

DOUBLE PREVENTION AND MENTAL CAUSATION

By

Sean Derek Johnson

A thesis

Submitted to the Victoria University of Wellington

In fulfilment of the requirements for the degree of

Master of the Arts

In Philosophy

Victoria University of Wellington

2019

1. Introduction
  - 1.1. Introductory Matters
  - 1.2. The Powers Theory of Causation
  - 1.3. The Causal Exclusion Argument
2. Double Prevention
  - 2.1. What is Double Prevention?
  - 2.2. Why is Double Prevention an Issue?
    - 2.2.1. Schaffer
    - 2.2.2. Hall
  - 2.3. Responses to Double Prevention
    - 2.3.1. Double Prevention is Not Causation
    - 2.3.2. Double Prevention is Causation
  - 2.4. Some Tentative Conclusions
3. Double Prevention and Mental Causation
  - 3.1. Gibb's Model of Double Prevention and Mental Causation
    - 3.1.1. Introduction to Gibb
    - 3.1.2. What Gibb Needs to Show
    - 3.1.3. Gibb's Proposed Case
    - 3.1.4. Gibb and the Powers Theory of Causation
    - 3.1.5. Is This a Counterexample to the Causal Exclusion Argument?
  - 3.2. Should We Believe in n0?
  - 3.3. My Adapted Proposal
    - 3.3.1. First Pass
    - 3.3.2. Is This a Counterexample to the Causal Exclusion Argument?
    - 3.3.3. Minor Objections and Clarifying Remarks
  - 3.4. Major Objections
    - 3.4.1. Objection to Gibb's Formulation of Closure
    - 3.4.2. Grasso
    - 3.4.3. Davies
  - 3.5. Conclusion on Double Prevention and Mental Causation
4. Overall Conclusion

## 1. Introduction

### 1.1. Introductory Matters

Double prevention is often mentioned in the causation literature but is not often discussed in depth. In this thesis my primary goal is to take a deep look at double prevention and evaluate one place it has been put to work. Briefly, a case of double prevention is a case where one event prevents another from preventing a third. While we have strong intuitions that such cases should be causally relevant at least, there is debate over whether they should be counted as fully causal. Sophie Gibb (2013) puts this concept to work by arguing that mental events act as double preventers to physical events. She frames this as an argument against the causal exclusion problem.

Before I start my discussion of double prevention proper it is useful to consider some preliminaries. In section 1.2. I will go over the causal exclusion argument, which Gibb thinks can be undermined through the use of double prevention. The Powers Theory of Causation is important because of its interesting approach to double prevention and also because Gibb endorses it in motivating her argument. I will explain the Powers Theory in section 1.3.

Then on to double prevention. Section 2. will be dedicated to explaining what it is and why we should care about it. I will start in section 2.1. by elucidating what double prevention actually is and provide a clear example. In section 2.2. I will consider several arguments that show that double prevention creates tension in our accounts of causation, the counterfactual analysis of causation in particular. Then, in section 2.3. I will consider the two broad responses that we take to double prevention. Is double prevention full causation or is it something else. I will make some tentative conclusions in section 2.4.

Sophie Gibb (2013) proposes that if we do not think of double prevention as fully causal, then it can be used to create a counterexample to the causal exclusion argument. If mental events are double preventers, then they are not causes and so do not compete with the physical. Section 3. will discuss this proposal in depth. In section 3.1. I will outline Gibb's proposal and show that under several assumptions it is a counterexample to the validity of the causal exclusion argument. However, her proposal rests on the

existence of some physical event that is a sufficient cause of both our mental events and any corresponding physical causal chain. Although experiments by Libet (1985) provide some evidence for her proposal, it remains deeply controversial. In section 3.2. I will outline the reasons why we should be skeptical of her claim. But that does not mean that the essence of the proposal is lost. In section 3.3. I argue that with some adaptations that are in line with the spirit of Gibb's proposal we can create a version of her model that has a much wider appeal and does not rest on the controversial assumptions she thinks it does. In section 3.4. I will consider objections to this proposal.

Double prevention as a concept is undervalued in the causation literature. In this thesis I hope to show not only that it is an interesting topic in its own right, but also that it can be put to work in our metaphysics.

## 1.2. The Powers Theory of Causation

The Powers Theory of causation is important to this thesis both in how it responds to double prevention and because it is the theory of causation appealed to by Gibb in her reply to the causal exclusion argument. A critical evaluation of the theory is beyond the scope of this thesis, but I will briefly summarise it here.

The Powers Theory asserts the primacy of powers, or dispositions (the terms are used interchangeably) in the causal picture. Various objects in the world have certain powers. For example, broken glass has the power to cut, and a ball has the power to roll down a slope. The Powers Theorist is a realist about these powers. Powers exist even if they are not manifested. What is causation then? Causation is a mutual manifestation of powers. Events that are caused are the changes produced when two dispositions come together and both manifest. The easiest way to elucidate this is by way of example.

To borrow Gibb's (2013, pg. 197-198) example, consider a vase. The vase has the power of fragility. We can imagine a case where the vase is dropped off the counter it was on and hits the hard floor and shatters. The vase's shattering involves it manifesting its power of fragility. But on the Powers Theory, this also involves the floor's hardness being manifested. The vase's shattering is caused by the mutual manifestation of the

vase's fragility and the floor's hardness. Both are needed to ensure the effect. The vase would not have manifested its fragility if the floor had been a deep shag pile carpet and hence soft.

Another way to think about causation on the Powers Theory is as a "passing around of powers" (Mumford and Anjum, 2011, pg. 5). That is, an important part of what is happening with causation is that powers are mutually manifesting, and this results in further powers being gained. The powers passed on need not be the same as the ones that were involved in the initial causation. In our vase case we previously had the powers of fragility and hardness. The porcelain shards that are left over now have a new power, the power to cut. On the Powers Theory all causation will involve the transference of powers.

Proponents of the Powers Theory think that it is an intuitive picture of causation. However, it is revisionary and has several commitments at odds with the typical picture of causation. The following are held to be important features in the most influential version of the Powers Theory, that espoused by Mumford and Anjum (2011). Gibb (2013) also endorses this version. Firstly, the Powers Theory is committed to simultaneous causation. This is at odds with a tradition in causation going back to Hume on which causes precede their effects. On such a view there is a temporal gap between the cause and effect. However, this sort of temporal gap does not sit well with the Powers Theory. Causation occurs when powers come together and mutually manifest and the effect is the joint manifestation. What gap could there possibly be between the two? Mumford and Anjum (2011, pg. 106-129) instead argue that the Powers Theory entails thinking of causation as a simultaneous, temporally extended process. Sugar dissolving in water is a causal process that takes place over the time from the sugar first being in contact with the water to the time where the sugar is fully dissolved. The causation is the process of the sugar's solubility and the water's power to be a solvent mutually manifesting over time and the effect lasts over the time that they manifest these powers. The view is simultaneous because the cause is contemporaneous with the effect.

In addition to this, the Powers Theory rejects causation by absences. The reason for this is simple, for the Powers Theorist causation consists in the mutual manifestation of

powers, but absences cannot have any powers. This is because they are committed to a view on which powers are real and instantiated in some particular (Mumford and Anjum, 2011, pg. 145). If powers have to be instantiated in some particular and absences are just that, the absence of a particular, then absences cannot have attached powers. Without powers absences cannot play any part in causation.

Lastly, the Powers Theory denies the transitivity of causation. Causation is transitive just in case if  $a$  is a cause of  $b$ , and  $b$  is a cause of  $c$ , then  $a$  is a cause of  $c$ . But this is not always true for the Powers Theory. Causation involves the cause having a disposition to bring about the effect. Causation occurs when two or more of these dispositions are mutually manifested. But not all cases where  $a$  has a disposition to cause  $b$ , and  $b$  has a disposition to cause  $c$ , is it the case that  $a$  has a disposition towards  $c$ . Mumford and Anjum (2011, pg. 172-173) ask us to consider the case where a fire starts in a building, which causes a sprinkler system to start up, that puts out the fire. In this case the fire is disposed to start the sprinkler system, and so is a cause of it starting. Additionally, the sprinkler system is disposed to put the fire out and so it a cause of the fire stopping. But, the fire is not disposed to put itself out and so cannot be a cause of it stopping<sup>1</sup>. On the Powers Theory then, causation is not transitive.

The Powers Theory is a plausible view of how causation works. However, it is not a popular view of causation. In part this is because people who hold strong views about simultaneous causation, causation by absences, and the transitivity of causation have reason to be skeptical of it. When we propose a theory of causation in part we are trying to propose the most intuitive overall picture. The Powers Theory is intuitive in some ways, but not in others.

### 1.3. The Causal Exclusion Argument

The causal exclusion argument originates in the work of Malcolm (1968) but was substantially revised and formalised in numerous papers from Kim (including 1989,

---

<sup>1</sup> This does not follow quite as straightforwardly as Mumford and Anjum suggest. It seems to me that at least plausibly the fire *is* disposed to put itself out, just in case there is a sprinkler system present. Nevertheless, a denial of transitivity is generally thought to flow from the Powers Theory.

1993, 2005). It is considered one of the most important problems within the mental causation literature and has been dubbed ‘the new mind body problem’. But what exactly is the causal exclusion argument?

Ordinarily, we like to think that the mental can be a cause of various things. Intuitively, the desire to wave my hand is a cause of my hand actually waving. Similarly, my belief that dodos are extinct is usually thought to be a cause of my answering questions about dodos in a certain way. However, when it comes to explaining exactly how this causation works there seems to be a problem. This is because we seem to be able to fully describe the causes of any event in terms of merely physical things. When we talk about hand waving we can talk of the causes being tendons moving, muscles contracting, and so on. We can follow this path back to neurons firing in the brain. Perhaps we can go deeper and describe the causes in terms of atoms and molecules interacting. However, if we accept this view of causation, then it seems like there is no work for the mental to do in the causal process. We can describe all of the causal process in physical terms. What is the mental doing?

The causal exclusion argument is often formally advanced as either an argument in favour of some sort of mental and physical identity, or as an argument for epiphenomenalism (Bennett, 2003, pg. 472). Epiphenomenalism is the thesis that it is not the case that mental properties are ever causally relevant to physical events<sup>2</sup> (Maslen, Horgan, and Daly, 2009, pg. 528). That is, although physical events could possibly cause mental ones, the reverse is not true. If one accepts epiphenomenalism, then they have to give up the intuitive ability referred to above for the mental to do any causal work with respect to the physical. A simple version of the argument can be roughly sketched as follows:

1. *Distinction*. Mental events are not physical events.
2. *Physical Causal Closure*: Every physical event contains only other physical events in its transitive causal closure.

---

<sup>2</sup> There is much debate in the literature as to what is the most appropriate causal relata. For the purposes of this thesis I will generally write as if events are the causal relata. I do not believe that any of the arguments and positions I offer are affected by the choice of causal relata (except where explicitly stated). Due to this, my use of events as the causal relata is not due to metaphysical commitment but rather ease of writing.

3. *Exclusion Principle*: No single event can have more than one sufficient cause occurring at any given time – unless it is a genuine case of overdetermination.
4. *Denial of Overdetermination*: Mental Causes do not systematically overdetermine their effects.
5. *Therefore*: Epiphenomenalism.

The argument can broadly take as its target anybody who accepts the premises.

However, it is usually taken to be an argument against any position that does not accept identity between the mental and physical. Notably, this includes non-reductive physicalism and dualism. Proponents of these positions who wish to avoid epiphenomenalism have some work to do in responding to the argument.

Part of the appeal of the causal exclusion argument is that each of the premises seem intuitively plausible while its conclusion is something that we intuitively want to avoid.

Distinction follows straightforwardly from the positions targeted by the argument. The distinctness of the mental and the physical is a defining commitment of traditional dualism.

Non-reductive physicalism makes a similar claim. Even though the mental supervenes on the physical, it should be considered as distinct and non-reducible. If non-reductive dualism thought that the mental and physical were not distinct, then it would just be a form of reductionism. The non-reductive physicalist often holds this view of distinctness because of the multiple realisability of the mental. That is, a mental state that we might have such as pain could be realised by many different physical states. If this is true, then the mental state cannot be reduced to any one physical state and should be considered something distinct.

Kim (2005, pg. 22) says that the Exclusion Principle is a straightforward metaphysical principle that we ought to accept. A case of overdetermination is any case where there is more than one sufficient cause at a given time. The Exclusion Principle simply states that in any case that is not overdetermined there is only one sufficient cause at any given time. This should be accepted readily because all cases that do have more than one sufficient cause at a given time are covered by overdetermination (Maslen, Horgan,



and Daly, 2009, pg. 528). Kim argues extensively in favour of the Exclusion Principle in his 1989 and 1993 papers, as does Malcolm (1968).

With regards to the Denial of Overdetermination, it is clear that the Exclusion Principle is not enough by itself to further the argument due to the possibility of overdetermination. As noted above, a case of overdetermination is one where there is more than one sufficient cause for some effect. The classic example of this case is where two assassins shot a target at exactly the same time. Both shots would be sufficient to kill the target. This would be a breach of the Exclusion Principle but is seemingly unproblematic whereas other breaches of Exclusion are. Hence the qualification to the Exclusion Principle. Nevertheless, it seems unlikely that the mental causal process is systematically overdetermining the physical. If the mental did overdetermine physical causes in every case, then we would have to postulate a pervasive global coincidence. This would be at odds with principles that say we ought to prefer simple solutions such as Occam's razor.

Physical Causal Closure is the most difficult of the premises to pin down, because there are several different versions that are endorsed in different arguments. I have chosen the variation proposed above, 'every physical event contains only other physical events in its transitive causal closure', because that is the version Gibb (2013) endorses in the argument that I will be responding to<sup>3</sup>. The basic idea that is being drawn out by an appeal to Physical Causal Closure is that for any physical event we ought to be able to look back at its causal history and see only physical events. A non-reductive physicalist ought to accept Physical Causal Closure on the grounds of their overall physicalist position. A dualist ought to accept Closure because it follows from our best physics and because of the success that physics has in making sense of the causal structure of our world.

All of the premises of the argument then seem plausible. But if this is true, and the argument is valid, then it is a compelling case for epiphenomenalism. But epiphenomenalism is in conflict with a deeply rooted picture of ourselves as agents. We also have powerful everyday intuitions that our mental states are causally relevant. It seems that the causal exclusion argument is a pressing one in need of a solution if we

---

<sup>3</sup> As I will show in section 3.4.1., Gibb's reply is compatible with many different formulations of Closure.

are to keep these sorts of commitments. In my final chapter I will explore an attempt by Gibb to do just that.

## 2. Double Prevention

Double prevention is an underdeveloped concept in the causation literature. It is my aim in this chapter to explain the concept in a precise and clear way and explain why it is problematic for many analyses of causation. I also wish to survey the responses that have been made to the tensions that it causes. My main goal is to show that double prevention is important and in need of further philosophical inquiry. I will start by explaining what double preventions are. I will then look at two arguments by Schaffer (2000) and Hall (2004) that seek to show that double prevention is in tension with widely held beliefs about causation. Finally, I will survey some of the ways that the literature has responded to the tensions caused by double prevention. While I am mainly trying to survey the place of double prevention in the literature, I will offer some tentative conclusions at the close of the chapter.

### 2.1. What is Double Prevention?

Double preventions occur when one event (*a*) prevents another event (*b*) that would have prevented a third event (*c*) from occurring. Thus, if (*a*) occurs, then in this scenario (*c*) also occurs due to a failing of (*b*) to eventuate. Furthermore, it seems intuitive to say that (*a*) is at least relevant in bringing about (*c*).

To illustrate this, consider the following example adapted from Ned Hall (2004). Carl is a bomber pilot and is flying a mission to bomb an enemy target. Aimee is flying in formation with Carl and is his lone escort for the mission. The enemy scrambles a Fighter in response to the threat with the goal of shooting down Carl's bomber. Luckily for Carl, Aimee sees the enemy Fighter before it gets to them and manages to shoot Fighter down before they can do any damage to Carl's bomber. If she had not done this, then the enemy Fighter would have shot down Carl's plane and prevented him from dropping his bombs on the target. However, Aimee prevents this, and Carl successfully completes the mission and bombs the target.

This scenario can be seen diagrammatically as follows. A letter represents an event. In this diagram, (*a*) corresponds to Aimee shooting the Fighter, (*b*) corresponds to the

Fighter shooting down Carl, (c) corresponds to Carl dropping bombs on the target, and (e) is the end effect, the target having been bombed. (x) and (y) are events that led up to Carl being in the position to drop the bombs, perhaps getting in to his plane, or an earlier point in the flight.

A letter with a circle represents a potential event that didn't occur. An arrow represents a causal connection and an arrow with a dotted line represents a causal connection that failed to occur. As we are talking about preventions a line with a dot at the end represents a prevention and a dot with a dotted line represents a prevention that failed to occur.

