

Keywords: causation, moral responsibility, moral luck, counterfactuals, omissions

Abstract: Both causation and moral responsibility seem to come in degrees, but explaining the metaphysical relationship between them is more complex than theorists have realized. This paper poses an original puzzle about this relationship and uses it to reach three important conclusions. First, certain natural resolutions of the puzzle reveal the existence of a new sort of moral luck called *proportionality luck*. Second, there is indeterminacy in the type of causal relation deployed in assessments of moral responsibility. Finally—and most importantly—leading theories of causation do not have the ability to capture the sorts of causal differences that matter for moral evaluation of agents' causal contributions to outcomes.

## **Causal Proportions and Moral Responsibility** (forthcoming in *Oxford Studies in Agency and Responsibility*)

Sara Bernstein

On the face of it, you are only morally responsible for what you cause. One thing that matters for the degree to which you are morally responsible for an outcome is your precise causal contribution to the outcome—intuitively, “how much” you contribute to the outcome’s occurrence. Having an adequate theory of such *degrees* of causal contribution is essential for unpacking the role of causation in attributions of moral responsibility, and in understanding the metaphysical relationship between causation and moral responsibility more generally. This paper poses a puzzle about the relationship between degrees of causation and degrees of moral responsibility, and uses it to argue that the metaphysical relationship between them is much more complex than theorists have realized.

The puzzle is brought out in the following pair of cases:

(Victim) Two independently employed assassins, unaware of each other, are dispatched to eliminate Victim. Being struck by one bullet is sufficient to kill Victim. Each assassin shoots, and Victim dies.

(Hardy Victim) Two independently employed assassins, each unaware of the other, are dispatched to eliminate Victim. Unbeknownst to both assassins, Victim is particularly hardy, and requires two bullets for his demise. Each assassin shoots, and Victim dies.

The cases differ causally insofar as Victim is a case of *causal overdetermination*, in

which there are multiple sufficient causes of an outcome, whereas Hardy Victim is a case of *joint causation*, in which there are multiple necessary causes of an outcome. But do the cases differ morally? That is: does each assassin's proportion of moral responsibility for Victim's death quantitatively differ between Victim and Hardy Victim? Call this *The Moral Difference Puzzle*.<sup>1</sup>

In section 1, I use the Moral Difference Puzzle to unearth and develop a new kind of moral luck I call *proportionality luck*. Proportionality luck involves an agent's proportion of moral responsibility for an outcome being out of that agent's control. In section 2, I argue that the central pair of cases in the Moral Difference Puzzle reveals indeterminacy in the causal relations deployed in assessments of moral responsibility. In section 3, I argue that no leading theory of causation can capture the principle that underlies the sorts of differences in causal contribution that matter for moral responsibility.

First, a few preliminaries. There are many aspects of agents' actions that are relevant for moral responsibility. Here I will be specifically interested in the relationship between causation and moral responsibility, so I will not be primarily concerned with features of the examples that would normally be considered highly morally relevant to such cases: for example, intentions of the agents, duties, and expectations. I will assume that we can make philosophical progress on the relationship between causation and moral responsibility even when setting these things aside. Thus, I will treat the agents in the central pair of cases as on par with respect to intentions, duties, expectations, and so on.

Second, I will not offer a general theory of what moral responsibility is. Though I take my results to be applicable to many theories of moral responsibility, I will assume here that moral responsibility at least partially tracks praiseworthiness and blameworthiness, that these can come in degrees, and that adequate sense can be made of degrees of praise and blame.

## 1.1 The Logical Space of Answers

Return to the Moral Difference Puzzle. Its possible resolutions are:

---

<sup>1</sup> Carolina Sartorio (forthcoming) discusses a similar puzzle that she had been working on independently.

(1) *No moral difference*. There is no moral difference between (Victim) and (Hardy Victim). Each assassin in (Victim) bears equal moral responsibility to the individual assassins in (Hardy Victim).

(2) *Moral difference: greater*. There is a moral difference between (Victim) and (Hardy Victim). Each assassin in (Victim) bears greater moral responsibility than each assassin in (Hardy Victim).

(3) *Moral difference: lesser*. There is a moral difference between (Victim) and (Hardy Victim). Each assassin in (Victim) bears less moral responsibility than each assassin in (Hardy Victim).

According to answer (1), there is no moral difference between the assassins in (Victim) and (Hardy Victim). One motivation for this view is that the intentions and actions of the assassins in both (Victim) and (Hardy Victim) are identical. If we hold that the way things turn out is irrelevant to the blameworthiness of agents<sup>2</sup>, then the cases of joint causation and overdetermination are identical with respect to amounts of moral blameworthiness.

But this approach is problematic insofar as it divorces moral evaluation of agents' causal contribution to outcomes from the way that things actually turn out. Retaining a relationship between actual causation and moral responsibility requires that quantitative differences in actual causal contribution lead to quantitative differences in moral responsibility between agents.

It is most plausible to hold that the greater or lesser an agent's actual causal contribution to an outcome, the greater or lesser her moral responsibility for that outcome (*ceteris paribus*). More formally, moral evaluation of an agent's contribution generally adheres to the following principle:

**Proportionality**. An agent's moral responsibility for an outcome is proportionate to her actual causal contribution to the outcome.

