Science Translational Medicine* 6(218), 218ra5.
- Kaptchuk, T. J., Friedlander, E., Kelley, J. M., Sanchez, M. N., Kokkotou, E., Singer, J. P., Kowalczykowski, M., Miller, F. G., Kirsch, I. & Lembo, A. J. (2010), 'Placebos without Deception: A Randomized Controlled Trial in Irritable Bowel Syndrome', *PLOS ONE* 5(12), e15591.
URL: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0015591>
- Kim, J. (1966), 'On the Psycho-Physical Identity Theory', *American Philosophical Quarterly* 3(3), 227–235.
- Kim, J. (1992), 'Multiple Realization and the Metaphysics of Reduction', *Philosophy and Phenomenological Research* 52(1), 1–26.
- Kim, J. (1993), *Supervenience and mind: selected philosophical essays*, Cambridge studies in philosophy, Cambridge University Press, Cambridge.
- Kim, J. (1998), *Mind in a physical world: an essay on the mind-body problem and mental causation*, Representation and mind, MIT Press, Cambridge, Mass.
- Kim, J. (2008), *Physicalism, or something near enough*, Princeton monographs in philosophy, 3rd print., and 1st paperback print edn, Princeton Univ. Press, Princeton.
OCLC: 611751299.

BIBLIOGRAPHY

- Knobe, J. & Fraser, B. (2008), Causal Judgment and Moral Judgment: Two Experiments, *in* W. Sinnott-Armstrong, ed., *'Moral Psychology'*, MIT Press.
- Kripke, S. A. (1980), *Naming and necessity*, Library of philosophy and logic Y, Blackwell, Oxford.
- Kroedel, T. (2008), 'Mental causation as multiple causation', *Philosophical Studies* **139**(1), 125–143.
- Kvart, I. (2004), Causation: Probabilistic and Counterfactual Analyses, *in* N. Hall, L. Paul & J. Collins, eds, *'Causation and Counterfactuals'*, MIT Press, Cambridge, Mass., pp. 359–387.
- Lemmon, E. (1996), Comments on D. Davidson's "The Logical Form of Action Sentences", *in* N. Rescher, ed., *'The Logic of Decision and Action'*, University of Pittsburgh Press, Pittsburgh, pp. 96–103.
- Levine, J. (1983), 'Materialism and Qualia: The Explanatory Gap', *Pacific Philosophical Quarterly* **64**(October), 354–61.
- Lewis, D. (1973), 'Causation', *The Journal of Philosophy* **70**(17), 556–567.
- Lewis, D. (1983), 'Postscript to "Mad Pain and Martian Pain"', *Philosophical Papers* **12**, 122–133.
- Lewis, D. (1986a), Postscripts to 'Causation', *in* D. Lewis, ed., *'Philosophical Papers Vol. II'*, Oxford University Press.

BIBLIOGRAPHY

- Lewis, D. (1990), What Experience Teaches, in W. G. Lycan, ed., *'Mind and Cognition'*, Blackwell, pp. 29–57.
- Lewis, D. (2000), 'Causation as Influence', *The Journal of Philosophy* **97**(4), 182–197.
- Lewis, D. K. (1986*b*), *On the Plurality of Worlds*, Basil Blackwell, Oxford.
- Marcus, E. (2005), 'Mental causation in a physical world', *Philosophical Studies* **122**(1), 27–50.
URL: <https://link.springer.com/article/10.1007/s11098-005-2204-x>
- Maxwell, J. C. (1863), 'A Dynamical Theory of the Electromagnetic Field', *Proceedings of the Royal Society of London* **13**, 531–536.
URL: <http://www.jstor.org/stable/112081>
- Mellor, D. H. (1995), *The Facts of Causation*, Routledge.
- Menzies, P. (1989), 'Probabilistic Causation and Causal Processes: A Critique of Lewis', *Philosophy of Science* **56**(4), 642–663.
- Menzies, P. (2014), Counterfactual Theories of Causation, in E. N. Zalta, ed., *'The Stanford Encyclopedia of Philosophy'*, spring 2014 edn, Metaphysics Research Lab, Stanford University.
URL: <http://plato.stanford.edu/archives/spr2014/entries/causation-counterfactual/>
- Nagel, T. (1974), 'What is It Like to Be a Bat?', *Philosophical Review* **83**(October), 435–50.
- Norton, J. d. (2008), 'The Dome: An Unexpectedly Simple Failure of Determinism', *Philosophy of Science* **75**(5), 786–798.

BIBLIOGRAPHY

Papineau, D. (2002), *Thinking about consciousness*, Clarendon Press ; Oxford University Press, Oxford : New York.

Puccetti, R. (1977), 'The Great C-Fiber Myth: A Critical Note', *Philosophy of Science* 44(2), 303–305.

URL: <http://www.jstor.org/stable/187355>

Reichenbach, H. (1956), *The Direction of Time*, Dover.

Rorty, R. (1965), 'Mind-Body Identity, Privacy, and Categories', *Review of Metaphysics* 19(September), 24–54.

Ryle, G. (1988), *The concept of mind*, repr edn, Penguin Books, London. OCLC: 248432191.

Schaffer, J. (2000), 'Causation by Disconnection', *Philosophy of Science* 67(2), 285–300.

URL: <http://www.jstor.org/stable/188725>

Schneider, S. (2017), 'Events', *Internet Encyclopedia of Philosophy*, ISSN 2161-0002'.

URL: <http://www.iep.utm.edu/events/>

Schwitzgebel, E. (2015), Belief, in E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', summer 2015 edn, Metaphysics Research Lab, Stanford University.

URL: <https://plato.stanford.edu/archives/sum2015/entries/belief/>

Sellars, W. S. (1963), Philosophy and the Scientific Image of Man, in R. Colodny, ed., 'Science, Perception, and Reality', Humanities Press/Ridgeview, pp. 35–78.

BIBLIOGRAPHY

- Sider, T. (2003), 'What's So Bad About Overdetermination?', *Philosophy and Phenomenological Research* **67**(3), 719–726.
- Smart, J. J. C. (1959), 'Sensations and Brain Processes', *The Philosophical Review* **68**(2), 141–156.
- Veit, L. & Nieder, A. (2013), 'Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds', *Nature Communications* **4**, ncomms3878.
URL: <https://www.nature.com/articles/ncomms3878>
- Vicente, A. (2006), 'On the Causal Completeness of Physics', *International Studies in the Philosophy of Science* **20**(2), 149–171.
- Vihvelin, K. (2013), *Causes, laws, and free will why determinism doesn't matter*, Oxford University Press, New York. OCLC: 864685327.
URL: <http://dx.doi.org/10.1093/acprof:oso/9780199795185.001.0001>
- Weir, A. A. S., Chappell, J. & Kacelnik, A. (2002), 'Shaping of Hooks in New Caledonian Crows', *Science* **297**(5583), 981–981.
URL: <http://science.sciencemag.org/content/297/5583/981>
- Woodward, J. (2003), *Making things happen: a theory of causal explanation*, Oxford scholarship online Y, Oxford University Press, Oxford.
URL: <http://dx.doi.org/10.1093/0195155270.001.0001>
- Woodward, J. (2013), Causation and Manipulability, in E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', winter 2013 edn, Metaphysics Research Lab, Stanford

BIBLIOGRAPHY

University.

URL: *http://plato.stanford.edu/archives/win2013/entries/causation-mani/*

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