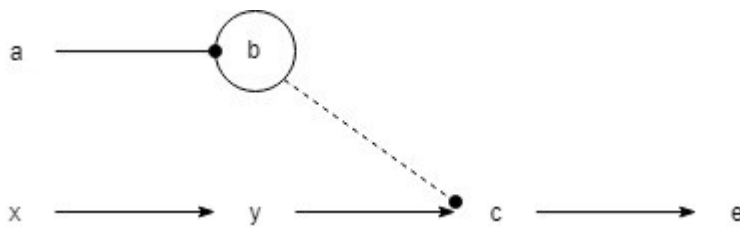


Figure 1

In this scenario, Aimee shooting down Fighter is acting as a double preventer for Carl bombing the target. Aimee shooting the Fighter prevents it from preventing Carl from dropping the bomb. It is clear in this scenario that Carl dropping the bombs (c) represents a cause of the target having been bombed (e). What is much less clear is whether (a) also counts a cause of the target being bombed. Intuitively, it seems clear that it had a role to play, after all if (a) had not occurred then (e) would not have occurred. In the following section I will consider some arguments as to why we might want to resist counting double prevention as fully causal.

## 2.2. Why is Double Prevention an Issue?

The major issue with double prevention is that it threatens to undermine some widely held beliefs about causation. Ned Hall (2004) and Jonathan Schaffer (2000) explore these issues separately. Schaffer shows how double prevention is in tension with the thesis that causation must involve some sort of direct connection, such as a flow of energy. He then shows that double prevention is commonplace in the usual cases we

think of as causes. This gives us a good reason to reject accounts of causation that rely on such connections. Hall argues that even though it might seem that intuitively a counterfactual analysis of causation has nothing to fear from double prevention, double prevention is in tension with some of the supplementary theses that go along with such views. Hall argues that a simple form of counterfactual analysis will involve belief in four theses – Transitivity, Locality, Intrinsicness, and Dependence. Causation by double prevention is in tension with the first three of these.

### 2.2.1. Schaffer<sup>4</sup>

One widely held approach to causation holds that causation must involve a connection from cause to effect (Schaffer, 2000, pg. 285). This sort of connection can be sketched out in a couple of different ways. One way is for that connection to be some sort of physical process such as an energy flow. In the classic case of billiard ball causation, kinetic energy from one ball is imparted into another one causing it to roll in a certain direction. This energy flow is the physical connection. If we endorse the physical connection thesis then this physical connection is necessary for causation to occur<sup>5</sup>.

These programs leave no room for causation by double prevention. They all require a real physical connection between cause and effect for causation to occur and that is simply not the case with double prevention. Consider our bomber example from earlier. For Aimee to cause the bombing to take place there must be a physical connection to

---

<sup>4</sup> In his paper, Schaffer does not talk explicitly about double prevention. He instead talks about 'causation by disconnection'. Nevertheless, it is clear that he is using the same concept that we are referring to by double prevention. Indeed, he notes this in his footnote on page 286. His main reason for this terminological change is that he uses the word 'preventions' standardly to refer to the case where one event stops another from occurring and didn't want to cause any confusion. I shall be proceeding using the term double prevention.

<sup>5</sup> It is worth noting that energy flows are not the only sort of possible physical connection. Schaffer (2000, pg. 289-290) briefly summarises three main research programs that focus on different sorts of physical connections. The program of Jerrold Aronson, David Fair, and Hector-Neri Castaneda takes energy transfer to be necessary for causation. Bertrand Russell, Wesley Salmon, and Phil Dowe are part of a second program where causal processes are basic entities that need to be specified. A third program postulates that the causal mechanism is "a process which underlies a regular sequence and each phase in which exhibits qualitative as well as spatio-temporal continuity" (Mackie, 1974, pg. 222). This program is that of J. L. Mackie and Douglas Ehring. Schaffer's argument attacks the physical connection thesis at a more fundamental level than the intricacies of these particular programs. As such I will not go in to them in any depth. I just want to make it clear that the energy transfer I use in my examples is not the only sort of possible physical connection that one could endorse.

the bombing itself, or at least an intermediary, such as Carl's plane. But what connection could this possibly be. Aimee simply doesn't interact directly with Carl in this way. To make things even more clear, it would still be a genuine case of double prevention if Aimee shot down the enemy Fighter 100 kms away from Carl, and Carl had no knowledge of this occurring. It seems implausible that there is a physical connection here because there is nothing in the space between the cause and the effect. Because of this, if we are to hold to a physical connection thesis then we ought to deny that double prevention counts as causation.

A different way of spelling out the connection from cause to effect is to appeal to an intrinsic relation between them<sup>6</sup>. For any cause *c* and effect *e* the thing that puts them in a causal relationship is determined by the internal properties of the events and their interaction together. Appealing to causal relations being intrinsic is useful in dealing with cases of pre-emption as the relationship between cause and effect is not affected by any alternative processes (Menzies, 1996, pg. 99).

However, this approach also leaves no room for double prevention to be causation. Causation by double prevention is entirely about extrinsic factors. There is no intrinsic connection between Aimee shooting down the enemy Fighter and the target being bombed. It is reliant on the extrinsic circumstance of Carl flying the mission in his bomber.

Schaffer has shown that two different ways of defining the widely held view that causation requires a connection between cause and effect. Both of these are inconsistent with double prevention being causation. This is not necessarily a problem, as the option is seemingly open to us to reject double prevention as causation. Proponents of a physical connection or intrinsic connection theory can simply bite the bullet and accept that there are some cases where no causation occurs because "there is no connection. Aronson (1971, pg. 425, in Schaffer, 2000, pg. 291) does exactly this:

Consider a weight that is attached to a stretched spring. At a certain time, the catch that holds the spring taut is released, and the weight begins immediately to accelerate. One might be tempted to say that the release of

---

<sup>6</sup> Both Peter Menzies (1996) and David Armstrong (1999) endorse this approach.

the catch was the cause of the weight's acceleration. If so, then what did the release of the catch transfer to the weight? Nothing, of course."

Schaffer notes that this would be an appropriate strategy if such cases were few and far between. However, he further argues that they are not. In fact, cases of double prevention are ubiquitous when talking about paradigmatic cases of causation and cannot simply be thrust aside. I will explore this argument in section 2.3.2.1. For now, what is important is to note that double prevention is problematic as it creates a tension between our intuitive want to count double prevention as genuine cases of causation, and the widely held view that it is a necessary condition of causation that there exists a connection between cause and effect.

### 2.2.2. Hall

It might seem that a counterfactual analysis of causation would have no problem accommodating double prevention. After all, at its simplest, a counterfactual analysis simply requires that an effect is counterfactually dependant on its cause. However, Ned Hall in "Two Concepts of Causation" (2004) argues that double prevention presents a threat to counterfactual theories of causation as well. Hall notes that there are four theses that are often endorsed as part of a full counterfactual theory of causation. These are Dependence, Transitivity, Locality, and Intrinsicness. Dependence is the crux of the counterfactual position. While the other three theses are added to ward off common objections to it. Hall (2004, pg. 225) lays out these theses as follows:

*Dependence:* Counterfactual dependence between wholly distinctive events is sufficient for causation.

*Transitivity:* If event *c* is a cause of *d*, and *d* is a cause of *e*, then *c* is a cause of *e*.

*Locality:* Causes are connected to their effects via spatiotemporally continuous sequences of causal intermediaries.

*Intrinsicness:* The causal structure of a process is determined by its intrinsic, non-causal character.

As noted before, Dependence is clearly consistent with double prevention as causation. In fact, following Dependence, double prevention must be causation. In the bomber case the target being bombed is counterfactually dependent on Aimee shooting the enemy Fighter down. If she had not done this then Carl would himself have been shot down and the target would never have been bombed. Without Aimee's intervention the end effect would never have occurred and so under Dependence it should count as a cause. However, Hall argues that double prevention as causation is in tension with the other three theses. In this section I shall explore those tensions<sup>7</sup>.

I will start with the tensions between Transitivity and double prevention. Transitivity is the thesis that allows causes to act through intermediaries and create causal chains. Hall (2004, pg. 235) also notes that Transitivity is extremely important in helping resolve problems with early pre-emption. Hall (2004, pg. 246-248) focus in on the issues with Transitivity by adding to the story of Aimee, Carl, and the enemy Fighter. Hall asks us to imagine that in the morning before the incident the enemy Fighter's alarm goes off, waking him up. If it had not gone off, then bad news for the Fighter, he would have been put on latrine duty and not been able to fly any missions. Thus, he would not have been on scene where he was in fact shot down. If the alarm had not gone off, then Aimee would never have shot down the enemy Fighter. But, if Aimee had not shot down the Fighter then the mission would not have been successful, and Carl would not have bombed his target. If we accept that Aimee shooting down the enemy Fighter was a cause of the target being bombed and that the alarm ringing is a cause of the enemy Fighter being shot down, then following Transitivity the alarm going off was a cause of the target being bombed.

But this is absurd. As we can see in the below diagrams of Halls<sup>8</sup>, whether the alarm goes off or not the target is still bombed. The enemy Fighter never has a chance of

---

<sup>7</sup> It is not my intention to defend the theses that Hall presents. In fact, one could quite happily reject them. It is merely my aim in this section to show that there is a tension between commonly accepted theses and double prevention as causation. This is sufficient to show that double prevention is a serious issue that needs to be confronted by many people in the debate around causation.

<sup>8</sup> Like in the previous diagram (a) refers to Aimee shooting down the enemy Fighter, (b) refers to the enemy Fighter shooting down Carl, and (c) refers to Carl dropping the bombs. (x), (y), and (z) refer to other events in the lead up to Carl being in a position to drop the bomb. (f) refers to the alarm going off, and (d) refers to some event on the enemy Fighters timeline that would cause him to be in a position to shoot down Carl, such as getting in his plane.



stopping it. Either their alarm never goes off and so they fail to shoot Carl down, or the alarm does go off and they are shot down themselves by Aimee, and so fail to shoot Carl down. Yet, if both Transitivity and Dependence are correct then the alarm going off is a cause of the target being bombed. Double prevention causes this consequence and therefore draws two theses that we want to hold into tension.

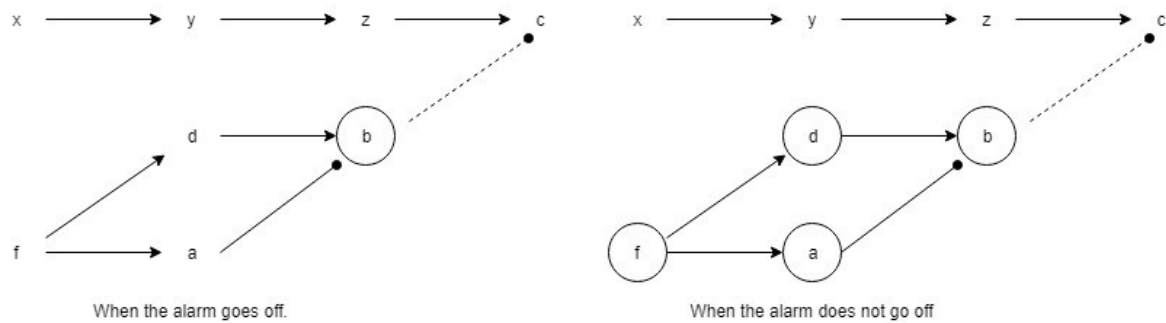


Figure 2: Adapted from Hall (2004), pg. 247, Figure 9.7.

Moving on now to problems with Locality. Locality is a thesis that tries to make sure that causes and effects are located appropriately to each other. Hall (2004, pg. 238) invokes it to help deal with the issue of late pre-emption. Imagine that Aimee and Carl are instead throwing rocks at a bottle, they both make a perfect throw that would smash it, but Aimee's rock reaches the bottle first and is the one that does the smashing. How might we determine which of the throws is the cause of the bottle breaking? The effect is not counterfactually dependant on either one or the other. Hall suggests that what might be important is that Aimee's throw is spatiotemporally connected in the right way to the bottle shattering. What is important is Locality, that the cause (Aimee's throw) is connected to the effect (the bottle shattering) by spatiotemporally continuous sequences of causal intermediaries (the various positions that the bottle flew through while in the air, culminating in actually hitting the bottle).

So, Locality seems like a reasonable sort of thesis to endorse. However, when confronted with cases of double prevention there again is tension. In cases of double prevention there is a spatiotemporal gap between the cause and effect. Taking our fighter pilot example where Aimee shoots down the enemy Fighter 100 kms away from Carl, and Carl has no idea that the event even took place. There is a gap both spatially, and temporally between the cause and effect. The enemy Fighter is shot down a long

way from the bombing and there are no connected events between the two. Equally, we can assume that the Fighter was shot down some time before the bombing and so there is also a temporal gap.

As with Transitivity, double prevention represents a case of causation that appears to fail Locality. This furthers the tension between Dependence and the other theses.

Lastly let's look at the tension between double prevention and Intrinsicness.

Intrinsicness claims that the causal structure of a process is determined by its intrinsic, non-causal character. Thus, if we have two cases where the intrinsic character of the scenario is the same, then they should share conclusions about what is a cause.

This needs some unpacking. What is intrinsic character? The intrinsic character of some effect is the structure of events in the causal history of the effect. We can choose some arbitrary time before *e* and take the causal chain from that time to the occurrence of *e*. That structure of events is the intrinsic character. What Intrinsicness is saying is that if we took that chain of events and put it in a new world with the same laws, then we would cause *e* in that world.

The Intrinsicness thesis is another way of dealing with problems caused by late pre-emption. Imagine the rock throwing, bottle smashing example. The bottle smashing is not dependent on Aimee's throw because if her throw failed to smash the bottle then it would still have been smashed by Carl's throw. Thus, the bottle smashing is not dependent on Aimee's throw.

We can however, imagine a second world in which Carl does not exist. In such a world Aimee still throws her rock and smashes the bottle. In this world the bottle smashing is dependent on her throwing the rock and is a clear case of causation. The causal chain in this world that lead to the bottle smashing is exactly the same as the one in the first world. It might include such things as Aimee's arm moving in a throwing motion, her releasing the rock, the rock's trajectory, and the rocks impact with the bottle. The two worlds intrinsic character, with respect to the bottle smashing, is identical.

The Intrinsicness thesis claims that because the intrinsic character is identical in both worlds, and that Aimee's rock throw is clearly the cause of the bottle smashing in the second world, that her rock throw counts as the cause in the first world.

However, this breaks down when considering double preventers. Let's use our bombing example. The intrinsic character of the bombing includes all the causal chain back to some arbitrary time point. So, in this case it would include things such as Aimee flying her plane, pulling the trigger, the bullets hitting the enemy Fighter. It would also include Carl flying his bomber. These sorts of things. If Intrinsicness holds, then any other world where all these causes are present we should also see the target being bombed.

We can imagine such a second world, with one change, that the enemy command centre was never going to order the enemy Fighter to shoot down Carl. Such a world would have the same intrinsic character as laid out above, but Carl bombing the target would not be dependent on Aimee shooting down the enemy Fighter. In this world Aimee's actions are not a cause of the target being bombed.

If this is correct, then Intrinsicness is false. We have a case of identical intrinsic character but different causal structure. It is tempting to respond that we must have not defined the right intrinsic character to start with. We should have included the fact that the enemy Fighter would be ordered to shoot down Carl as part of the initial causal chain. However, this seems ridiculous. In doing so we would be forced to admit unintuitive things such as that the enemy Fighter being ordered to shoot down Carl caused the target to get bombed. As before there seems to be a great tension between double prevention and Intrinsicness.

The arguments of Schaffer and Hall show that double prevention is a major issue that persons wanting to provide a complete analysis of causation need to address. Schaffer showed that counting double prevention as fully causal is inconsistent with analyses of causation that involve a physical or intrinsic connection from cause to effect. Hall has argued that counterfactual analyses of causation are also in conflict with double prevention. A simple version of a counterfactual analysis that only involves an appeal to Dependence might still work. But to make a robust counterfactual analysis we often appeal to the supplementary theses of Transitivity, Locality, and Intrinsicness. Hall has

shown that double prevention is in conflict with the conjunction of Dependence and these theses. In the case of the counterfactual analysis of causation the issue of double prevention is even starker than other analyses because the central thesis of Dependence clearly seems to indicate that double prevention should be counted as genuine cases of causation. If we are going to count cases of double prevention as real causes, then they expose several tensions with other widely held beliefs about causation. In response to this we either need to accept double prevention as not being fully causal or argue for an analysis that makes sense of the tensions it creates. In the next section I will look at several responses to double prevention.