This principle is a precisification of the intuitive doctrine with which we started: if one is morally responsible only for the outcome that one causes, one should also only be

---

<sup>2</sup> This position is termed *rationalism* by Wolf (2001).

morally responsible for the *part* or *proportion* of an outcome that one causes.

**Proportionality** extends an everyday intuition about moral responsibility for whole outcomes to moral responsibility for parts of outcomes. Accepting this plausible principle requires that we reject answer (1).

## 1.2 Proportionality Luck

On to the latter two answers to the puzzle. According to both answers, there is a moral difference between the agents in (Victim) and (Hardy Victim). Committing to a moral difference between these cases in either direction is a commitment to the existence of moral luck. Specifically, it is a commitment to the existence of resultant moral luck. *Resultant luck* is the sort of moral luck that arises when the ways that our acts turn out, or the consequences of our acts, are beyond our control. Resultant luck is often demonstrated by appeal to example. The paradigmatic pair of contrast cases that illustrate resultant moral luck involve a lucky driver and an unlucky driver. Both are negligent drivers, but a pedestrian appears in the unlucky driver's path whereas one does not appear in the lucky driver's path. The unlucky driver kills the pedestrian; the lucky driver returns home without incident. Only the unlucky driver is morally responsible for a pedestrian's death.

**Proportionality** is an extension of the reasoning that undergirds these more general cases of moral luck. Just as the unlucky driver is morally responsible for what she actually causes—a pedestrian's death--and the lucky driver is not morally responsible in virtue of the death's not having occurred, an agent is morally responsible only for the proportion of an outcome that she actually causes.<sup>3</sup> The unlucky driver is unlucky precisely because the consequences of her negligent driving are not under her control. Similarly, Victim's hardiness is a matter out of the assassins' control. Each assassin in (Victim) and (Hardy Victim) sets out to eliminate the victim himself: the assassins' intentions are the same in both cases. Since Victim's hardiness lies out of the assassins' control, it is a matter of luck that each assassin's portion of causal responsibility in

---

<sup>3</sup> The "actual" in "actual causal contribution" is meant to stand in contrast to an agent's merely counterfactual causal contribution.

(Victim) differs from each assassin's portion of causal responsibility in (Hardy Victim). Assuming that **Proportionality** is true, this quantitative difference in causal responsibility generates a quantitative difference in moral responsibility owing to resultant luck.

Call this sort of moral luck *proportionality luck*. Proportionality luck involves circumstances out of an agent's control either increasing or decreasing that agent's proportion of moral responsibility for an outcome. Proportionality luck can be thought of as a type of resultant luck, since it involves the way that an outcome turns out. It is a natural outgrowth of the idea that an agent's moral responsibility is proportionate to her actual causal contribution to an outcome.

Proportionality luck is everywhere. Almost every one of us in the first world produces a portion of carbon emissions sufficient to do some portion of environmental damage; it is a matter of luck which of us do more damage than others. Many people use antibiotics even when they are not necessary; it is a matter of luck which users contribute to antibiotic resistance more than others.

In addition to moral blameworthiness, moral praiseworthiness is also subject to proportionality luck. Charities claim that "every little bit helps". Consider a case in which ten individuals donate to charity, and it only takes five of their contributions to alleviate hunger in the stricken area. (The excess funds go to smaller, but still important, tasks, such as providing clean places to bathe.) Here, which agents are responsible for the actual alleviation of hunger is a matter of causal luck—i.e., which donations actually end up helping the victims.

Note that cases of preemption, as when Assassin 1 shoots and kills victim first, thus preempting Assassin 2's bullet from causing Victim's death, are cases of traditional resultant moral luck. For presumably Assassin 2's bullet makes no actual causal contribution to Victim's death. In contrast to these traditional cases, what is distinctive about proportionality luck is that there are multiple actual causal contributions to the outcome.

Proportionality luck is unique along several dimensions. First and most obviously, there is a different basis for moral differentiation between lucky and unlucky agents. Whereas traditional cases of moral luck involve moral differentiation between agents

based on the occurrence of an outcome (e.g. the pedestrian's death) versus the non-occurrence of that outcome, proportionality luck involves a fixed outcome with quantitatively different causal contributions. The question is not *whether* the relevant outcome occurs, but rather, what comparative causal proportions contribute to the outcome's occurrence.<sup>4</sup>

Proportionality luck is more intricate than other types of resultant moral luck. This complexity is made apparent when one attempts to apply Sartorio's (2013) distinction between two senses of "how things turn out" in cases of resultant moral luck. One sense is *outcome-driven*, tracking whether and how an outcome occurs after the agent acts. Another sense is *agent-driven*, tracking how the agent ends up being related to an outcome.

Consider the distinction at work in typical case of resultant luck involving drunk drivers. Outcome-driven moral luck captures the occurrence of the bad outcome (the pedestrian's death) apart from how the agent is related to it. Agent-driven moral luck captures the fact that the unlucky agent ended up being causally responsible for the outcome. What is out of an agent's control in outcome-driven moral luck is whether a particular outcome occurs. What is out of an agent's control in agent-driven moral luck is whether the agent's behavior ends up resulting in (that is, being causally related to) the particular outcome—in this case, the pedestrian's death. For example, suppose that the pedestrian, hit by the truck of a negligent driver, dies of a totally unrelated health condition a short while later. Here, though the driver is unlucky that he hit a pedestrian, his negligence does not actually result in Victim's death.

As a genus, proportionality luck straddles the distinction between outcome-driven and agent-driven resultant moral luck. Some species are outcome-driven and other species are agent driven. In (Hardy Victim), whether the outcome occurs is out of an agent's control: had the other assassin not been present, Victim would not have died. The very *occurrence* of the outcome is out of each assassin's control. But in analyzing the moral difference between (Victim) and (Hardy Victim), what is out of each assassin's

---

<sup>4</sup> Elizabeth Harman astutely pointed out to me that cases of "traditional" moral luck might also be thought of as cases of proportionality luck in the following sense: one agent is responsible for the entire outcome and another agent is responsible for none of the outcome. Here I will restrict my focus to cases in which both agents causally contribute to the outcome in some way, and are thus each responsible for some of the outcome.

control is her quantitative causal contribution to the outcome. For it is a matter of luck whether she is partly or fully causally responsible for the outcome.

Proportionality luck is unique because it hinges on the idea of differential causal contribution, i.e., multiple agents being differentially causally responsible for the same outcome. Proportionality luck is interesting partly because its existence follows naturally from a commitment to **Proportionality**. Aside from being a basis of much of our thinking about moral luck, this principle also infuses other parts of moral theory-- for example, formulations of consequentialism. It is therefore important to have a careful account of *what it is* for an agent to be more or less causally responsible for an outcome, as in the case of Victim and Hardy Victim. But I will now suggest that such an account faces several obstacles: first, an intractable conflict between two types of causal intuitions; and second, a resistance to being captured by leading theories of causation.

## 2.0 A Causal Ambiguity in Differences of Contribution

The Moral Difference Puzzle asks: is there a moral difference between the assassins in both cases? Assuming the answer is *yes*, there is a prior question: are the assassins in (Victim) *more* or *less* causally responsible than those in (Hardy Victim)? I have stayed silent on this topic until now.

*Prima facie*, it is plausible to say that each assassin in (Victim) is *more* causally responsible than each assassin in (Hardy Victim). The idea is that each assassin is “fully” responsible for an outcome in a case of overdetermination, but only “partly” responsible in a case of joint causation. Joint causation is often thought of as a kind of “causal teamwork”, whereas overdetermination is often thought of as “double the causation”. These sorts of causal intuitions implicitly employ a “productive” or *oomph* theory of causation.<sup>5</sup> According to that sort of theory, causation is a matter of energy transference between a cause and an effect. Because it takes *two* assassins to transfer the necessary energy to Victim in the joint causation case, each agent is partially responsible for the death. And because either transfer of energy would have been sufficient to kill Victim in the overdetermination case, both assassins are, in a sense, “fully” responsible for

---

<sup>5</sup> Here I rely on Ned Hall’s (2004) distinction between dependent and productive concepts of causation.

Victim's death. This sort of thinking motivates answer 2: *Moral Difference: Greater*.

But there is another causal reading of the case. Rather than invoke "productive" causal intuitions, consider "dependent" or counterfactual causal intuitions about our cases. In a case of joint causation, each cause is necessary to bring about an effect. If one cause fails to occur, then the effect fails to occur. In a way, then, each assassin is more essential to the victim's death in (Hardy Victim) than in (Victim), for the victim's death would have occurred no matter what the second assassin did in (Victim), but not if the second assassin had opted out in (Hardy Victim). The difference is powerfully brought out in the following omissive variant on the case:

(Conscientious Objector) Two assassins on the same team have been dispatched to eliminate Victim. Each knows that Victim is particularly hardy and requires both bullets for his demise. Assassin 1 shoots. Assassin 2 is gripped by a crisis of conscience, and does not shoot. Victim survives.

In this case, had Assassin 2 fired his bullet, Hardy Victim would have died. But the failure of Assassin 2 to fire causes Victim's survival. In joint causation cases, counterfactual dependencies hold between each conjunct of a conjunction of necessary causes, whereas in overdetermination cases, counterfactual dependence does not obtain between either conjunct and the effect. In joint causation cases, a single conjunctive cause is just as important to the occurrence of an effect as a necessary cause in a case with a single sufficient cause. Consider the following additional example:

(Soldiers) Five soldiers in a military installation are each positioned at a button. All five button-pushes are necessary to bring about the launching of a weapon, and the soldiers are all aware of the launch requirements. A call comes down from on high to launch the weapon at a group of innocent children. Four soldiers push their buttons, and one soldier abstains from pressing his button. The weapon does not launch.