### 2.3. Responses to Double Prevention

The literature regarding double prevention is relatively underdeveloped. In the following section I will explain the three major attempts to fit double prevention into an analysis of causation. All have their trade-offs. The first attempt I will consider is that of Mumford and Anjum (2009). Mumford and Anjum argue for a Powers Theory of causation. On the Powers Theory double prevention is straightforwardly not causation. The Powers Theory is not a popular theory of causation, but its ability to provide a plausible account of double preventers is argued to be a point in its favour (Gibb, 2013, pg. 202). Hall (2004) and Schaffer (2000) both endorse versions of an approach to causation that make room for double prevention as fully causal. Schaffer argues that an account of causation that relies on necessary and/or sufficient conditions such as Hume's constant conjunction or Lewis' counterfactual approach can work with double prevention as long as it allows for causation by absences. Hall has argued against the counterfactual approach being consistent with causation and instead endorses a hybrid approach on which causation come in two forms. Dependence is one of these, the other he calls Production. After explaining all of the approaches and the different ways that they deal with double prevention I will make some tentative conclusions regarding it.

### 2.3.1. Double Prevention is Not Causation

In light of the tensions argued for in the previous section, counting double preventer as not fully causal should be the default position. Hall and Schaffer argue that double prevention presents major problems for connection based and counterfactual based theories of causation. As we shall see shortly, both of them propose radical changes to those theories in order to accommodate double prevention. Both also argue for theories that allow causation by absence and Schaffer in particular thinks that this is necessary for an account that can make sense of double prevention. Thus, those who want to hold an unmodified connections based or counterfactual based view, and those who deny causation by absences, should be skeptical of double prevention being fully causal. However, At this time analyses of double prevention on those views are lacking and so I will turn to a view that does have a full analysis of double prevention as not fully causal.

Mumford and Anjum's (2009, 2011) Powers Theory of causation emphatically does not count double prevention as a form of causation. Mumford and Anjum argue that considering the tensions caused by counting double prevention as causation, we ought to prefer a theory of causation that does not count double preventers as causes, all else being equal. They claim to have just that in the Powers view.

Recall that the Powers view asserts that causation can be understood as a 'passing around of powers'. That is, causation occurs when a disposition of an object (or objects) manifests with the result that there is some new disposition in the resultant object. For instance, in the vase smashing case, the effect of the vase smashing is caused by the mutual manifestation of the vase's disposition towards fragility in the face of hard objects, and the floor's disposition towards hardness. The resulting shards of glass have new powers, such as the disposition to cut.

A key aspect of their Powers view is that causation is that cause and effect must happen simultaneously or in an overlapping fashion. This is because causation involves a mutual manifestation of powers. If the rock was not manifesting its disposition towards hardness at the same time that the bottle was manifesting its disposition towards fragility in the face of a hard object, then the bottle wouldn't smash. Because of this a case of double prevention cannot be a case of direct causation. Double prevention can,

and often does happen at a spatiotemporal distance and this falls foul of the above rules. It is also the case that double prevention cannot be a case of indirect causation. This is because such indirect causation would require an intermediary absence to carry the causation (Gibb, 2013, pg. 201). Accept for a moment that double prevention is causal. In our bombing case Aimee shooting down the enemy Fighter would be a cause of the bombing. In the Powers Theory of causation absences cannot be causes and therefore a causal chain making use of absences cannot be causation either. Absences are not the sorts of things that can carry powers and so cannot be causes.

This rejection of double prevention allows proponents of the Powers view to avoid the tensions that double prevention brings with it. But how does it classify double prevention then. It seems clear that even if double prevention isn't actual causation it still has some role to play. Gibb (2013) argues that even though double prevention is not causation it is still causally relevant in that it allows or permits causation to take place. It is causally relevant in that it permits the causation to occur. This is not just an explanatory role. This role in permitting the causation is an objective one that is independent from the attitude we take to the case and our interest in it (Gibb, 2013, pg. 202). Gibb sees this distinction between events that are causes and events that permit causation as another boon for the Powers view. Such a distinction can help us explain double prevention without any of the baggage, while still maintaining its importance. The distinction will also form the crux of her solution to the causal exclusion argument.

The response of the Powers Theory seems well and good on the face of it. Although it does not count double prevention as fully causal, it does make a strong case for why it is not given acceptance of the theory. In addition, it does make some sense of the fact that we think of double prevention as causally important by giving it the important role of permitting causation. It is not a cause, but that does not mean that it is causally worthless. Overall the Powers Theory makes a strong case that it can include an account of double prevention within its analysis of causation without making many sacrifices to do so.

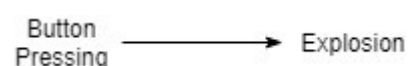
### 2.3.2. Double Prevention is Causation

In contrast to the Powers Theory, Hall (2004) and Schaffer (2000) both argue that cases of double prevention ought to count as real cases of causation. To make this compelling they must do two things. First, they must show that cases of double prevention are central to our concept of causation, and then they must resolve the tensions that I discussed above. They both attempt to do just that.

#### 2.3.2.1. Schaffer

##### 2.3.2.1.1. Double Prevention is Important as Causation

Schaffer (2000) argues that cases of double prevention are paradigmatic of causation and any theory that cannot accommodate them as such is therefore deficient. He asks us to imagine the simple case of pressing a button that sets off an explosion. There is nothing special in this case, some technician pushes the button and because the button was pushed the explosion occurs. Surely this case counts as causation. To reject it would be to reject everything about causation as we typically understand it. However, this case could potentially be one of double prevention. The mechanism of the button could work in a couple of different ways. It could simply work such that by pressing the button a pulse is emitted that which connects to the bomb making it explode.



*Figure 3*

Alternatively, it could be the case that pushing the button inhibits an electrical current already in place that was preventing an independent source from setting off the explosion. Pushing the button could act as a double preventer.

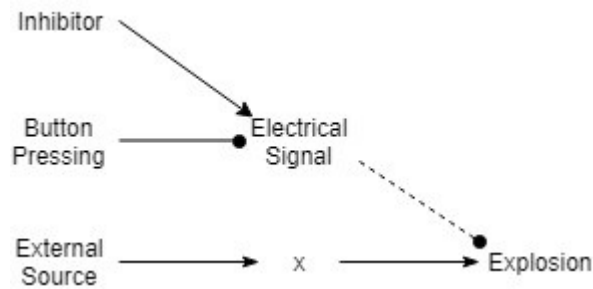


Figure 4

We do not ask how the button works before saying that pressing the button causes the explosion. From this example it seems that we are intuitively willing to grant causal status without considering how exactly the scenario is wired. Without wondering if there is double prevention at play. This suggests that whether something is causal or not is irrelevant of the question of whether it is double prevention. Therefore, cases of double preventions ought to count as genuine causation, all else being equal.

There is one other option open to us. We could accept that the button pressing case looks like causation at first glance. But, upon looking closer at it and discovering that in this case it is actually a case of double prevention we ought to reject our initial intuition. So, when we see it is the case the pressing a button seems to cause an explosion, we ought to initially say that yes it seems like causation is occurring. If it happens to be the case that the button is wired such that it connects directly to setting of the explosion then great, our intuition was correct. However, if it turns out to be a case of double prevention then we just must accept that our intuitions were wrong.

This might seem like an odd approach. After all, our initial intuition that pressing the button is straightforwardly causal was a strong one. We don't want to throw it away. But, if you are committed to an approach that denies causation to double prevention then it might be plausible to bite the bullet, claim that this is simply an odd case, and that even though we had a strong intuition it turns out we were wrong.

The problem, says Schaffer (2000, pg. 286-289) is that this is not an odd case. It is actually an extremely typical one. Schaffer investigates the extended case of an angry man shooting another and notes that there is double prevention rampant throughout all of the intermediary causal steps. Overall the causal chain might go as follows (Schaffer,



2000, pg. 286): “The killer gets angry, pulls the trigger, fires a bullet through the victim’s heart, and the victim dies.”

We can zoom in on one part of this, the bullet piercing the victims heart causing them to die. This is a case of double prevention. The victim needs a steady supply of oxygen to the brain in order to prevent oxygen starvation. The heart plays a part in supplying this oxygen. However, the piercing prevents the heart from carrying out its role. The piercing prevents the heart from preventing oxygen starvation. Which causes death. In other words, the piercing causes death through a process of double prevention.

I will not go into more detail here, but Schaffer shows that this is the case for all the extended causal chain above. Seemingly this could apply to any number of other paradigmatic cases of causation as well. Cases of double prevention are not odd cases that we can ignore or dispense with, they are common and too important to ignore.

So, according to Schaffer, we need to accept double prevention as causation. How can he resolve this with the tensions that we discussed earlier? He (2000, pg. 294, italics in original) suggests that to deal with the tensions any adequate account of causation must “involve necessary and/or sufficient *conditions*, and relate *absences*”.

#### 2.3.2.1.2. Conditions Based Approaches

Schaffer suggests that approaches to causation that involve necessary and/or sufficient conditions, such as counterfactual approaches to causation or Hume’s constant conjunction account, have no difficulty dealing with double prevention. He thinks this as long as such accounts also allow for causation by absence. I will focus on counterfactual approaches as I have discussed them in previous sections, but the following applies to other similar approaches to causation as well. Schaffer takes the counterfactual approach to be asserting something like Dependence from our previous section. That is, counterfactual dependence between wholly distinct events is sufficient for causation. Indeed, as mentioned in my section on Hall, Dependence by itself can easily accommodate double prevention. The issue Hall raised was that even if double

prevention is easily accommodated by Dependence, it lies in tension with the related theses of Transitivity, Locality, and Intrinsicness.

Schaffer (2000, pg. 294-296) argues that we can deal with the problem of Locality by enforcing his second criterion of an adequate causal theory, that we allow causation by absences. What this means is that just as we allow causation to occur between two obtaining events we should also allow the absence of something to be a cause or effect. Causation by absences is a controversial topic within the causation literature. However, Schaffer deems it essential to deal with the tension with Locality.

Recall that Locality states that “Causes are connected to their effects by via spatiotemporally continuous sequences of causal intermediaries”. This is in tension with double prevention because in cases of double prevention there is a potentially large spatiotemporal gap between cause and effect. If we allow causation by absences, then this worry goes away. In our fighter plane example there was a gap between Aimee shooting down the enemy Fighter, and Carl completing his mission. This was touted as a failure of Locality. However, if we allow causation by absences then this gap is filled by the absence of the enemy Fighter close to Carl’s plane (realistically the gap is filled by a chain of absences of the enemy Fighter at various points between where they were shot down by Aimee and where they would have in turn shot down Carl). Thus, there is no tension with Locality.

In summary, Schaffer argues that a conditions based approach to causation, coupled with allowing causation by absences, is one adequate way that we can deal with the tensions created by double prevention. However, this approach is not without its issues. Although Schaffer does deal with the tensions with Locality by embracing causation by absences, he does nothing to address the tensions with Transitivity and Intrinsicness that were brought up by Hall. Now, to be fair, Schaffer never claims to try to do this, and we can grant his conclusion that causation by absence is at least a necessary factor in any conditions based account of causation that wants to deal favourably with double prevention. Nevertheless, any such account that wants to be compelling clearly has more work to do. In the absence of that argument we should be skeptical that appealing to a basic sort of conditions based account will succeed.

### 2.3.2.1.3. Hybrid Approach

Schaffer also suggests that a hybrid strategy that combines the above with a connection based approach might also deal adequately with double prevention. As with the above, such a strategy would have to make a place for causation by absences. A connection based approach is one similar to the variety that were discussed in section 2.2.1. on which there is some sort of physical or intrinsic connection between the cause and effect<sup>9</sup>. The hybrid view that Schaffer argues for notes that such connections are important for causation. Cases involving connection from cause to effect are causation. But that there is also causation in cases of conditional connection.

Schaffer starts with the hybrid approach of Fair (1979, pg. 274), who after noting that an energy-flow based account of causation cannot accommodate causation by absence, looks to counterfactuals about energy flow and suggests that there are broadly four different sorts of causal process’.

1. C causes E (Connection).
2. C causes  $\sim E$  (Prevention).
3.  $\sim C$  causes E (Omission).
4.  $\sim C$  causes  $\sim E$  (Omission of Prevention).

Causation of type (1) is a standard case of connection. Causation of type (2) occurs when there is a causal connection between C and E’ that is incompatible with E. Causation of type (3) occurs when C does not occur and if it had then it would have a causal connection with some E’ that is incompatible with E. Causation of type (4) occurs if had C occurred then there would be causal connection to E.

In this way, Fair takes the notion that causation necessarily involves connection and expands his overall causal analysis to include problematic cases concerning absences by endorsing counterfactuals about connection. Schaffer notes that those wanting to endorse a hybrid strategy and accommodate double prevention need to allow for causation by absences for the same reason as those who endorse a conditions approach do. To deal with problems of Locality. Double prevention raises the possibility of

---

<sup>9</sup> There could well be other sorts of connection that are appealed to.

causation at a distance, a gap in the spatiotemporal continuous sequence of causal intermediaries that is thought to connect cause and effect. If we allow causation by absences, then this gap is bridged.

Schaffer (2000, pg. 296-297) argues that there are issues with Fair's approach. The first is that our normal causal semantics do not map nicely to the model. Earlier, I explained how Schaffer argues that any paradigmatic case of causation of the that we think of as being in the first form could actually be double prevention. Recall that case of the button pushing causing the explosion. As read, this is causation of the first, standard form. The occurrence of A causes the Occurrence of B. However, this case could possibly be one of double prevention where the button pushing causes the explosion through inhibiting an electrical current that was blocking an independent source from setting the explosion off. In this case there is causation of the second form where pushing the button causes the non-occurrence of the current flow. There is also causation of the third form, where the non-occurrence of the current flowing causes the independent signal to get to the explosive device.

So what form should the causal chain from the button pushing to the explosion take? In our normal semantics it is of form one. But form one is defined as direct connection which is not what is happening here. It might be tempting to claim that the hybrid theorist can chain the different forms in various ways. This would be such that the four forms above would define direct causation and causation generally would simply be various chains of direct causations. In this system, double prevention would simply be a chain of form two and form three. However, Schaffer argues that there are issues with this approach. We are in no position to assert that all the different possible n-length chains (e.g. (1)-(2)-(3)-(1)-..., (4)-(3)-(4)-(2)-... etc.) are all really causation. He also notes that this approach heightens any problems with disunity of causation. If we are to allow any number of different chains of direct causation to count as causation overall, then it is hard to see what makes causation special as a singular concept.

Schaffer's proposes that an adequate hybrid approach must unify presences and absences into a single form and then use this to unify the four different forms of causation. Thus, any chaining that we do would be simple repetition rather than an entirely new type of chain of causation. This solves any disunity problems and we have

good reason to think that any length chain of these direct causations will be causation generally due to Transitivity.

Schaffer's suggestion to achieve this is to take the causal relata to be  $\langle \text{Property, Region} \rangle$  pairs that include absences by allowing for negative properties. In our button pushing causing explosion example the relevant causal relata might be pairs such as  $\langle \text{Pushed, Button} \rangle$ ,  $\langle \text{No Current, First Wire} \rangle$ ,  $\langle \text{Current, Second Wire} \rangle$ , and  $\langle \text{Explosion, Outside} \rangle$ . If there is to be causation from pushing the button to the explosion, then there must just be a chain of causation linking these relata together. In this case:

1.  $\langle \text{Pushed, Button} \rangle \rightarrow \langle \text{No Current, First Wire} \rangle$
2.  $\langle \text{No Current, First Wire} \rangle \rightarrow \langle \text{Current, Second Wire} \rangle$
3.  $\langle \text{Current, Second Wire} \rangle \rightarrow \langle \text{Explosion, Outside} \rangle$

Because such a chain exists, there is causation. Formally, Schaffer (2000, pg. 298) proposes the following relation for causation:

“C directly causes E iff (a) C and E are actual, distinct events, and (b) had C not occurred, then an (actual) connection to the E-manifestation<sup>10</sup> would not have occurred.”

Schaffer has argued that a modified version of a hybrid approach would be adequate for dealing with double prevention. He argues that such an approach would need to allow causation by absences and involve a counterfactual approach to connection. Such an approach would have to unify presences and absences into one form. Schaffer's suggestion for this is to consider the causal relata to be  $\langle \text{Property, Region} \rangle$  pairs that allow for absences as properties.

---

<sup>10</sup> The E-manifestation is the disjunction of a positive relata and the positive proxy of a negative relata. The positive proxy of the negative relata  $\langle \text{No Beer, Fridge Region} \rangle$  could be  $\langle \text{Water, Fridge Region} \rangle$ , to borrow Schaffer's example.

### 2.3.2.2. Hall

#### 2.3.2.2.1. A Hybrid Approach

Ned Hall (2004) argues for a different sort of hybrid analysis of causation. He argues that to solve the tensions of double prevention we ought to endorse two different concepts of causation, Dependence and Production.

He notes that one obvious response to the tension is to claim that counterfactual dependence is not causation. Counterfactual dependence is not by itself necessary for causation, and the best way we have of adding to it to create a full-blown analysis is to add the theses of Intrinsicness, Transitivity, and Locality. But as has been shown, double prevention shows that these theses work against counterfactual dependence being sufficient for causation rather than in its favour. So perhaps we should just recognise that the tensions are too much and give up counterfactual dependence being causation. The issue of course, is that counterfactual dependence is the driving force of counterfactual analyses of causation, and to give it up would be a hard bullet to bite.