Here, the conscientious soldier's failure to press the button has the same outcome as *no* soldiers pressing their buttons. Since the weapon-launching depends on all of them jointly, a single failure to press the button prevents the occurrence of the launch. Intuitively, the abstaining soldier is the "hero" of the case: his failing to press the button prevents the death of the innocent children. Similarly, the assassin with a crisis of

conscience single-handedly prevents the death of Victim by failing to shoot.

These sorts of examples motivate the idea that each assassin in (Hardy Victim) is *more*, rather than *less*, causally responsible than each assassin in (Victim). The counterfactual profile of the cases illustrates a sense in which an outcome depends more on a joint cause than on an overdetermining cause, since the joint cause is required for the outcome to occur. On this way of thinking, overdetermining agents are less causally responsible than jointly causing agents, since either overdeterminer alone was sufficient for the outcome. If either overdetermining assassin had abstained, it would not have made a difference to Victim's death. Yet if either jointly causing assassin had abstained, the death would not have occurred.<sup>6</sup> This is a motivation for the truth of Answer 3: *Moral Difference: Lesser*.

Whether the assassins in (Victim) bear more moral responsibility than those in (Hardy Victim) depends on which causal concept one employs.<sup>7</sup> Here I do not come down on either side. Rather, I take the lesson to be that these contrast cases reveal a kind of semantic indeterminacy in the causal relation underlying **Proportionality**—one that can be resolved in two motivated, but incompatible, ways. A major reason why there is “more than meets the eye” to the idea that moral responsibility is proportionate to causal responsibility is that there are multiple causal relations that can be employed in moral assessments of differential causal contributions, and there are no clear, principled rules for which type of causal relation should be used. As I shall argue in the next section, the relationship between causation and moral responsibility faces an even deeper problem than this indeterminacy: no leading theory of causation can fully account for the relevant moral intuitions.

### 3. Proportionality

---

<sup>6</sup> One might still resist the intuition that agents in a case of overdetermination are less responsible than agents in a case of joint causation. But consider an example involving, say, 100 assassins, each poised at individual buttons. Any individual button-push will cause a fatal missile strike to Victim. All 100 assassins push their buttons, and Victim dies. Here, any single button-push is sufficient to kill Victim, and Victim's death is overdetermined 99 times over. Many have the intuition that the sheer number of assassins involved lessens the moral responsibility of each one.

<sup>7</sup> In my “Causal and Moral Indeterminacy” (forthcoming), I argue that this is a type of indeterminacy that afflicts moral responsibility as well.

Above, I outlined a new type of resultant moral luck, *proportionality luck*, and suggested that it is a natural outgrowth of **Proportionality**, or the idea that an agent's moral responsibility for an outcome is proportionate to her causal contribution to it. I also suggested that attempting to apply **Proportionality** to the Moral Difference Puzzle reveals a kind of indeterminacy with respect to whether a productive or a dependent causal theory should be used: the sort of causal relation being employed determines whether we find an agent to be more or less morally responsible for a particular outcome.

As it stands, **Proportionality** is fairly imprecise. Though we have an intuitive idea of what it is for an agent to be more or less a cause of a particular outcome, and though we have an intuitive idea of what it is for an agent to bear greater or lesser moral responsibility for an outcome, it would be helpful to have a more careful metaphysical account of these ideas. Doing so requires, first and foremost, a formulation of what it is for one agent to be “more of a cause” of an outcome than another agent, or for anything to be “more of a cause” of an outcome than anything else. Call this idea “causal differentiation.”<sup>8</sup> I will now suggest that accounting for causal differentiation is a serious obstacle to a more careful formulation of **Proportionality**.

### 3.1 Causal Quantity

Causal differentiation is best illustrated by example. Suppose that I hoist a heavy Oxford English Dictionary onto a shelf. Now, how much did I cause the book to be lifted onto the shelf? A natural answer is: I caused *all* of the outcome.<sup>9</sup>

In contrast, if a friend helps me lift the book, it is natural to say that each of us partially contributes to the placement of the book on the shelf. In that case, each of us caused the book to be lifted in virtue of causing part of the book to be lifted. In outcomes with multiple causes, locutions such as “x caused y” mean “x caused y in virtue of causing a *part* of y”. We often take something to be a cause in virtue of its making *a* causal contribution— but not the entire contribution— to an outcome.

Such cases make it plausible to think of causation as having quantity-like

---

<sup>8</sup> Michael S. Moore (2009: pp. 118-119) terms a similar concept “the scalar nature of legal causation”. Harms with multiple causes are “divisible harms”.

<sup>9</sup> Here I assume that we can distinguish background conditions from causes.

properties. *Quantity-like properties* are features that everyday quantities, like a portion of milk, exhibit. I can add to it (by adding more milk), subtract from it (by pouring some out), or divide it (by pouring some milk into another glass). Similarly, causal contribution can be divided (as when two friends lift a book together), added (as in the case of adding more friends to the lifting of the book), and subtracted (as in a case where a causal helper is removed.)

Not all causal contributions are equal. If five construction workers of varying strength and physical aptitude lift a heavy wooden plank, it seems like the weaker construction workers causally contribute less to the lifting than the stronger construction workers, who shoulder more of the weight.