Furthermore, it seems unintuitive when considering the usual ways we use causation. If we take two distinct events and note that had the first not happened then the second wouldn't have, then we can make all sorts of normal causal statements about it. The first event explains the second or is partly responsible for it. These just seem like appropriate things to say, and as Schaffer (2000) pointed out they are appropriate whether after the fact we discover it was a case of double prevention or not. So, it seems like counterfactual dependence must be causal.

Hall suggests that there is only one option available to us. There are two sorts of causation. The first is counterfactual dependence, and in this sort of causation Transitivity, Locality, and Intrinsicness are all false. But there is another sort, Production, where those theses are true.

Thus, we can resolve the problem cases with double prevention. They are simply one example of times where the two sorts of causation come apart. Cases of double prevention exhibit Dependence and so are causation in this manner. But from the issues

they expose in Transitivity, Locality, and Intrinsicness we can conclude that they do not exhibit Production.

If Halls argument is sound, then we have a good way of folding cases of double prevention into our understanding of causation where they have full causal powers. At least as far as Dependence goes.

#### 2.3.2.2.2. What is Dependence?

Now that we have pulled apart the concepts of Dependence and Production it is useful to see how exactly they might be spelled out. Dependence is the easier of the two and can be defined as simple counterfactual dependence however you might want to spell that out formally. A simple version might go something like ‘for any two wholly distinct events  $c$  and  $e$ , if had  $c$  not happened then  $e$  would not have happened, then  $c$  is a cause of  $e$ ’. It does not matter to Halls overall argument exactly how this gets spelt out.

As shown above, double preventers are clearly causes under Dependence. If Aimee had not shot down the enemy Fighter, then Carl would not have successfully bombed the target. So, the successful bombing is counterfactually dependant on Aimee shooting down the enemy Fighter. Thus, Aimee shooting down the enemy Fighter is a (Dependence) cause of the successful bombing.

#### 2.3.2.2.3. What is Production?

Hall has a harder job trying to define Production. He ends up endorsing a speculative attempt of fleshing the concept out but stresses that his attempt is just that, speculative. He notes that the analysis he offers needs further defence. What is really important to Hall is just that there is some plausible way of spelling out an analysis of causation on which Transitivity, Locality, and Intrinsicness are all true. This analysis would have to not rely on Dependence due to the tension between that thesis and the others. If such an analysis plausibly exists, then we do not have to think of Hall’s appeal to the existence of Production as an appeal to an unanalysable causal primitive.

The main thrust behind Hall's analysis of causation by Production is that some cause  $c$  is a Productive cause of effect  $e$  just in case it is part of some set  $S$  that is minimally sufficient for  $e$ . Often it seems like the problem cases for causation involve components that are separate from the 'standard' case of causation where an effect straightforwardly follows from its causes (Hall, 2004, pg. 260). For example, in the case of prevention, there is some standard causal chain that is infringed upon by the preventer. It doesn't make much sense to talk about some preventer preventing an effect if that effect were not going to be produced had the preventer not obtained. Cases of double prevention are similar, with a standard causal process on one hand and a double prevention process off to the side. In the diagrams the standard causal process is along the bottom and the infringing 'problematic components' are on top.

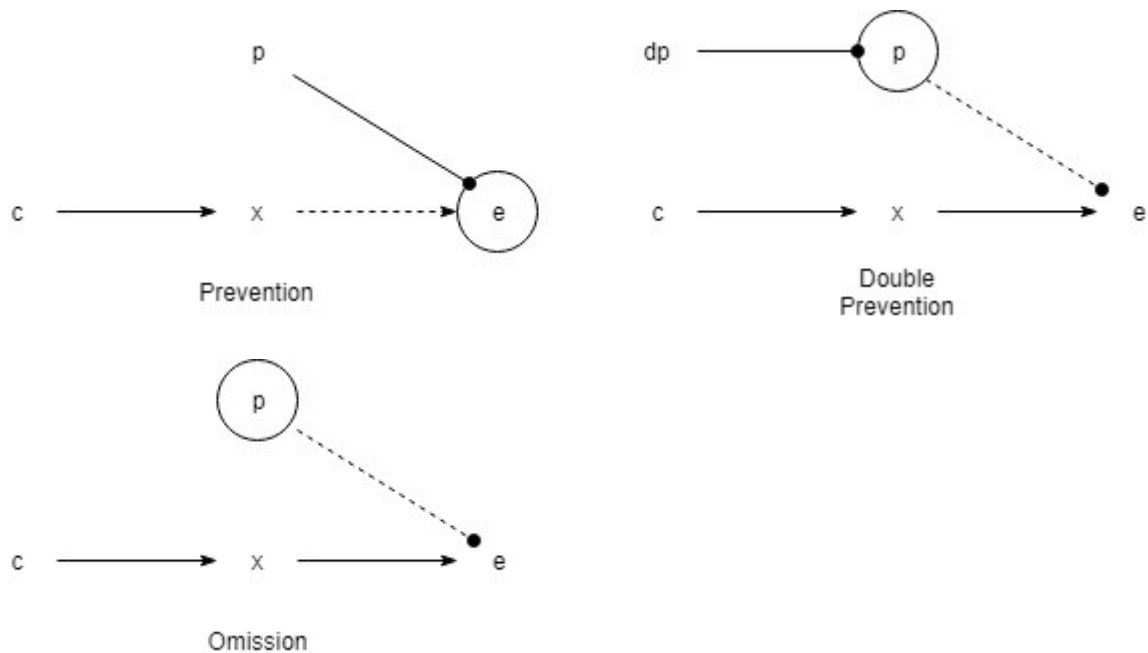


Figure 5

What Hall is trying to capture in his notion of Production is these 'standard' or 'nice' causal connections, relegating the others to causation by Dependence. Although he ends up qualifying it somewhat, Hall (2004, pg. 261) thinks that the minimally sufficient set of causes at some time  $t$  will contain all of the Producers of  $e$  and only those Producers.

On Hall's picture,  $S$  is sufficient for  $e$  just in case  $e$  follows from any laws, the premise that all of the members of  $S$  are present at  $t$ , and the premise that no other events occur



at  $t$ . Presumably though, any sort of analysis of sufficiency will work here.  $S$  is minimally sufficient just in case there is no proper subset of  $S$  that is also sufficient for  $e$ .

On our case of double prevention above, the set  $\{x\}$  is minimally sufficient for  $e$ . It is sufficient because if  $x$  occurs and no other events do, then  $e$  will follow. It is clearly minimally sufficient because it only contains one member. Thus,  $x$  is a Producer of  $e$ . The set  $\{dp\}$  is not minimally sufficient for  $e$ . In the case that  $dp$  obtains but  $x$  does not,  $e$  does not follow. Thus,  $dp$  is not a Producer of  $e$ .

This notion of Production also excludes cases of omission and prevention. Omissions are excluded for a similar reason to double preventions.  $\{x\}$  is the minimally sufficient set for  $e$  and the set  $\{\sim p\}$  is not part of any minimally sufficient set for  $e$ . Preventions are excluded because the minimally sufficient set for  $\sim e$  is nothing happening at all and so any potential  $c$  cannot be part of the minimally sufficient set.

Hall has so far conducted a first pass of analysing Production as the 'nice' parts of the causal structure. His original aim was for this analysis to be consistent with Transitivity, Intrinsicness, and Locality. Hall does not explicitly say how Locality fits in to the above picture, but he seems to take it as given that 'nice' varieties of causation will be consistent with it. Transitivity and Intrinsicness both extend the reach of the above analysis. Suppose that the above analysis is silent as to whether  $c$  is a Productive cause of  $e$ . We can apply Intrinsicness to see if that helps resolve the issue. We can look at a world with the same intrinsic character, if  $c$  is a Productive cause in that world then it is in this one. The same goes for Transitivity, if  $c$  is a Productive cause of  $d$ , and  $d$  is a Productive cause of  $e$ , then  $c$  is a Productive cause of  $e$ . By implementing these theses, Hall takes his analysis of direct Productive causation to a full analysis of Production.

The addition of Transitivity to Hall's analysis creates some issues for the simple version of his analysis that prompts him to suggest a more sophisticated version. The problem is that there are cases where there are Producers that do not belong to the minimally sufficient set, and there are cases where there are non-Producers that do belong to the minimally sufficient set.

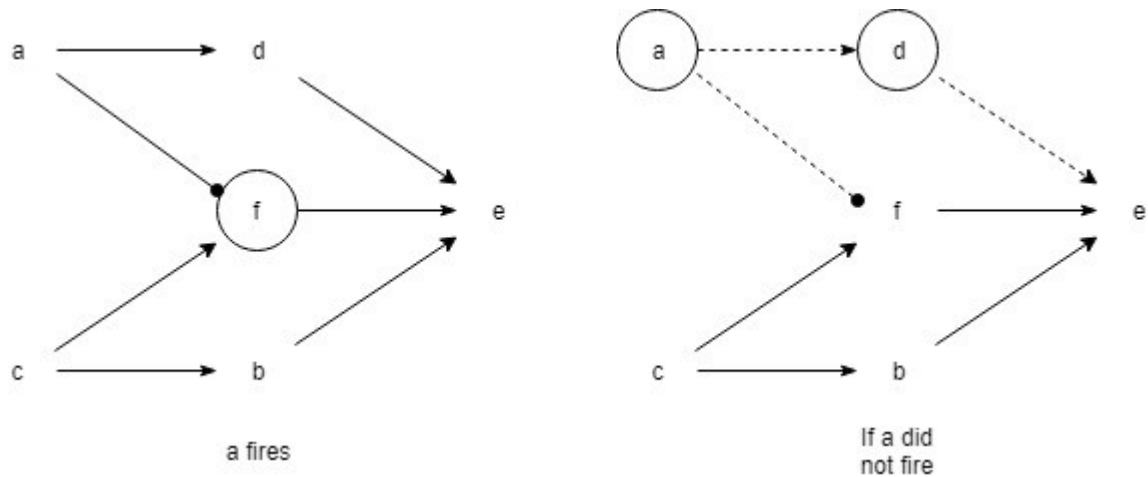


Figure 6: Adapted from Hall (2004), pg. 263, Figure 9.12

In this case, *e* is a stubborn neuron that will only fire with two stimulatory inputs. *a* and *c* ought to both be Producers of *e*. *d* is part of the minimally sufficient set at its time of occurrence for *e* and so is *a* Producer of *e*. *a* is part of the minimally sufficient set for *d* and so is *a* Producer of *d*. So, by Transitivity *a* ought to be a Producer of *e*. However, the minimally sufficient set for *e* contains just *c* as evidenced by the case where *a* does not fire. So, it seems that *a* is not a Producer of *e*.

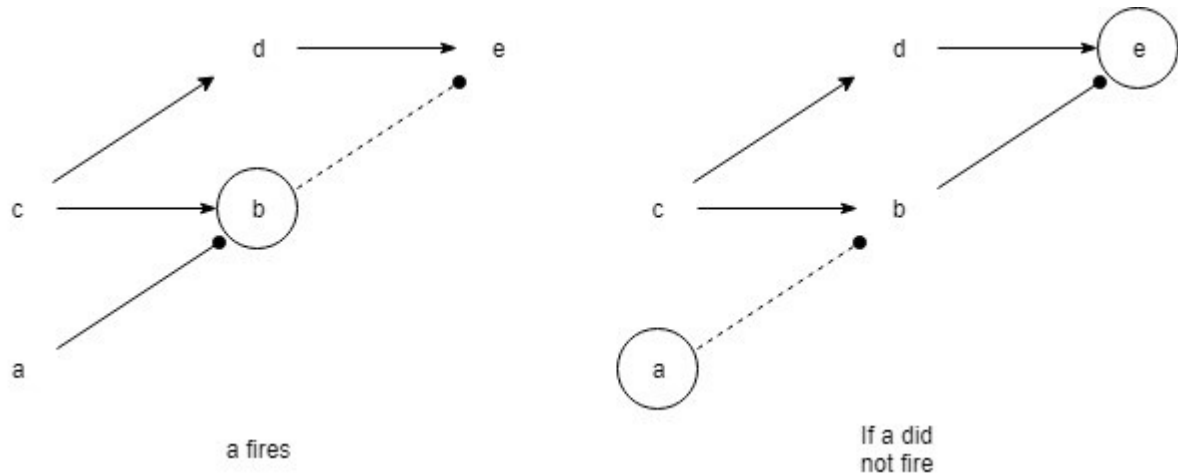


Figure 7: Adapted from Hall (2004), pg. 264, Figure 9.13

Part of Hall's appeal to Production was an attempt to single out the 'nice' cases of causation and omit cases such as double prevention. *a* is not the sort of cause that he was wanting to capture in the analysis. But, in this special case of double prevention, *a* is part of the minimally sufficient set for *e* at the time of its occurrence.

Hall suggests that we can deal with these issues by building in constraints to his analysis of Production. Suppose that *e* is at time  $t'$  and that we are investigating minimally

sufficient sets for it at various previous times  $t$ .  $t_0$  and  $t_1$  are two such previous times that have corresponding minimally sufficient sets  $S_0$  and  $S_1$ . Hall (2004, pg. 265) completes his analysis of Production by adding the requirements that:

1. For each element of  $S_1$ , there is at  $t_0$  a unique minimally sufficient set for it; and
2. The union of these minimally sufficient sets is  $S_0$ .

Hall is saying that when we are looking at the minimally sufficient sets of two different times we ought not simply identify the minimally sufficient set at the earlier time and go with that. Instead, we should identify the minimally sufficient sets of the components of the minimally sufficient set at the later time and the union of those sets is the minimally sufficient set at time  $t_0$ . In our first problem case we should start by identifying the minimally sufficient set at time  $t_1$ . This is the set  $\{b, d\}$ .  $b$  has the minimally sufficient set  $\{c\}$  and  $d$  has the minimally sufficient set  $\{a\}$ , the union of these is the set  $\{a, c\}$  and so both  $a$  and  $c$  are Producers of  $e$  at time  $t_0$ . Thus, the adapted theory does what we want it to do and counts  $a$  as a Productive cause of  $e$ .

In our second problem case, the minimally sufficient set for  $e$  at  $t_1$  is  $\{d\}$ . The minimally sufficient set for  $d$  is  $\{c\}$  and so the only Productive cause for  $e$  at  $t_0$  is  $c$ . Again, this is consistent with not allowing double prevention as Production.

Hall has presented an analysis of Production as a form of causation that is consistent with Transitivity, Intrinsicness, and Locality. It is not clear that this is the right analysis, but Hall does not need to achieve that. His main commitment is to the idea of Productivity rather than this particular analysis of it. What he has shown is that we can spell out the notion in a plausible way.

#### 2.3.2.2.4. Dependence and Production

Hall has provided a picture on which we can make sense of cases like double prevention as being fully causal even though they expose tensions between Dependence and the supplementary theses of Transitivity, Intrinsicness, and Locality. He has appealed to a hybrid analysis of causation on which Dependence is one sort of causation and Production is another. Under normal circumstances our notions of Dependence and

Production will coincide. However, it is precisely in the problem cases that they come apart, and that is in fact why they are problem cases.

## 2.4. Some Tentative Conclusions

Now that we have looked at several available analysis of double prevention we are in a position to offer some tentative conclusions about them. I say tentative because the literature on double prevention is underdeveloped and the analysis developed so far only offer a explanations of double prevention as it pertains to a few theories of causation. In addition, I say tentative because there is no clear answer to which analysis is right. All of the options are plausible. But they each force us to make concessions and assumptions about how causation works that we might not want to make.

### 2.4.1. Double Prevention is Not Causation

The Powers Theory provides a place for double prevention within the overall framework that is clear and consistent with the picture of causation it provides. Gibb is right in claiming that the positive account the Powers Theory can provide of double prevention is a positive of the theory. However, the Powers Theory is not a popular theory because of its commitment to simultaneous causation and non-transitivity. To accept both of these theses would be a major departure from the pictures of causation endorsed by most people. Even though the picture of double prevention on the Powers View is an appealing one I doubt it is enough to convince people skeptical of the Powers View to abandon their current commitments with respect to causation. To be fair, it is not trying to. Additionally, insofar as the overall picture presented by the Powers View strives to be a compelling one, its attitude towards double prevention helps.

Where I think the response of the Powers View shines is as a template for others who are skeptical of double prevention as being fully causal. The main thrust of their argument is that double prevention is causally relevant without being actually causal. This provides support to our intuitions that double preventers are important without having to give up other intuitions about causation proper. I imagine that other who are

skeptical of double prevention, such as proponents of connection views of causation, those who deny causation by absences, and counterfactual theorists who reject Schaffer's and Hall's hybrid theories, could make a similar distinction. Davidson (1967) and Beebe (2004, pg. 301-304) argue for a distinction between causation and causal explanation that fits nicely with this move, although they do not relate it to double prevention.