Clearly, there is an intuitive sense in which something can be more or less of a cause than something else. But, as I shall now argue, giving a *metaphysical* account of the causal intuition behind **Proportionality** faces steep obstacles. There are several different problems. Some theories of causation cannot account for causal differentiation in general. The theories simply do not have the resources to model the idea that there can be comparative degrees of causation, or quantitative differences in causal contribution. Other theories of causation can account for causal differentiation, but not in a way that does justice to our moral intuitions about differential causal contribution. I will now run through several leading theories of causation and show how each falls short of being able to fully account for **Proportionality**. The discussion is not meant to be exhaustive, but rather to provide an idea of the obstacles to developing a careful account of the proportionate relationship between causation and moral responsibility.

### 3.2 Productive Theories of Causation

Productive theories of causation hold that causation is a matter of “production” or “mark transfer” between  $c$  and  $e$ . There are many varieties,<sup>10</sup> but here I will focus on Dowe’s (2000) theory of physical causation, according to which  $c$  is a cause of  $e$  if there is a transfer of conserved quantity (such as force or energy) from  $c$  to  $e$ . For example, one domino causes another to fall in virtue of the transfer of energy from one to another.

---

<sup>10</sup> Salmon (1994) is another prominent proponent of so-called “physical causation”.

This view does have resources to account for a sort of causal differentiation. We might say that  $c_1$  is more of a cause than  $c_2$  if  $c_1$  transfers greater conserved quantity than  $c_2$ . For example, if assassin 1's bullet hits Victim with greater impact than that of assassin 2, assassin 1 is a greater cause of Victim's death.

But, as one might guess, this view does not generally track the kind of causal differentiation relevant to moral responsibility. Consider the following case:

(Bad Politician) Bad Scientist hoists a weapon onto a launch pad for mere testing. But Bad Politician gains access to the weapon's computerized control system, and presses the launch button, launching the weapon.

Here, Bad Scientist is responsible for transferring far more conserved quantity to the weapon than Bad Politician (who merely presses a button), and yet Bad Politician is intuitively more responsible for the weapon's launch.<sup>11</sup>

Here is another example:

(Enthusiastic Soldier) It takes two button-pushes to launch a weapon. The call comes down from above, and one soldier presses his button. The other soldier does a dance on top of her button in order to push it.

In this example, the enthusiastic soldier isn't intuitively more of a cause of the weapon's launch than the unenthusiastic soldier. The difference in conserved quantity is largely irrelevant to the intuitive notion of more or less causal contribution as it relates to our concept of differential moral responsibility.

Certainly, we can imagine counterexamples. Suppose that multiple evil scientists are poised on highly sensitive buttons. The more pressure exerted on the button, the stronger and more painful the shock delivered to Victim. Here, the amount of force transferred to each button does seem relevant to moral evaluation. But this is a mere accident of the case.

The lesson is that though the conserved quantity theory has the resources to

---

<sup>11</sup> One might hold that the Bad Politician case leans too heavily on normative intuitions, especially given that the case includes malevolent intentions. But other cases absent such intentions can easily be constructed, with the same results: amount of energy transfer is often irrelevant to moral responsibility.

account for something's being more of a cause than something else, this notion of causation isn't usually the salient causal concept in moral evaluation.

### 3.3 Counterfactual Theories of Causation

On the simple counterfactual view, causation is the ancestral of counterfactual dependence, which roughly holds that  $c$  is a cause of  $e$  if had  $c$  not occurred,  $e$  would not have occurred. For example: if Billy hadn't thrown his rock at the window, the window would not have shattered. Lewis (1973b, 1979) is the most famous proponent of this theory.

The law largely utilizes the simple counterfactual account to analyze legal responsibility for outcomes.<sup>12</sup> According to the law, an agent is generally responsible for an outcome if it would not have occurred "but for" the agent's actions. The simple counterfactual account of causation is generally considered the gold standard for judging moral responsibility for outcomes.

Aside from well-known problems that involve preemption and overdetermination, the simple counterfactual account does not have the resources to account for causal differentiation. Either a counterfactual dependence holds or it doesn't; either something is a cause or it's not. Context can play a role in determining which causes are *salient*, but on the simple counterfactual view, it does not play a role in determining if one is more of a cause than the other. In (Enthusiastic Soldier), for example, the launching of the weapon is counterfactually dependent on both button-pushings. But neither comes out as more of a cause than the other. The launching is counterfactually dependent on both.

In contrast to the simple counterfactual account, Lewis' (2004) "influence" account admits differential causal contribution of causes. According to this account, causation is a matter of counterfactual covariation between modally fragile alterations, such that

"C influences e if and only if there is a substantial range of  $c_1, c_2, \dots$  of different not-too-distant alterations of  $c$  (including the actual alteration of  $c$ ) and there is a range of  $e_1, e_2, \dots$  of alterations of  $e$ , at least some of which differ, such that if  $c_1$  had occurred,  $e_1$  would have occurred, and if  $c_2$  had occurred,  $e_2$  would have

---

<sup>12</sup> See Hart and Honore's classic (1985), or Moore's more recent (2009).

occurred, and so on.” (Lewis, 2004)

Influence is meant to track “a pattern of dependence of how, when and whether [e occurs] upon how, when and whether [c occurs].” (2004 p. 190)

With this theory in hand, consider a case of preemption: Suzy and Billy both shoot Victim at the same time; Suzy’s bullet kills Victim; Billy’s bullet hits Victim slightly after Suzy’s. Here, Suzy’s bullet preempts Billy’s bullet from killing Victim. According to Lewis, alterations of the preempting cause (Suzy’s bullet) covary counterfactually with alterations of the effect more than the preempted cause (Billy’s bullet) covaries with the effect. Make small changes in, say, the time or manner of Suzy’s shooting, and the result is a group of corresponding changes in time or manner of Victim’s death. Not so with Billy’s shooting: unless it is altered so radically that it impacts Victim instead of Suzy’s, alterations in whether, when, and how it occurs do not result in major changes in whether, when, and how Victim’s death occurs.

But does influence capture the concept of differential causal contribution relevant to moral responsibility? Consider (Bad Politician) under the influence view. An alteration on Bad Politician’s button-pushing (for example, a slightly earlier and slightly more vigorous button-pushing) results in a small variation in the way that the weapon-launching occurs: it occurs a few seconds earlier. When we alter Bad Scientist’s placement of the weapon on the platform, there are many more differences in the launching of the weapon: it launches in a much different way, and with many more physical differences, than if Bad Politician had just pressed his button one second earlier. There is a stronger pattern of counterfactual covariation between alterations on Bad Scientist’s actions and the weapon launching than between Bad Politician’s actions and the weapon’s launching. And yet, Bad Politician is an intuitively greater contributor to the launching than is Bad Scientist. Obviously, this is due in part to his malicious intent. But setting that aside, it still seems obvious that the strength of patterns of counterfactual dependence between modally fragile alterations has very little to do with the moral analysis of the case.

The problem is that whether, when, and how dependence doesn’t necessarily track the kind of differential causal contribution relevant to moral assessment. The account is bound to yield bizarre judgments in examples such as (Bad Politician), where patterns of

counterfactual covariation between alterations are correlated with whether, when, and how energy is transferred to the outcome. While Lewis admits that patterns of counterfactual dependence between alterations need not track physical energy transfer, it is hard to see how we can massage the theory to give an account of causal differentiation relevant to moral responsibility.<sup>13</sup> Thus while the influence account includes within it a notion of differential causal contribution, it doesn't capture the notion of causal differentiation that plays into moral assessment.

Finally, I will briefly assess the potential of the interventionist approach vis à vis causal differentiation. Roughly, the interventionist approach holds that causation occurs when intervening on A produces a change in B owing to the manipulation of A.<sup>14</sup> Interventionists posit *causal models* of counterfactual relationships between variables. Variables represent causal *relata*. Causal relations are *model relative*: relations hold between variables in some models but not in others, depending on how the models are constructed and what the variables represent. Interventionists are counterfactual theorists, since the central causal relation in question within each causal model involves counterfactual dependence.<sup>15</sup> While simple counterfactual theorists must posit what must be held fixed in evaluating counterfactuals, interventionists must posit what to hold fixed so that a change in B is due to the manipulation in A rather than background causal history or irrelevant causal factors.

Interventionist counterfactuals gain explanatory power over the simple counterfactual approach in two ways. First, interventionists can account for finer-grained counterfactuals insofar as their antecedents can be made more complex than those of the simple counterfactual theory. For example, rather than simply evaluating the claim "Is it the case that if Suzy hadn't shot, Victim wouldn't have died?", the interventionist can ask "Is it the case that if Suzy hadn't shot, but Billy had shot, would Victim have died?" The interventionist approach can thus, in some cases, be used to model quantitative differences in causal contributions, as well as account for preemption and

---

<sup>13</sup> Notable exceptions include cases in which hastening or delaying an outcome is morally significant.

<sup>14</sup> Hitchcock (2001) and Woodward (2003), among others, develop the interventionist account in detail.

<sup>15</sup> For an argument that causal models don't add much explanatory power to counterfactuals, see Briggs (2012).

overdetermination in a way that the simple counterfactual approach cannot.<sup>16</sup>

The second way that interventionism gains explanatory power over the simple counterfactual approach is that it can account for moral intuitions about causation. Because the default and deviant variables in causal models are decided relative to our causal interests, they can be set in such a way that they represent the appropriate moral intuitions. For example, in attempting to account for causal differentiation in Bad Politician, the interventionist might include intentions in the setting of the variables themselves. Similarly with problems about causation by omission that I discuss in the following section: whether or not a particular omission counts as a cause of an outcome depends on whether there is a norm or expectation in place that the omitted event occur.