#### 2.4.2. Double Prevention is Causation

Schaffer and Hall also both present plausible places for double prevention in our causal metaphysics. Both present different sorts of hybrid accounts of causation. Schaffer's view is a hybrid in virtue of trying to find the middle ground between connections based and conditions based views of causation. Hall's view is a hybrid because it proposes two different sorts of causation. Dependence and Production. Like the analysis of the Powers View, while each is plausible I think that they each fall short of being really convincing. Each proposes a major departure from the usual theories of causation that their views are based on. They argue that our current theories do not allow double prevention as causation, and that double prevention is causation, so we ought to give up those theories. But it is just as legitimate to reaffirm our current theories of causation and instead deny that double prevention is causation. Double prevention might seem like causation, but on our best theories it is not. Too bad for double prevention. What Hall and Schaffer both propose is that we ought to be wary about double prevention because it is in tension with widely held theses about causation. However, their solution to that problem is to propose an analysis of causation that is in tension with widely held theses about causation. They have not really made any progress then. We still face the same issue.

One prime example of this is that both Schaffer and Hall are committed to causation by absences. Schaffer appeals to it explicitly. Hall is committed to it insofar as causation by absences follows for Dependence. But whether we should allow causation by absences is a deeply controversial debate in the causation literature. We might have strong intuitions about double prevention, but people have strong intuitions about causation

by absence as well. The arguments in favour of double prevention as causation are not strong enough to unroot these. This is especially true because there is an alternative. We can follow the general strategy proposed by Mumford and Anjum (2009) and count double prevention as being causally relevant. This allows us to go some way towards counting double prevention as important while still allowing our usual commitments in the metaphysics of causation.

As it stands there has not been enough debate on double prevention to determine if double prevention should be causal or causally relevant. Certainly though, there should be a group of people who wish to resist the claim that it is. Those who want to hold on to non-modified versions of the most popular analyses of causation should think of double prevention as causally relevant but not fully causal.

### 3. Double Prevention and Mental Causation

#### 3.1. Gibb's Model of Double Prevention and Mental Causation

##### 3.1.1. Introduction to Gibb

In several recent papers, Sophie Gibb (2013, 2015a, 2015b, 2015c) proposes the ingenious idea that we can use double prevention to solve the causal exclusion argument. She argues that if we take the relation between mental and physical events not to be straightforwardly causal, but instead one of double prevention, then the causal exclusion argument is invalid. If we accept this sort of view of double prevention, then mental events can be causally relevant to the physical without violating any of the premises of the causal exclusion argument.

Her argument consists of presenting a test case where the causal exclusion argument does not hold up. If there is one case where the causal exclusion argument breaks down, then it is likely that it breaks down for all cases. In this section I will first explain the sort of case that Gibb appeals to. I will then show how that case renders the causal exclusion argument invalid.

Gibb is a dualist and proposes her model with this in mind. She does not think that her model is available to a non-reductive physicalist due to issues with Distinction. In explaining why her model is a counterexample to the causal exclusion argument I will also show that her reasoning for thinking that it is not an option for the non-reductive physicalist is flawed. This means that her model should be appealing to a wider audience than she anticipated.

Gibb also endorses the Powers Theory of causation. The Powers Theory is not the most popular view of causation. Gibb endorses the version of the Powers Theory espoused by Mumford and Anjum (2011). This version claims that causation is simultaneous and non-transitive. I will show that Gibb's general model can be endorsed without endorsing the Powers Theory and so holders of other views of causation will still be able to use this model.

So far, so good. However, Gibb's argument deals with issues concerning Closure by appealing to the existence of a prior neural event that is a sufficient cause for every

mental event in question and the corresponding physical event. The existence of this prior neural event precludes libertarian free will. Libertarians often hold that the mental must be the initiator of a causal chain for us to have free will. The existence of a prior neural event that sufficiently causes all of our mental events means that this can never be the case. Such a neural event is supported to an extent by the research of Libet (1985), which Gibb appeals to in order to motivate her claim. However, Libet's research is controversial. If we can modify Gibb's proposal such that it need not rely on such a premise, then it would have a wider appeal. I will do just that.

With some minor tweaking we can take the essence of Gibb's proposal and make it less controversial. After having done this I will consider some objections that have been raised regarding it and respond to them.

### 3.1.2. What Gibb Needs to Show

Gibb's proposal is targeting the validity of the causal exclusion argument. In order to show that it is invalid it is sufficient to have a case in which mental events are causally relevant (refuting epiphenomenalism) and all of the other premises of the argument hold true. Before moving on let us recall our formulation of the causal exclusion argument quickly:

1. *Distinction*. Mental events are not physical events.
2. *Physical Causal Closure*: Every physical event contains only other physical events in its transitive causal closure.
3. *Exclusion Principle*: No single event can have more than one sufficient cause occurring at any given time – unless it is a genuine case of overdetermination.
4. *Denial of Overdetermination*: Mental causes do not overdetermine their effects.
5. *Therefore: Epiphenomenalism*. Mental events have no causal role with respect to physical events.



### 3.1.3. Gibb's Proposed Case

Gibb's proposal is to straightforwardly map a potential case of mental causation to the broader idea of double prevention. We can imagine that Carl (on break from flying bomber missions) happens to see his friend Aimee across the airstrip. He wants to get her attention and forms the desire to wave his hand at her to do so. He then does in fact wave his hand at her. A seemingly normal case where it appears that we have mental causation. Just as susceptible to the causal exclusion argument as any other.

However, we can think of this example in a different way. In terms of double prevention. In the following discussion let n1 refer to the neuron N1 firing, let n2 refer to the neuron N2 firing. Additionally, let b1 refer to Carl waving his hand and b2 refer to him keeping it still, while m1 is Carl's desire to wave his hand and m2 the conflicting desire to keep his hand still. We can plug this in to a double prevention picture as follows:

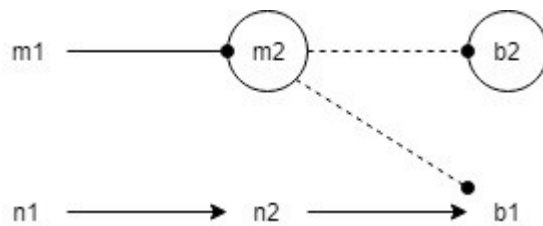


Figure 8

In this case, Carl's hand waving (b1) has both n1 and n2 in its causal chain. This just represents the physical causal chain leading to his hand waving. If Carl had the desire to keep his hand still, then it would be the case that that desire would prevent n2 from causing Carl's hand to wave (b1) and instead his hand would stay still (b2). Luckily for Carl, he does have the desire to wave his hand, which prevents the desire to keep his hand still from obtaining, and so the causal process from n2 to b1 can continue unimpeded. Carl waves at Aimee.

Now that we have set up the case that Gibb proposes serves as a counterexample to the causal exclusion argument we can turn to investigating if it succeeds in that role.

### 3.1.4. Gibb and the Powers Theory of Causation

Gibb endorses a Powers Theory of causation to advance her claim. This is important to her counterexample, because it allows her to accept all the premises of the causal exclusion argument. This is most obvious with Closure. However, as I will soon argue, acceptance of the Powers Theory is not essential to accepting her claims.

Is there a breach of Closure with respect to b1? If a physical effect has a mental cause, then Closure is broken<sup>11</sup>, because there would be a non-physical event in the transitive causal closure of a physical event. This is a problem for those who endorse a view of double prevention on which double preventers are fully causal. That is, if one endorses the views of Hall or Schaffer as discussed earlier. If double preventers are fully causal, then we can simplify a double prevention relation into a simple causal relation. Carl's desire to wave his hand would be a cause of his hand waving. If we take Gibb's proposed case and simplify it in line with this then we get the following:

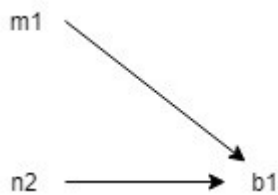


Figure 9

In this simplified case, there is a physical effect (b1) that has a mental cause (m1). This is a breach of Closure. If such a breach stands then clearly Gibb's case cannot solve the causal exclusion argument.

Gibb endorses a Powers Theory of causation to get around this worry. On the Powers Theory, double preventers are not causes (Gibb, 2013, pg. 200. Mumford and Anjum, 2009, pg. 286-287). Instead, double preventers are merely causally relevant. If we accept that double preventers are not causes, then the problem just addressed does not worry us. Because mental events act as double preventers they cannot be causes. This means that there is no danger of m1 causing b1, and no worry about a breach of Closure in this

---

<sup>11</sup> Assuming *Distinction*, that mental events and physical events are different.

regard. The only straightforward causes of  $b_1$  are physical ones, namely  $n_1$  and  $n_2$ . This is entirely consistent with preserving Closure.

Gibb is committed to the Powers Theory of causation. However, any theory which does not count double prevention as fully causal should work with her overall case. The specifics of the Powers Theory are not doing any work here (although it is a plausible response to the tensions of double prevention). In order to avoid a breach of Closure there just cannot be a causal link from a mental event to a physical one. Such a link will not occur if we accept any view where double preventers are not causal because the double prevention relation stands on its own. This would also be the case if we accepted a Process view of causation, as discussed earlier. Gibb's model then, should not just be of interest to people endorsing the Powers Theory. Anybody who is skeptical of the full causal power of double preventers could potentially make use of it<sup>12</sup>.

The Powers Theory is not a popular theory of causation. By divorcing Gibb's reply from her endorsement of the Powers Theory it opens up the reply to be endorsed by those who endorse other theories. Even if we do not want to commit to any particular theory of causation, there are good reasons to think that causation is not simultaneous (see Maslen, 2018) and transitivity of causation is widely held<sup>13</sup>. The Powers Theory denies the transitivity of causation and upholds simultaneous causation. By arguing that Gibb's reply does not need to endorse the Powers Theory, I have opened it up to be endorsed by people who reject these claims in particular.

### 3.1.5. Is This a Counterexample to the Causal Exclusion Argument?

Bearing a non-causal view of double preventers in mind, we are now able to take a first swipe at the causal exclusion argument. We can go through the premises and conclusion one by one and see if they are acceptable.

---

<sup>12</sup> Accepting this view of double prevention where it is not fully causal is not enough by itself to accept Gibb's full picture. You would also have to accept her empirical picture of conflicting desires.

<sup>13</sup> Lewis for example (1986, pg. 167) simply asserts "[c]ausation must always be transitive"

*Epiphenomenalism.* Mental events have no causal role with respect to physical events.

In the above scenario epiphenomenalism is false. Although exactly what causal role mental events play could vary based on what theory of causation you endorse. Any explanation of why Carl waved his hand would be remiss in excluding his desire to do so from the contributing factors.

If we merely focused on the physical chain of causes that lead up to Carl waving his hand, then we would have the complete set of causes of his hand waving. However, this would not be the whole picture. It is in virtue of the presence of his desire to wave his hand that the causal link between the neurons firing and Carl's arm waving is able to obtain. The double preventer is causally relevant, even if it is not a cause. This is enough to show that epiphenomenalism is false (in the case presented). Mental events *do* have a causal role to play with respect to physical events, that role is just not as a straightforward cause.

*Distinction.* Mental Events are not physical events.

Gibb is a dualist. If we follow her then in this case mental events ( $m_1$ ,  $m_2$ ) are distinct from physical events ( $n_1$ ,  $n_2$ ,  $b_1$ ,  $b_2$ ). The diagrams show that the mental events have different causal powers from the physical ones and so they cannot be identical. So, in the scenario above, Distinction holds true.

Gibb (2013, pg. 205) argues that her model will not appeal to the non-reductive physicalist. She argues that most accounts of non-reductive physicalism are committed to a picture of realisation where the causal powers that realised events have are a subset of the causal powers that the realiser event has. This picture of realisation is known as the subset view of realisation and defended notably by Shoemaker (2001, 2007, 2013) and Wilson (1999, 2011). Furthermore, says Gibb (2013, pg. 2005), this is not the case for  $m_1$  and  $n_1$ ; " $m_1$  has a causal power that  $n_1$  does not have, namely, the power to affect  $m_2$ ". If this is right, then  $n_1$  cannot be the realiser of  $m_1$ .

However, it is unclear that most non-reductive physicalists are committed to this sort of picture as she assumes. Baysan (2015) identifies four different sorts of realisation that defended in the causation literature, of which the subset view is only one. Without further argument, Gibb's model should be open to those who endorse a different view.

Additionally, even if we do support the subset view we should be able to accept the model. A key part of Gibb's argument is that we accept a view of causation where double prevention (and thus prevention) is not fully causal. So even on the subset view  $m_1$  does not have any causal powers that  $n_1$  does not have. The link from  $m_1$  to  $m_2$  that Gibb appeals to in order to dismiss non-reductive physicalism is not a causal one.

It should follow straightforwardly from non-reductive physicalism that the mental and physical are distinct. If that was not the case, then it would instead be a form of reductive physicalism. Given that, they ought to accept this premise. There is nothing about accepting this model that forces the non-reductive physicalist to give that commitment to Distinction up and so non-reductive physicalists can make use of this model.

*Exclusion Principle:* No single event can have more than one sufficient cause occurring at any given time – unless it is a genuine case of overdetermination.

Endorsing the Exclusion Principle is another case where our commitment to a theory of causation that does not permit double prevention to have full causal powers does some work for us. If we endorse an analysis of causation that does grant double prevention full causal powers, then we can change the double prevention relation in the diagram to a simple causal one. This means that we are again left with the following simplified picture of the case:

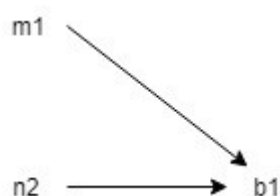


Figure 10

Taken at face value, it seems like Exclusion is not met. We are already treating m1 and n1 as distinct events. So, there are two causes of b1 that superficially both seem sufficient and Exclusion requires that there be only one sufficient cause. Possibly there is room to argue that m1 and n1 together represent a sufficient *total* cause. In any case, endorsing the view of double prevention that we do means that we do not have to worry about this issue.

If this is the case then Gibb 2013, pg. 207) notes that we cannot have any problem with Exclusion because m1 is not a cause and therefore cannot overdetermine any physical effect. There is only one cause and so there can only be one sufficient cause.

*Denial of Overdetermination:* Mental Causes do not overdetermine their effects.

There is no overdetermination at play in the above case. If double prevention is not causal, then there is only one cause. If there is only one cause, then there cannot be overdetermination.

*Physical Causal Closure:* Every physical event contains only other physical events in its transitive causal closure.

Closure is the most difficult premise for Gibb to show holds true under her case. We resolved some of the potential issues of Closure between m1 and b1 earlier by assuming a theory of causation that does not count double prevention as fully causal. However, this is not the only potential issue that we have with Closure.

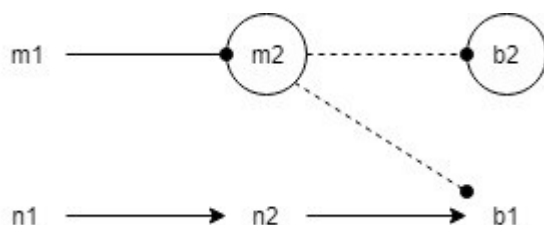


Figure 11

If we look at our diagram again all of the obtaining physical events contain only other physical events in their transitive causal closure, just as this premise requires. However,

even though in this world there is not actual breach of Closure there seems to be a large potential for there to be, in the prevented causal link from  $m_2$  to  $b_2$ . We can easily imagine the case where  $n_1$  fires but  $m_1$  is not there to prevent  $m_2$ . In this case, Carl does not have the desire to wave his hand and this directly causes his hand to stay still. This is a breach of Closure as a physical effect will have a mental event in its causal history.

To avoid this problem with Closure, Gibb (2013, pg. 205) argues that it must be the case that every time  $n_1$  obtains it must also be the case that  $m_1$  obtains. In that case there could be no breach of Closure between  $m_2$  and  $b_2$  because such a causal link would never happen. Although there would still be a slightly worrying potential breach of Closure, it would not matter because the breach would never actually obtain.

In order to make it the case that every time  $n_1$  obtains,  $m_1$  also obtains, Gibb proposes the existence of  $n_0$ .  $n_0$  is another neurological event that is a sufficient cause of both  $n_1$  and  $m_1$ .

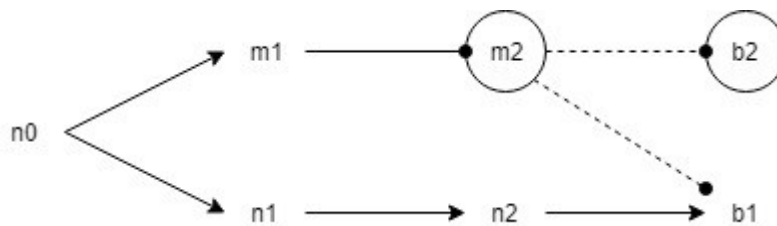


Figure 12

If  $n_0$  is a sufficient cause of both  $n_1$  and  $m_1$ , then there can be no case where  $n_1$  obtains without  $m_1$  also obtaining. Meaning that there is no chance that the problematic causal link from  $m_2$  to  $b_2$  obtains. If this link does not obtain, then there is no actual issue with Closure in Gibb's proposed case.