However, interventionism falls heavily on the side of *causal explanation* rather than *causation*. Let the project of *causal explanation* be that of describing causal relations through the lens of human-relative interests and uses. Let the project of the metaphysics of *causation* to be that of modeling the mind-independent world as it actually is. Approaches to causation that take the relation to directly depend on human-created concepts such as norms, expectations, contextual judgments, and “normal” versus “abnormal” causal paths, all collapse the distinction between causation and causal explanation. Interventionist theories, like contextualism, pragmatism, and contrastivism, intertwine human judgments and contexts with the causal relation itself. Interventionist models require mind-dependent judgments about variables, including which parts of a causal scenario are manipulable in the relevant sort of way. Interventionism and other anthropomorphic approaches to causation differ from metaphysical theories of causation in their *explananda*: the former uses human intuition to model causal claims, whereas the latter tries to model causation apart from those interests.

Now, it might turn out that only causal explanation, rather than causation, can offer an account of causal differentiation that comports with our moral intuitions. Whether to abandon causation for causal explanation is a complicated methodological topic for a different time. For the present purposes, I see little difference between interventionists that lean heavily on human intuition, and straightforward contextualist and pragmatist approaches that hold that whether or not *c* is a cause of *e* is mostly a

---

<sup>16</sup> For interesting empirical work on this issue, see Gerstenberg and Langnado (2015).

matter of collective human mental activity.<sup>17</sup> Such approaches can and do model causation in a way that comports with our moral intuitions, but that's because causation, on those views, is bound up with the moral intuitions themselves. So while interventionism better explains causal differentiation than the simple counterfactual account, much of its extra explanatory power regarding moral responsibility owes to its inclusion of human intuition rather than its extra causal resources.

### 3.4 Omissions and Proportionality

It is particularly hard to pin down causal differentiation in cases of collective causation by omission. Consider the following example:

(Negligent Subway Buddies) Two buddies—one a Black Belt in martial arts, and the other untrained and inept—witness a mugging in a subway. Together, both could have stopped the mugging, though Black Belt would have been more useful. Neither alone could have stopped the mugging. But both buddies stand by while the mugging occurs and do nothing.

Intuitively, Black Belt is more causally responsible for the occurrence of the mugging than is Inept. For though both Black Belt and Inept failed to save the victim, one had more ability to do so than the other. It is natural to hold that Black Belt is more morally responsible for Victim's mugging than Inept. And note that famous cases of causation by omission share a similar structure. For example, wealthy people who fail to give to charity are intuitively more morally responsible for the deaths of starving children than those who have very little to give.<sup>18</sup>

Consider another example of causal differentiation in causation by omission, this time involving proportionality luck:

(Debtors) Joe's rent is \$800. Joe has no ready cash, but four friends owe him

---

<sup>17</sup> In my "Causal Idealism" (forthcoming), I argue that contextualism, contrastivism, and pragmatism about causation are explanatorily on par with causal idealism, the view that causation is entirely a projection of human mental activity.

<sup>18</sup> See Singer (1972) for this famous argument, and also Bernstein (2014) for a metaphysical model of such cases.

money. Two friends owe him \$300 each, and the other two friends owe him \$500 each. Every friend promised to pay Joe back by the start of the month, but none do. Joe does not make rent, and ends up homeless.

There are several notable features of this case. First, it involves causal differentiation: intuitively, the individuals who failed to pay back \$500 are more responsible than the individuals who failed to pay back \$300. Second, it involves proportionality luck. For each debtor, it is a matter of luck which or how many other debtors repaid or failed to repay Joe. For example, had a friend who owed Joe \$500 and a friend who owed Joe \$300 each paid Joe back, then Joe would not have been homeless, and the other two debtors would not have been responsible for the homelessness. Assuming that the friends did not know about each other, it was a matter of bad moral luck that each other friend was delinquent on his or her debt. For had the other friends made good on their debts, Joe would have paid his rent, and any individual debtor would have been off the (moral) hook for Joe's ensuing homelessness.

As in non-omissive cases, no leading theory of causation accounts for such a difference in causal responsibility. For some theories of causation, the problem is the same as in ordinary, non-omissive cases of differential causal contribution. Others face unique problems with causation by omission.

The transfer of energy theory does not apply to causation by omission, for there are no causal *relata* to or from which energy can be transferred. One might try to apply the theory counterfactually. One could ask, for example: *if* the debtors had repaid Joe, which of the debtors would have transferred more energy to him? Or: *if* Black Belt and Inept had interfered in the mugging, which would have transferred more energy to Victim? But this line of reasoning only reinforces the absurdity of employing an energy transfer account of causation to account for causal differentiation in moral contexts. Clearly, differences in energy transfer have little relevance to moral intuitions in these cases.

The simple counterfactual theory of causation cannot account for differential causal contribution by omission in precisely the same way that it cannot account for "normal" causal differentiation. Either the outcome is counterfactually dependent on a particular cause, or it is not. In *Negligent Subway Buddies*, Victim's mugging

counterfactually depends on both buddies' failures to intervene. And in Debtors, Joe's failure to pay the rent counterfactually depends on a random collection of any three friends' omissions. But the account does not capture the intricacies or specifics of each agent's causal contribution.

Nor does Lewis' influence theory of causation do the job. Though Lewis extends his influence account to causation by omission by holding that there can be alterations on events *or* omissions, he also admits that "It doesn't make sense for two distinct absences to differ slightly in detail." (2004, p. 102). And even if one tries to make sense of alterations on omissions—for example, altering the way Black Belt sits in the subway car ever so slightly—such a change would not result in a corresponding change in the way Victim's mugging occurs.