The existence of  $n_0$  is a broadly empirical claim. It either is or is not the case that our mental events are caused by specific neural activity. I will turn to Gibb's defence of that claim shortly. However, if it turns out to be true that  $n_0$  does exist, then Closure is true in Gibb's case.

Gibb has provided a counterexample to the causal exclusion argument. Mental events act as double preventers of physical events and so remain causally relevant to them.

Thus, epiphenomenalism is false. Furthermore, all of the other premises of the causal exclusion argument can still be held true under Gibb's case. I have explained how each of these can still be held in turn. The biggest issue was with Closure. However, if we posit the existence of  $n_0$  then we can deal with this issue too. The issue for Gibb now is to argue that the existence of  $n_0$  is likely. It is to that argument I now turn.

### 3.2. Should we Believe in $n_0$ ?

The existence of  $n_0$  is an empirical claim. For Gibb's claims about  $n_0$  and its effect on closure to be true it would have to be the case that all of our mental events (or at least all of the causally relevant ones) are sufficiently caused by a neurological event that also sits in the physical causal path of the resulting physical effect. If empirical evidence suggested otherwise, then her account would be straightforwardly false as it stands.

One worry about the existence of  $n_0$  is that it precludes libertarian free will. Standard accounts of libertarianism are committed to the agent being a causal initiator. But if this were the case, then  $m_1$  should start a causal chain. On Gibb's account, it does not. Instead the causal chain is started at  $n_0$  or some previous physical event. If all of our desires are prefaced by a neurological event, then our free will does not have the important role of instigator that we often suppose it to have. It would be preferable to find an alternate solution that does not place undue restrictions on our conception of free will. In particular, if we do not have to endorse the existence of  $n_0$ , then we open the reply up for dualist libertarians.

Gibb looks to the famous experiments by Libet (1985) in order to support her claim. Libet conducted a series of experiments in which he asked subjects to flick their wrist or hand and note down the time when they felt the urge or intention to do so before the act. His experiments found that the subjects felt the urge to act on average 190ms before the act took place. However, activity in the brain (in this case called "readiness potential") had an onset on average 535ms before the act. The onset of the readiness potential was 345ms *before* the conscious desire to act (Libet, 1985, pg. 532). These experiments lead Libet (1985, pg. 536) to conclude that "the performance of every conscious voluntary act is preceded by a special unconscious cerebral process that



begins about 500 ms or so before the act". In other words, that all of our voluntary mental events are initiated by a neurological event. This is great for Gibb, Libet's experiments provide exactly the empirical backing that she needs for her claims about  $n_0$  to be true.

However, Libet's claims are far from universally accepted. Daniel Dennett (2003a, 2003b) for one argues that Libet's results are mistaken due to a misunderstanding about where the self exists. We will have to adjust the timings of the various parts of the experiments based on how long the information takes to get to *us*. But the 'we' doing the work could feasibly exist in any part of our brain (or all of it). Libet's experiments do not prove anything one way or the other, because we cannot access the proper data.

Another argument against Libet is advanced by Mele (2015). Mele presents a multipronged attack on Libet's reasoning. Firstly, Mele claims that there is no particular reason to think that our decision is made at the time of the onset of readiness potential. An experiment conducted by Haggard and Magno (1999) participants were instructed to perform an action after receiving a signal to do so. The mean time between the signal and the action was 231ms. This seems to indicate that the decision to perform the action must have happened less than 231ms before the action, not 535ms as Libet claims. Libet's time of the urge to act being 190ms before the action is in line with this finding and so it lends support to this urge being the decision-making point.

Secondly, even if the participants in Libet's experiment were not making conscious decisions, we cannot generalise the findings to show that we never do. In Libet's experiment the subjects were asked to flex their wrists as they wanted with no reasoning process weighing up pros and cons to the action. Even if there was no free will showcased in the experiment it does not mean that we would not have it in more complex cases.

Lastly, Mele asks why it should matter that there is some lag between a between when a free decision is made and when we become aware of it. Perhaps the decision *is* made at the onset of readiness potential. If we undergo a conscious decision-making process that leads to a decision and then an action then surely that action was made freely, even if it takes us a small amount of time to be aware that we made the decision.

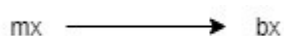
It is well outside my area of expertise to provide a compelling argument against Libet. However, the existence of arguments against his position (for example Dennett and Mele) and the fact that it places restrictions on our conception of free will limit the appeal of any analysis that relies on it. If we can find an adaption of Gibb's case that does not rely on Libet's experiments, then that would be preferable. The issue of course is Closure. Without Libet the existence of  $n_0$  loses much of its empirical backing. But without  $n_0$  there will be a major potential issue with Closure on Gibb's model. In the following section I will propose an alternate way to deal with the issue of Closure using Gibb's case. One that does not rest on a controversial understanding of the world.

### 3.3. My Adapted Proposal

#### 3.3.1. First Pass

When considering alternate ways for Gibb to deal with the issues caused by Closure I think that it is best to step back and look at Gibb's overall project. What Gibb has done is propose that we turn what we intuitively thought of as straightforward mental to physical causation into a more complicated relationship involving double prevention.

Our naive approach to mental causation is simply to assert that mental events can be causes of physical ones. If you were to go up to somebody on the street and ask them if their desire to wave their hand caused their hand to wave then an overwhelming majority will say that it does. This is what makes the causal exclusion argument so appealing. Of course, we always have the option of accepting the argument and giving in to epiphenomenalism. But, for most people, our intuitions about mental causation are so strong that this is deeply unappealing. Our intuitions point to the following relationship being the case:



*Figure 13*

When confronted with additional intuitive premises about the physical world we take on the following, slightly more complicated picture:

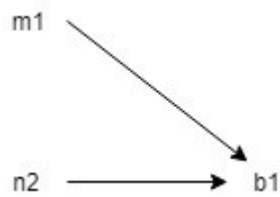


Figure 14

Gibb's project (before n0) has been to suggest that this picture is mistaken. But mistaken in a way that is compatible with our original intuitions that mental events are causally relevant. This gives us the following:

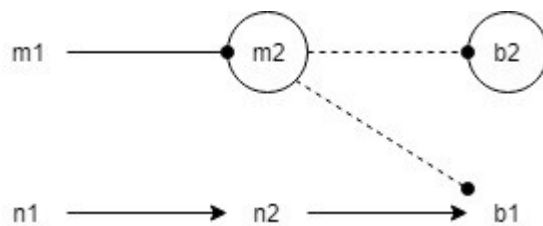


Figure 15

A slightly different way of looking at her proposal is to say that in every case where we thought we had a mental to physical causal link, we are actually mistaken. That link should instead be thought of as a double preventer. By undertaking this process, we take problematic mental to physical causal links and transform them into non-problematic double preventers. The only problem with this move is the existence of a potential, but not realised, problematic mental to physical causal link.

But Gibb has just given us a way of dealing with those links to make them unproblematic. Change them to be double preventers. We can do this to the link between m2 and b2 giving us the following expanded diagram:

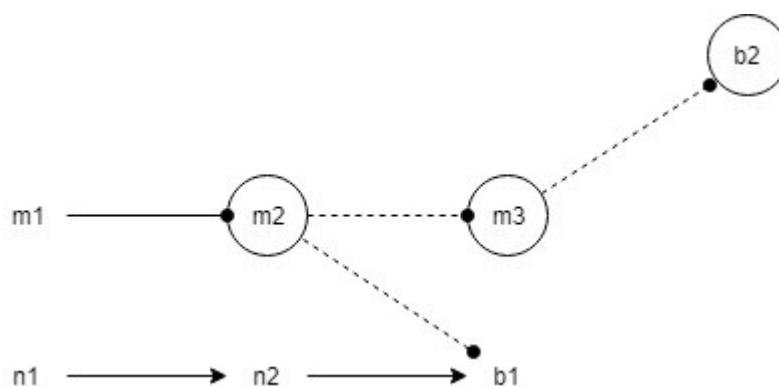


Figure 16

This diagram is largely the same as Gibb's. There is a physical chain of neurons firing that leads to Carl's hand waving. If Carl had the desire to keep his hand still, then this would prevent the firing of N2 from causing his hand to wave. However, in this case, Carl did have the desire to wave his hand, which prevented him from having the desire to keep it still, permitting the firing of N2 to cause his hand to wave.

Off to the side, we can note that had Carl had the desire to keep his hand still then this would have been causally relevant to the potential case of his hand staying still through a double prevention relation as well. Some desire to move his hand (m3) would have prevented him from keeping his hand still had it not been the case that this desire was itself prevented by m2 obtaining.

### 3.3.2. Is This a Counterexample to the Causal Exclusion Argument?

With this adjusted case in hand we can turn back to the causal exclusion argument and see if this is a good counterexample.

#### *Epiphenomenalism*

As with Gibbs case mental events are causally relevant to physical events by virtue of being double preventers. Carl's hand would not wave without him having the desire for it to do so. It seems clear that his desire has an important causal role.

#### *Distinction*

Again, mental events are clearly distinct from physical events. They have different causal powers.

#### *Mind-Body Supervenience*

This premise is still not affected by the case Gibb and I have proposed.

#### *Exclusion Principle*

Mental events are not causes. Every effect in the diagram only has one sufficient cause at any given time.

*Denial of Overdetermination*

Again, there is only one cause at any given time and so there cannot be any overdetermination.

*Physical Causal Closure*

Just like in Gibb's case there is no breach of Closure with respect to Carl waving his hand because mental events are not causes. There are only physical causes in the transitive causal closure of b1.

With respect to potential breaches of Closure regarding b2, this is no longer an issue. m2 cannot even potentially cause a breach of Closure here because there is no causal relation. Importantly, my adapted proposal deals with Closure without positing the existence of n0, some neurological event that precedes and causes both m1 and n1.

My adaption of Gibb's model serves as a counterexample to the causal exclusion argument. Specifically, it targets the validity of the problem and provides a case where all of the premises are true and the conclusion, epiphenomenalism, is false. I view my adaption as the natural extension of Gibb's thesis when taken to its full extent. By doing this Gibb's general proposal should be more appealing because it does not have to rely on controversial premises about whether mental events are directly caused by neural processes in the brain. However, there are some potential objections to consider, and I will turn to those now. I will start by looking at some minor objections and clarify my position in response to them. Then I will look at major objections levelled against Gibb by other authors and see if my particular version of the model can answer them.

### 3.3.3. Minor Objections and Clarifying Remarks

#### 3.3.3.1. What if I Like Libet?

My argument thus far has been to assert that we can provide a better version of the double prevention model if we remove its reliance on the controversial existence of n0 supported by Libet. However, Libet's work, while controversial, does enjoy a large

following. If you are a fan of Libet then you might want to reject my version of the model. There is no need to do this. My version is perfectly compatible with the existence of  $n_0$ , it just does not require it. In addition, my adaptations should be considered the logical extension of Gibb's proposal when fully applied and so even with the existence of  $n_0$  we should still think that they apply.

Consider my version of the model with  $n_0$  added:

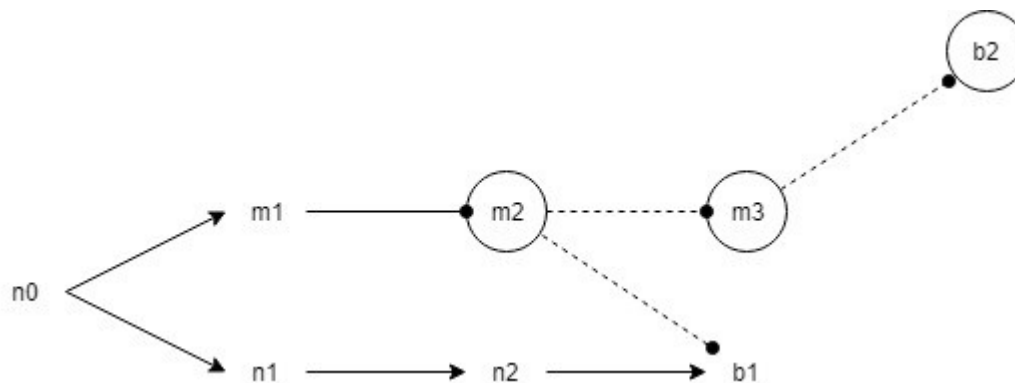


Figure 17

There is nothing inconsistent here and it still provides a counterexample to the causal exclusion argument. The double prevention relation between  $m_2$  and  $b_2$  will never be realised. Perhaps it is a bit redundant to draw it out in this way. However, as I explained earlier, I have just done to this relation what Gibb is proposing we do to all mental to physical causal relations. If we accept Gibb's overall model, then this is the sort of relation we should expect here.

### 3.3.3.2. Is it a Problem that there is a Hanging $b_2$ ?

You might be concerned that the double prevention relation from  $m_2$  to  $b_2$  exists without any supporting physical causes. There could be any number of different sorts of physical events attached to this relation in several ways depending on what position you want to take regarding how mental events map to physical ones. I have intentionally left them out for several reasons. Firstly, for clarity. Such physical events are ultimately clutter and do not contribute to my overall point that the causal relationship from a mental to physical event ought to be a double preventer. Secondly, because I did not want to commit myself to any particular version of how mapping mental to physical

events might work. To draw those links in would be to do this. My account is independent of this concern. Thirdly, because on my account this relationship does not actually obtain. To draw in a number of physical events that do not obtain seems pointless.

If you are worried about the physical events that would not occur and the way that they pertain to that relationship, then you are welcome to add them. I cannot see how they would affect my claim. However, I would welcome any objection if they do.

### 3.3.3.3. Is 'Causally Relevant' Enough to Defeat Epiphenomenalism?

You might object that mental events being *only* causally relevant is not enough to defeat the charge of epiphenomenalism. Formally you could endorse a version of epiphenomenalism that claims that mental events are not *causes* of physical events. If you endorsed this sort of epiphenomenalism, then the double prevention account would not be a good counterexample. On the double prevention case epiphenomenalism of this sort would be true, because mental events are *not* causes of physical ones, they are merely causally relevant.

I confess that I find this sort of statement of epiphenomenalism unappealing. On the informal version of the causal exclusion argument our worry is that our mental events are not causally relevant at all, cast aside by the physical causal chains. The double prevention model provides a causally relevant place for our mental events to do work and this seems to be enough to satisfy an informal account of the problem. If you want to endorse a stronger version of epiphenomenalism, then you will find this argument unappealing. But I think that most people will not and so the argument still retains wide appeal.

### 3.4. Major Objections

#### 3.4.1. Objection to Gibb's Formulation of Closure

There are many different versions of the Causal Closure Principle that are defended in the literature, and one objection to the double prevention model might be that although it works on Gibb's formulation of Closure it does not work on all of them.

Closure comes in weaker and stronger varieties. At the weaker end, it might be formulated as something like:

"Every physical phenomenon has a physical cause" (Montero, 2003, pg. 174)

Both Gibb's and my reply to the causal exclusion argument upholds this version of closure. Every physical event, except  $n_1$ , has a physical cause. Although it is not shown on the diagram, we can imagine that  $n_1$  has some causal history stretching back as far as we desire.

This formulation of Closure can be considered a weaker variant in that it allows for the possibility of a mental cause as well as the physical one. In a different way, this sort of formulation seems too strong for the physicalist to accept because it does not account for the possibility of uncaused physical phenomenon (the big bang for example).

In response to these worries the following formulation might be preferred:

"Every physical phenomenon that has a sufficient cause has a sufficient physical cause" (Montero, 2003, pg. 174)

This version of Closure is also not breached by the acceptance of our reply.  $n_2$  is the only physical event with a sufficient physical cause,  $n_1$ . That sufficient cause is a physical one.  $b_1$  does not have a sufficient cause. So, there is no worry that it does not have a sufficient *physical* cause.

This sort of formulation is still relatively weak because it allows that physical events might have a sufficient physical cause and also a non-sufficient non-physical cause. For example, some physical effect might have a sufficient physical cause that causes the effect through a non-physical causal intermediary. The above formulation would still be



met because there is a sufficient physical cause at some point in the causal chain. Gibb's formulation of Closure represents a stronger version of Closure:

"Every physical event contains only other physical events in its transitive causal closure" (Gibb, 2013, pg. 193)

Her version of Closure is a strong one because it *only* allows for physical causes in the causal chain. The case we had above where a sufficient physical cause brings about its effect through a non-physical intermediary would be a breach of Closure on this formulation. This strong formulation excludes non-physical causes altogether. As I have explained previously in this chapter, this version of Closure is met in our replies. Thus, our replies will still work on the assumption of quite a strong version of Closure.