Famously, counterfactual theories of causation also face the problem of profligate causation by omission. That problem involves potentially infinite numbers of true omissive counterfactual claims. For example, the counterfactual "Had Barack Obama not failed to pay Joe \$800, Joe would not have been homeless" is true, as is the counterfactual "Had Steven Segal not failed to appear out of nowhere and stop the subway mugging, the subway mugging wouldn't have occurred". Yet, intuitively, neither of these is a cause of the relevant outcomes. This is another way in which counterfactual theories fail to accommodate differences in causal contribution: too many omissions come out as (equal) causes of each outcome.

I conclude that none of the preceding leading theories of causation have the resources to account for the kind of causal differentiation relevant to moral responsibility. This is an obstacle to the creation of a more precise formulation of **Proportionality**, and also an obstacle to a more careful analysis of proportionality luck. Though there are several other causal theories I have not discussed here, I suspect that they fall short in similar ways.

#### 4.0 Lessons

My goal has been to show that the relationship between causation and moral responsibility so often used in moral assessment is much trickier than previously

imagined. It is also particularly methodologically fraught if current causal theories are to be the guides. Even the simple idea that an agent's causal responsibility for an outcome is proportionate to her moral responsibility for an outcome generates far-reaching consequences and numerous causally weighty interpretations. While there are interesting positive lessons to be drawn from **Proportionality**, including a heretofore unexplored kind of resultant moral luck, the principle reveals that our leading metaphysical theories of causation are generally not cut out for the ambiguities and intricacies of moral evaluation. There is much work to be done before we can trust that linking moral responsibility to metaphysical theories of causation clarifies our theories rather than obfuscates our thinking on these matters.<sup>19</sup>

---

<sup>19</sup> I owe thanks to Carolina Sartorio and an anonymous referee for helpful feedback on this paper. I also thank Caroline Arruda, Elizabeth Harman, Hallie Liberto, Geraldine Ng, and Brad Skow for their commentaries. Finally, I would like to thank audiences at Rocky Mountain Ethics Congress, the 2015 Pacific APA Moral Overdetermination Symposium, the University of Ghent, the 2015 Bellingham Summer Philosophy Conference, Arizona State University, the New Orleans Workshop on Agency and Responsibility, the Collège de France, the Institut Jean Nicod, Australian National University, and Collective Action and the Law for feedback on this paper in the various forms it has taken.

## References

- Bernstein, Sara. (2014). "Omissions as Possibilities". *Philosophical Studies*. 161:1-23.
- (forthcoming). "Causal and Moral Indeterminacy" *Ratio*.
- (forthcoming). "Causal Idealism", Idealism: New Essays in Metaphysics, eds. Tyron Goldschmidt and Kenny Pearce, Oxford University Press.
- \_\_\_\_ (book ms) What Might Have Been: Causation and Possibility.
- Briggs, Rachael. (2012). *Philosophical Studies* 160 (1):139-166.
- Dowe, Phil. (2000). Physical Causation. Cambridge: Cambridge University Press.
- Lagnado, D. A. & Gerstenberg, T. (2015). A difference-making framework for intuitive judgments of responsibility. In Oxford Studies in Agency and Responsibility (Volume 3), David Shoemaker (Ed.).
- Hall, N. (2004). "Two Concepts of Causation". In J. Collins, N. Hall, and L. A. Paul (eds.), Causation and Counterfactuals. MIT Press, Cambridge, MA.
- Hart, H.L.A., and Tony Honoré. (1985). Causation in the Law, 2nd ed., Oxford: Clarendon.
- Hitchcock, Christopher. (2001). "The Intransitivity of Causation Revealed in Equations and Graphs", *The Journal of Philosophy*, 98(6): 273–99.
- Lewis, David. (1973b). "Causation", *Journal of Philosophy*, 70: 556–67. Reprinted in his (1986a).
- (1979). "Counterfactual dependence and time's arrow". Reprinted in *Philosophical Papers Vol. II*, NY: Oxford, 32-66.
- (2004). "Causation as Influence", in Collins, Hall, and Paul, pp. 75–106.
- Moore, Michael S. (2009). Causation and Responsibility. Oxford: Oxford University Press.

Sartorio, Carolina. (2013). "Resultant Luck", *Philosophy and Phenomenological Research* (84).

----- (forthcoming). "A New Form of Moral Luck" in A. Buckareff, C. Moya, and S. Rossell, Agency and Moral Responsibility. New York: Palgrave-Macmillan.

Salmon, Wesley. (1994). "Causality without Counterfactuals" *Philosophy of Science* 61:297-312.

Singer, Peter. (1972). "Famine, Affluence, and Morality" *Philosophy and Public Affairs* (1) 3:229-243.

Wolf, Susan. (2001). The Moral of Moral Luck. *Philosophic Exchange*. Volume 31: Issue 1.

Woodward, James. (2003). Making Things Happen: A Theory of Causal Explanation, Oxford: Oxford University Press.