All of the version of Closure I have mentioned so far conceive of it as acting in one direction. Physical effects are closed from mental causes. However, Kim (1996, pg. 147) endorses a two-way variant of Closure<sup>14</sup>:

"[N]o causal chain will ever cross the boundary between the physical and non-physical"

This formulation of Closure is met by my reply to the causal exclusion argument but not Gibb's. Gibb's version has a causal link from  $n_0$  to  $m_1$  which is an example of a causal link crossing the physical to mental boundary. On my version there is no  $n_0$  and so this is not an issue. My version of the reply has no cases where a causal link links a mental to a physical event or vice versa. There are prevention links, but the model is grounded on the assumption that prevention and double prevention are not causal. So, there is no breach of even this extremely strong version of Closure on my reply.

To formulate a version of Closure that my reply is in breach of you would have to endorse something like the following:

Every physical event has only physical events in its transitive causal Closure and has at least one sufficient physical cause.

---

<sup>14</sup> Although this is the variant of Closure he endorses in that book elsewhere he defends other formulations more in line with the formulation proposed by Gibb. For example, in *Physicalism, or Something Near Enough* (Kim, 2005, pg. 50) he defends a formulation closer to Gibb's, "Any cause of a physical event is itself a physical event".

However, it is hard to see what sort of empirical evidence could be offered for a version of Closure that is this strong. In particular, as was pointed out with our first formulation of Closure, a formulation such as this is too strong to explain uncaused physical events such as the big bang. Formulating Closure in this way also seems to implicitly assume determinism as every physical event has a sufficient physical cause.

I have shown that both Gibb's and my reply to the causal exclusion argument is compatible with a variety of formulation of Closure and that the version endorsed by Gibb to start with was a strong one. My version of the reply has the additional benefit of being compatible with a two-way formulation of Closure such as endorsed by Kim (1996). In order to find a version of Closure that my reply is not compatible with you would have to endorse an extremely strong version of the principle that would be hard to find support for.

### 3.4.2. Matteo Grasso

#### 3.4.2.1. Physical Double Preventers

Another objection comes from Matteo Grasso (2015). Grasso's claim is that when we add in some missing parts to Gibb's neural diagram the causal efficacy of the mental events goes away. Usually, in neural diagrams, mental events form a linked pair with a physical event. Physical events can stand on their own. In Gibb's diagram  $m_1$  and  $n_1$  represent this sort of pair. However,  $m_2$  is not linked to  $n_2$  in the same sort of way.  $m_2$  is a desire that is ultimately overridden while  $n_2$  is a connecting neural event. Insofar as  $m_2$  is a desire that obtains it ought to have a corresponding physical event  $m_1^*$ . This is going to be the case in at least some double preventions where there is competing desires. To preserve symmetry further there also ought to be a corresponding  $n_2^*$  linking  $n_1^*$  to  $b_2$ <sup>15</sup>.

---

<sup>15</sup> The vertical arrows in this diagram represent supervenience or realisation relations.

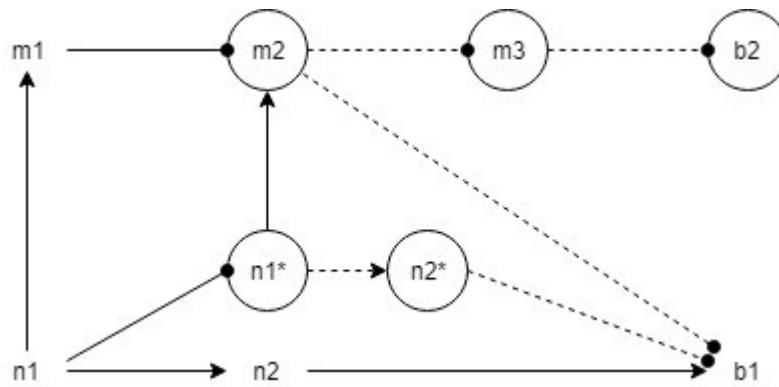


Figure 18

If we accept the amendments to Gibbs picture then Grasso thinks that not only does  $n1$  ultimately cause  $b1$ , it also prevents  $n2^*$  (by preventing  $n1^*$ ) from preventing  $n2$  from causing  $b1$ . If this is true, then the double prevention relationship that Gibb outlines is also present in the physical world.

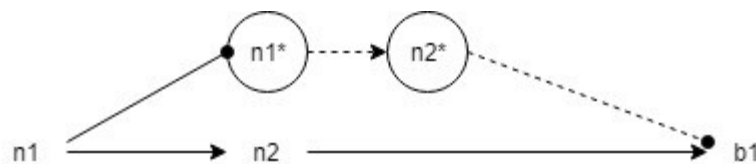


Figure 19

Under this picture, Grasso claims that mental events are again causally irrelevant. To claim otherwise would be to breach 'Denial of Overdetermination'. 'Denial of Overdetermination' claims that there is no systematic overdetermination of causal efficacy. If we accept this, we should also accept that there is no overdetermination of causal relevance. But Grasso has pointed out that there is overdetermination of causal relevance in Gibb's model. For every mental double preventer there is also a physical double preventer.

There are two ways to respond to Grasso's objection. Firstly, the dualist such as Gibb will straightforwardly deny Grasso's view of the world. In claiming that mental events always correspond to a physical event, Grasso is presumably appealing to a non-reductive physicalist notion of supervenience or realisation. A dualist is not committed to such a correspondence. If  $m2$  floats free of any physical event, then there need not be any corresponding physical double prevention path that Grasso appeals to. Despite his intentions, for the dualist, Grasso's objection holds no weight.

However, this is a plausible objection to the non-reductive physicalist endorsing this model. A non-reductive physicalist will believe that every mental event has a corresponding physical one. However, Grasso's jump from denial of causal overdetermination to denial of causal relevance overdetermination is undermotivated. We treat causes as a protected class with only so much causal work to go around. But there is no particular reason to think that the same is true of causal relevance. The key intuitive issue that the causal exclusion argument raises for mental causation is that there is no work left for the mental to do in causing some physical event. A sufficient cause usually excludes anything else from being a cause. But causal relevance is not protected in the same way. All sorts of things might be causally relevant for various reasons and that is simply the case that is happening with mental and physical double preventers.

Grasso's objection to my reply is in my opinion the most promising as an attack on how well the reply works for a non-reductive physicalist. Despite this, I do think that more work needs to be done to show that causal relevance of one sort excludes other causal relevance.

### 3.4.3. Kim Davies

#### 3.4.3.1. Implausibility of m2 Always Being Accompanied by m1

Kim Davies (2016) presents a number of objections against Gibb's proposal, and therefore also my own. The first three of these are concerns with how Gibb's model must work in order for her to uphold Closure. Davies (2016, pg. 40) first reaffirms the objection that lead to Gibb positing the existence of  $n_0$ . He notes that Gibb's model rests on the idea that mental events have the real possibility of causally intervening in the physical realm, it's just that they never do so in the actual world. Gibb (2013, pg. 207) puts this plainly, (her example refers to Fred not Carl) it is "naturally impossible for Fred's desire to manifest all it was disposed for... to prevent  $n_2$  from causing Fred's hand's moving".  $n_2$  has the causal power to prevent  $n_2$  from causing  $b_1$ . But this never happens in the actual world on Gibb's model because every time that the chain leading to  $n_2$  causing  $b_1$  is initiated,  $m_1$  also occurs due to  $n_0$ . On Gibb's version of the model all

of this has to be the case, because if there ever was a time where m2 did in fact prevent n2 from causing b1, then there would be a breach of Closure because m2 would cause b2.

Davies simply does not think that this picture is very plausible. She (2016, pg. 40) says that this leaves us with “a remarkable universal coincidence, namely that “preventer” desires, which would otherwise intervene in the physical chain of events, are always accompanied by the countermanding double preventers”.

However, on my version of the model, this is a non-issue. I have dealt with the potential breach of Closure between m2 and b2 by proposing that that relation also be understood as double prevention. Therefore, I do not need to worry about the case where Carl has the desire to keep his hand still and therefore N2 firing is not able to cause Carl’s hand to wave. Consider our double prevention case but without m1.

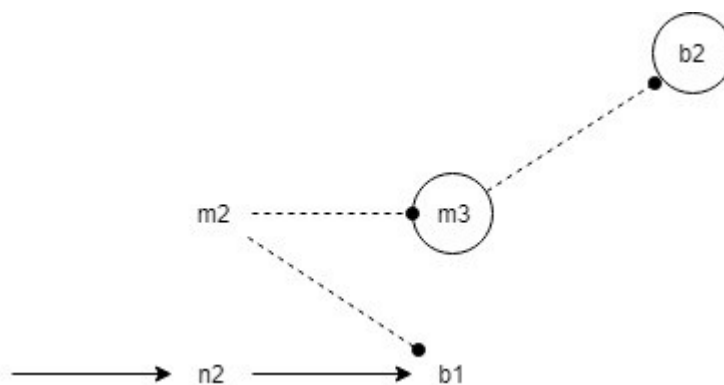


Figure 20

It is still the case that N1 firing causes N2 to fire which would cause Carl to wave his hand. But, Carl has the desire to keep his hand still, and due to this that last causal connection is prevented. Instead, a double prevention relation is made connecting Carl’s desire to keep his hand still to his hand actually staying still<sup>16</sup>. There is no breach of

<sup>16</sup> It’s worth noting that even if this double prevention obtains it is not necessarily the case that Carl’s hand actually does stay still. There would have to be some physically sufficient causes of this occurring as well. We can consider a case where this relationship obtains but is prevented by a third preventer. Carl has the desire to keep his hand still and perhaps under normal circumstances this would be associated with a physical (neural) process that would bring this state of events about. However, it might be the case that Carl’s friend Aimee grabs his hand and waves it for him. In this case the double prevention relation is still there, it is just that it’s last link is itself prevented by Aimee waving Carl’s hand. This seems like a plausible picture and shows that the double prevention relation can obtain no matter if it is successful or not. Because the existence of the double prevention relation is independent of whether it does any actual work or not, we are able to assess it separately of what transpires in the actual world.

Closure in this modified scenario. Because there are no mental events in the transitive causal closure of any physical events.

But, the worries brought up by Gibb and Davies are also dealt with. Gibb worried that the desire to keep the hand still was not able to manifest all it was capable of. That it was not able to stop N2 firing from causing Carl's hand to wave. But on my proposal, it can do just this without any breach of closure. In the modified case with my amendment to the model Carl's desire to keep his hand still prevents N2 firing from causing Carl's hand to wave. I have shown that it can do this without any breach of Closure. Therefore, Gibb's worry that m2 cannot manifest all it is capable of is not an issue.

Furthermore, Davies was concerned that in order to keep our commitment to Closure, it must be the case that there is a "remarkable universal coincidence". That the desires that would do single preventing such as m2 always have the double preventer m1 attached. He thought that this state of affairs was implausible. However, I have shown that this need not be the case. The case we have just examined was one where the single preventer (with respect to the causal connection from n2 to b1) obtained without a corresponding double preventer. In this case there was no breach of Closure. It might be the case that Davies has a reasonable point that to assume that for every preventer there is a double preventer is implausible. However, the implausibility of this has no effect on my version of the model, because it straightforwardly does not assume that this is the case.

You might worry that in the case of m2 obtaining without m1 there is a breach of Closure because m2 has a direct prevention link from m2 to b1. After all, in Gibb's original model there was only ever double prevention actually manifested between mental and physical events, no single prevention. This would be an issue if you thought that prevention was a straightforward cause. However, remember that Gibb's model is predicated on accepting a theory of causation that does not count double prevention as causation. The two main theories that we have looked at that do not count double prevention as causation are the Powers Theory and Production theories. Both of these

do not generally accept causation of absences<sup>17</sup>. Because prevention is a sort of causation of absence we should not accept that prevention is causal under these theories. Therefore, if we are to take Gibb's model on seriously we ought not be concerned about a breach of Closure in the single prevention relation between  $m_2$  and  $n_1$ .

### 3.4.3.2. The Problem of Many $m_2$ 's

Davies is also concerned about another aspect of Gibb's model. She (2016, pg. 40) notes that it is not obvious that there is only one possible desire that could occupy the role of  $m_2$ . In fact, it seems plausible that any number of varying desires could fulfil this role. To look back at our example,  $n_1$  and  $n_2$  will (absent of preventers) cause Carl's hand to wave in a certain way. There is a whole swath of desires that would prevent this and therefore could occupy the role of  $m_2$ . Carl having the desire to keep his hand still is one of them, but so is the desire to move his hand in a different way. Having the desire to keep his whole body still would still prevent his hand from waving, while being a markedly different desire to just keeping his hand still. Davies concern is that in the face of such varying options for  $m_2$  is that we would either need to admit that there would be a correspondingly large number of  $m_1$  like desires, or that a single  $m_1$  would have to be very powerful in that it could prevent all of the possible  $m_2$ 's.

Davies again thinks this does not seem very plausible. However, I think that thinking about our desires as powerful double preventers that can prevent many different  $m_2$  like single preventer desires is exactly the right way to think about them. What is really going on when we have a  $m_1$  like desire? We have a desire that we think does some causal work, and under Gibb's model it does that causal work by preventing at least one contrary desire that would have prevented the physical effect. In order for our desire to do any causal work at all two things must be the case. Firstly, it must be the case that we actually have that desire. This should be obvious. If we did not have a particular desire how could we expect it to do any work. But this is not enough by itself. It seems

---

<sup>17</sup> The notable exception being Fair (1979), but it is unclear that he would accept double prevention as not causal to start with, so we need not worry about this.

completely reasonable to assume that we can have competing desires. I might have the desire to go out to town and spend time with my friends and a competing desire to finish up some work at home. To be a desire like  $m_1$  we need to both actually have that desire, and also for it to be the desire that 'wins out'<sup>18</sup>. The desire has to be the one that rises above potentially several conflicting desires. If these two conditions are met, then we have the sort of desire that we naively want to claim is causal. I might have the conflicting desires to go to town and stay home. The desire to go out to town with my friend might win out in the conflict. Then it seems straightforward that I went out to town because I had the desire to do so (among other reasons).

But if one of several desires wins out in a conflict then it seems fair to say that any conflicting desires that we might have had lost that conflict. That is not to say that we never had that desire. Just that the desire was prevented from being causally relevant. Further, it seems fair to say that in the case of one desire winning out, *all* conflicting desires that you might plausibly have had are prevented from winning. It would be impossible for the desire to stay at home to succeed in its causal work and keep us at home if instead the desire to go to town wins and that is what we do. But this is true of any conflicting desire. If I desire to go to town and do so then it is also the case that my desire to go for a run at the track did not win, and that my desire to play chess did not win. This will be the case for all desires that are in conflict with whichever one wins. Importantly, if a desire meets these conditions then it both prevents any actual conflicting desires that we have from doing causal work, and it would prevent any other plausible conflicting desires from doing so.

For a desire to be an  $m_1$  like desire it must be the case that we have that desire and that it wins out in competition between plausible desires. If it wins that competition, then it necessarily prevents those conflicting desires from exercising any causal power they have. That is not to say that we do not have those desires at all. We might well have conflicting desires. But even if we did, their causal power would be prevented. So, it seems like powerful  $m_1$  like desires that prevent a large number of  $m_2$  like desires is

---

<sup>18</sup> When I talk about competition between desires that competition can be thought about in terms of the strength of the competing desires. In this case the strength refers to the motivational strength of the desires. That is, their ability to actually make me act on them.



exactly right. Desires are not narrow things that focus on preventing a specific contrary desire. They are large positions that have the ability to prevent a wide range of conflicting desires from manifesting their causal power. Davies' suspicions about this being an implausible position are misplaced.

#### 3.4.3.3. No Independent Support for Ever-Present Double Preventers

Davies (2016, pg. 40-41) further worries that we have no independent support that mental events are causally relevant by way of double prevention. When we use mental causes in our ordinary discourse we simply do not refer to double prevention. We simply say that my desire to wave my hand caused me to do so. Without support we simply do not have a good reason to think that the double prevention solution is true.

Gibb responds to this objection by appealing to the phenomenology of William James. She asks us to consider his famous passage describing getting out of bed in the morning:

We know what it is to get out of bed on a freezing morning in a room without a fire, and how the very vital principle within us protests against the ordeal [...] We think how late we shall be, how the duties of the day will suffer; we say, 'I *must* get up, this is ignominious', etc.; but still the warm couch feels too delicious [...] Now how do we *ever* get up under such circumstances? If I may generalize from my own experience, we more often than not get up without any struggle or decision at all. We suddenly find we *have* got up [...] It was our acute consciousness of both the warmth and the cold during the period of struggle, which paralyzed our activity then and kept our idea of rising in the condition of *wish* and not of *will*. The moment these inhibitory ideas ceased, the original idea exerted its effects. (James, 1981, pg. 1133. In Gibb, 2013, pg. 210-211. Ellipses Gibb's)

This sort of scenario fits nicely with Gibb's model. I desire to stay in the warm bed and this desire prevents me from getting up. But this desire is temporarily stopped by my desire to get up. When my desire to stay in the warm bed is prevented, only then do I

get up. The desire to get out of bed, by double prevention, allows the physical process to take over and so I get up.

If we had this sort of scenario in every case where there was supposed mental causation, then this would be good independent evidence in favour of the double prevention model. However, as Davies rightly points out, it is far from obvious that this is always the sort of situation going on. While it does seem like in some case we have one desire that is trumped by some other desire this is not the usual way that we talk about mental causation. Instead we usually claim that our mental events are straightforwardly causes of the physical. There is no appeal to competing desires.

It seems then that James' case, while interesting, is not sufficient evidence that Gibb's model is correct. But that does not mean there is no evidence at all, contrary to Davies' claims. Firstly, the fact that the model provides a compelling response to the causal exclusion argument in line with our intuitions about mental causation is good evidence to suggest that it is correct. The causal exclusion argument is *the* problem of mental causation. Most responses to it rely on rejecting one of its premises and, insofar as all of the premises are intuitive, involve biting the bullet to at least some extent. With Gibb's model we do not have to do this. Because she is questioning the validity of the argument we are able to keep all of our intuitions about the premises and also our intuitions that mental events are causally relevant. The fact that her model is able to respond to the causal exclusion argument in this way where most others cannot do so is a good reason to think that the model is true.

But this might not be a satisfying response. After all this is not an *independent* reason to think the model is how it works in the real world. Davies objection hinges on the fact that the Jamesian explanation of events that Gibb appeals to is not good enough to capture all that we mean when we talk about mental events being causally relevant. I will aim to show that the double prevention model need not rely on such an understanding of events. In fact, mental causation by double prevention could occur in at least two different ways. If the model does not need the Jamesian picture to always be the case, then the sting goes out of Davies objection because the double prevention model is at least consistent with how we view the world.

In my response to the previous objection I argued that in virtue of being an m1 like desire all contradictory desires are prevented from manifesting their causal powers. This prevention could happen in a number of different ways. The most straightforward of these ways is that discussed in James and appealed to by Gibb. There could be two (or more) competing desires and one desire 'wins out' the contest and in doing so prevents an obtaining contradictory desire from manifesting any causal powers. But Gibb misses that this is not the only way to prevent a desire from manifesting its causal powers. Another way that this could happen is by preventing the contradictory desire from obtaining at all.

We can make this clear by looking at a modification of our standard case of double prevention. Carl is flying his bomber mission to bomb the enemy target. The enemy Fighter is dispatched to shoot him down. However, along the way the enemy Fighter encounters Aimee. They have a dogfight which ultimately results in Aimee shooting down the enemy Fighter. Carl, unaware of all of this, successfully completes his mission and bombs the enemy target. This story of double prevention parallels what is going on in James' example that Gibb is appealing to. Both the would-be preventer (the enemy Fighter) and the double preventer (Aimee) are out there having a contest to see which will prevail and be causally relevant. Aimee is a double preventer in virtue of winning this contest.

But this is not the only way that Aimee can be a double preventer. Carl is still a bomber flying his mission to bomb the enemy target. The mission is successful, and Carl bombs his target. At the time of the mission Aimee is happily sitting back at base enjoying her dinner. The day before though she flew a bombing mission of her own. A mission to destroy the enemy's air field. She was successful in her mission too, knocking out all of the enemy's planes. If Aimee had not been successful in her mission, then the enemy Fighter would have been able to use one of these planes to intercept Carl and prevent his mission. Aimee prevented this from happening because she bombed all of the planes.

In this second scenario Aimee is no less a double preventer. Without her mission being successful, Carl's would have also failed. Clearly her mission is causally relevant. But in this case, she is a double preventer in a very different way. In the first example she is a

double preventer by virtue of winning the contest. In this case she is a double preventer in virtue of preventing the contest from even happening. Neither of these options is any less a relevant double preventer than the other.

Under this sort of picture of double prevention, the double prevention model does not need to insist on a Jamesian model of how it would work in the world. Perhaps there are some cases where that is what is happening. That is perfectly fine. Davies was correct in pointing out that this is unlikely to always be the case. However, with a new understanding of double prevention the model can be used in other types of scenarios. My proposal here does not provide the independent support for double prevention occurring in all cases of mental causation that Davies was asking for. However, his main reason for asking for that evidence was that he thought that Gibb's picture of competing desires was implausible to *always* be the case. I agree. However, Gibb's picture is not the only way that we can have double preventers. I have shown that mental events can also act as double preventers by preventing potentially competing desires from being in a position to compete. Taking double preventers to act in either of these sorts of ways should be more consistent with the empirical picture than Gibb's proposal.

#### 3.4.3.4. Objection to 'Denial of Overdetermination'

Davies final objection concerns Exclusion. Consider the case where my desire does not conflict with the event being caused by the physical causal chain. My physiological/neural state (feeling hot) leads me to wave my hand in a certain way (vigorously to fan my face), as does my desire to cool myself down. This is a problem for Gibb's version of the account. There is no desire preventing the physiological chain. On Gibb's view the way in which desires can be double preventers is to contest an obtaining contradictory desire and win. But in this case, such a contradictory desire does not obtain. So, my desire to wave my hand cannot be acting as a double preventer. But double prevention is the way in which mental events can be causally relevant and so it must be the case that in these sorts of cases my desires are not causally relevant.

Gibb endorses this sort of picture, where a desire will be causally relevant or not depending on whether there exists a suitable preventer for it to double prevent. If there

is a suitable preventer, then the desire is a causally relevant one. But the exact same desire could be causally irrelevant if the preventer did not exist. Gibb (2013, pg. 212) notes that it is precisely for this reason that our actions feel 'resistless' (as James puts it) when there is no desire preventing them.

But Davies (2016, pg. 41) claims this surely cannot be right. Although I can choose to act on my desires or not, a desire cannot choose to be causal. Instead he offers that what is really going on is that my eventual action in these cases is overdetermined by my desires and the physiological situation. In addition, this is a systematic overdetermination because whenever there is a case of this sort it will be one of overdetermination. If Davies picture is right, the double prevention model will be in breach of the 'Denial of Overdetermination' premise.

My version of the account is not subject to this same objection. As I argued above there are different ways that a mental event can be a double preventer. Not just by overcoming a contradictory obtaining desire. So, in the case that I have a desire and no competing one, my desire can still be causally relevant through double prevention. It just has to be a different sort of double prevention. This means that there is no overdetermination because my mental event is not a cause, it is merely causally relevant. This undermines the thrust of Davies objection.

### 3.5. Conclusion

In this section I have presented Gibb's model of a counterexample to the causal exclusion argument. The model works by proposing that mental events are double preventers of physical ones. In this way, mental events are causally relevant to those physical events without being actual causes. I have shown that Gibb's model does represent a counterexample to the validity of the argument. However, the model rests on several controversial claims. Most notably the existence of some  $n_0$  which is a sufficient cause of both the mental event in question and a corresponding physical causal chain.

Gibb endorsed the model as a dualist, but I have argued that it can be used by both dualists and non-reductive physicalists. I also argued that one need not endorse a Powers View of causation to make use of the model. Instead any view which treats double prevention as not fully causal will suffice. Both of these arguments should widen the appeal of the model.

I proposed two major adaptations to Gibb's version of the model which will also broaden its appeal. Firstly, I showed that by adapting a problematic mental to physical causal link to one of double prevention we can end the reliance of the model on some  $n_0$ . I believe that this adaptation is in line with the spirit of Gibb's initial proposal. By ending the reliance on  $n_0$  I have made the model more appealing to those who found its existence problematic, including those who endorsed a libertarianism on which the agent initiates causal chains.

Secondly, I argued that Gibb's conception of the phenomenological picture on which we always have extant competing desires was not the only way in which the mental can act as a double preventer. The mental can also act as a double preventer by preventing the potentially competing desire from manifesting at all. This means that there is a much broader and more plausible picture on which mental double preventers could be operating.

Accepting this model of mental causation does require some assumptions. Notably, it only works if we do not count double preventers as fully causal. However, I believe that with my adaptations it is a broadly compelling picture of how mental causation operates.

#### 4. Overall Conclusion

In my thesis I have explored the thorny issue of double prevention and one potential place that it can be of use in our metaphysics. I clearly and concisely presented a view of what double prevention is and provided an example. I then explained why double prevention is in tension with several widely held theses about causation. Double prevention seems to be clearly causally relevant at least, but there is some contention as to whether it should count as fully causal or not. There have been several attempts to resolve this contention, notably by Mumford and Anjum (2009), Schaffer (2000), and Hall (2004). I detailed these attempts and reached some tentative conclusions about double prevention. None of the current attempts to resolve double prevention are fully satisfactory. They all are still in conflict with widely held theses about causation, which was the problem with double prevention in the first place. Nevertheless, Mumford and Anjum have provided us a general strategy that could be used to resolve the tensions created by double prevention. If we could double prevention as causally relevant but not fully causal, then we create an important place for double prevention in our causal analysis while dodging many of its issues.

If we do count double prevention as not fully causal, then it has the ability to do some work in our metaphysics. Sophie Gibb (2013) proposes that the mental could act as a double preventer on the physical, providing a counterexample to the causal exclusion argument. I believe that Gibb's proposal shows great promise. However, it relies on controversial assumptions, such as accepting the Powers Theory of causation and the existence of  $n0$ . I propose that with some adaptations, Gibb's reply can become a much more appealing response to the causal exclusion argument. I defend my proposal from objections levelled at Gibb's original reply and some others.

It is clear that there is still much work that could be done on double prevention. It remains underdeveloped in the literature. I hope that my thesis can be of use in clearing up the issue and providing a starting point for moving forward with research.

## 5. Bibliography

- Armstrong, D. (1999). The Open Door: Counterfactual versus Singularist Theories of Causation. In H. Sankey, *Causation and Laws of Nature* (pp. 175-186). Dordrecht: Kluwer.
- Aronson, J. (1971). On the Grammar of Cause. *Synthese*, 22(3/4), 414-430.
- Baysan, U. (2015). Realization Relations in Metaphysics. *Minds and Machines*, 25(3), 247-260.
- Bennett, K. (2003, September). Why the Exclusion Problem Seems Intractable, and How, Just Maybe, to Tract It. *Nous*, 37(3), 471-497.
- Broadbent, A. (2012, April). Causes of Causes. *Philosophical Studies*, 158(3), 457-476.
- Davidson, D. (1967). Causal Relations. *The Journal of Philosophy*, 64(21), 691-703.
- Davies, K. (2016). Powers, Double Prevention and Mental Causation. *Metaphysica*, 17(1), 37-42.
- Dennett, D. (2003). *Freedom Evolves*. New York: Penguin Books.
- Dennett, D. (2003). The Self as a Responding—and Responsible—Artifact. *Annals of the New York Academy of Sciences*, 1001(1), 39-50. doi:doi:10.1196/annals.1279.003
- Fair, D. (1975, November). Causation and the Flow of Energy. *Erkenntnis*, 14(3), 219-250.
- Gibb, S. (2009). Explanatory Exclusion and Causal Exclusion. *Erkenntnis*, 71(2), 205-221.
- Gibb, S. (2010). Closure Principles and the Laws of conservation of Energy and Momentum. *Dialectica*, 64(3), 363-384.
- Gibb, S. (2012). Nonreductive Physicalism and the Problem of Strong Closure. *American Philosophical Quarterly*, 49(1), 29-41.
- Gibb, S. (2013). Mental Causation and Double Prevention. In S. C. Gibb, E. J. Lowe, & R. D. Ingthorsson, *Mental Causation and Ontology* (pp. 193-213). Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199603770.001.0001
- Gibb, S. (2014). The Entailment Problem and the Subset Account of Property Realization. *Australasian Journal of Philosophy*, 92(3), 551-566.
- Gibb, S. (2015). Physical Determinability. *Humana Mente*, 8(29), 69-89.
- Gibb, S. (2015). The Causal Closure Principle. *The Philosophical Quarterly*, 65(261), 626-647.
- Gibb, S. (2015). VIII-Defending Dualism. *Proceedings of the Aristotelian Society*, 115(2pt2), 131-146.
- Grasso, M. (2015). The Metaphysics of Free Will: A Critique of Free Won't as Double Prevention. *Rvista Internazionale Di Filosofia E Psicologica*, 6(1), 120-129.



- Grygianiec, M. (2016). How to Get Rid of Closure. *Diametros*, 48(1), 1-17.
- Haggard, P., & Magno, E. (1999). Localising Awareness of Action with Transcranial Magnetic Stimulation. *Experimental Brain Research*, 127, 102-107.
- Hall, N. (2004). Two Concepts of Causation. In J. Collins, N. Hall, & L. A. Paul (Eds.), *Causation and Counterfactuals* (pp. 225-276). Cambridge, Massachusetts, United States of America: MIT Press.
- Hitchcock, C. (2003). Of Humean Bondage. *British Journal for the Philosophy of Science*, 54(1), 1-25.
- James, W. (1981). *The Principles of Psychology*. (F. H. Burkhardt, Ed.) Cambridge, MA: Harvard University Press.
- Kim, J. (1989). Mechanism, Purpose, and Explanatory Exclusion. *Philosophical Perspectives*, 3, 77-108.
- Kim, J. (1993). *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press.
- Kim, J. (1996). *Philosophy of Mind*. Boulder, CO: Westview Press.
- Kim, J. (2005). *Physicalism, or Something Near Enough*. Princeton: Princeton University Press.
- Kim, J. (2010). Thoughts on Sydney Shoemaker's Physical Realization. *Philosophical Studies*, 148(1), 101-112.
- Lewis, D. (1986). *Philosophical Papers Vol. II*. Oxford: Oxford University Press.
- Lewis, D. (2004). Void and Object. In *Causation and Counterfactuals* (pp. 277-290). Cambridge, MA, USA: MIT Press.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8, 529-539.  
doi:10.1017/S0140525X00044903
- List, C., & Stoljar, D. (2017). Does the Exclusion Argument Put Any Pressure on Dualism. *Australasian Journal of Philosophy*, 95(1), 96-108.
- Lowe, E. J. (2000). Causal Closure Principles and Emergentism. *Philosophy*, 75(294), 571-586.
- Mackie, J. L. (1974). *The Cement of the Universe: A Study in Causation*. Oxford: Clarendon Press.
- Malcolm, N. (1968, January). The Conceivability of Mechanism. *Philosophical Review*, 77(1), 45-72.
- Maslen, C. (2018). The Case for Widespread Simultaneous Causation. *The Philosophical Quarterly*, 68(270), 123-137.

- Maslen, C., Horgan, T., & Daly, H. (2009). Mental Causation. In H. Beebe, C. Hitchcock, & P. Menzies (Eds.), *The Oxford Handbook of Causation* (pp. 523-553). Oxford: Oxford University Press.
- Mele, A. (2014). *Free: Why Science Hasn't Disproved Free Will*. Oxford University Press USA.
- Menzies, P. (1996, January). Probabilistic Causation and the Pre-emption Problem. *Mind*, 105(417), 85-117.
- Montero, B. (2003). Varieties of Causal Closure. In S. Walter, & H.-D. Heckmann (Eds.), *Physicalism and Mental Causation* (pp. 173-187). Exeter: Imprint Academic.
- Mumford, S., & Anjum, R. L. (2009). Double Prevention and Powers. *Journal of Critical Realism*, 8(3), 277-293.
- Mumford, S., & Anjum, R. L. (2011). *Getting Causes From Powers*. Oxford: Oxford University Press.
- Nesic, J. (2013). Dispositional explanations in dualism. *Filozofija I Društvo*, 24(4), 218-241.
- Paul, L. A., & Hall, N. (2013). *Causation: A User's Guide*. Oxford: Oxford University Press.
- Russo, A. (2016). Kim's Dilemma: Why Mental Causation is not Productive. *Synthese*, 193(7), 2185-2203.
- Schaffer, J. (2000, June). Causation by Disconnection. *Philosophy of Science*, 67(2), 285-300.
- Shoemaker, S. (2001). Realization and Mental Causation. In C. Gillett, & B. Loewer (Eds.), *Physicalism and its Discontents* (pp. 74-98). Cambridge: Cambridge University Press.
- Shoemaker, S. (2007). *Physical Realization*. Oxford: Oxford University Press.
- Shoemaker, S. (2013). Physical Realization Without Preemption. In S. C. Gibb, E. J. Lowe, & R. D. Ingthorsson (Eds.), *Mental Causation and Ontology* (pp. 35-57). Oxford: Oxford University Press.
- Wilson, J. (1999). How Superduper does a Physicalist Supervenience need to be? *Philosophical Quarterly*, 49(194), 33-52.
- Wilson, J. (2011). Non-reductive Realization and the Powers-based sub-set strategy. *The Monist*, 94(1), 121-154.