

# Responsibility and the ‘Pie Fallacy’

Alex Kaiserman  
Draft of 26/11/2020

**Abstract:** Much of our ordinary thought and talk about responsibility exhibits what I call the ‘pie fallacy’ – the fallacy of thinking that there is a fixed amount of responsibility for every outcome, to be distributed among all those, if any, who are responsible for it. The pie fallacy is a fallacy, I argue, because how responsible an agent is for some outcome is *fully grounded* in facts about the agent, the outcome and the relationships between them; it does not depend, in particular, on how responsible anyone else is for that same outcome. In this paper, I explore how the pie fallacy can arise by considering several different kinds of case in which two or more agents are responsible for the same outcome. I’ll end with some brief remarks on the potential consequences of my arguments for how to think about responsibility in war.

## 1. Introduction

In his “Sharing Responsibility”, Michael Zimmerman (1985a) drew attention to an important fallacy in our ordinary thinking about responsibility:

Why can it be misleading to talk of people sharing ... responsibility for an outcome? Some would say that this is so because “our idea of responsibility requires that it should be uniquely ascribed.” [Quoting Weinryb (1980: 9)] But to say this seems to me quite mistaken; surely more than one person can be responsible for the same outcome. I would say, on the contrary, that talk of people sharing responsibility can be misleading because such talk seems often, perhaps always, to suggest that the responsibility of each person involved ... is *diminished* simply by virtue of the fact that others are also involved and are responsible. The suggestion seems to be that there is just so much responsibility to go around – the “pie” is just so large – and the more people involved, the smaller the share for each. *If* this is ever the suggestion, it is one that I deny to be correct; I deny that there is ever any such diminution. (Zimmerman 1985a: 115).

Call the fallacy Zimmerman is drawing attention to here – according to which there is a some fixed ‘quantity’ of responsibility available for every outcome, to be distributed among all those, if any, who are responsible for it – the ‘pie fallacy’. The pie fallacy is a fallacy, Zimmerman argues, because an agent’s degree of responsibility for an outcome is not in general dependent on how responsible anyone else is for that same outcome. More precisely:

**Individualism:** An agent A’s degree of responsibility for some outcome *o* is *fully grounded* in facts about A, *o* and the relations between them.

Note that Individualism is consistent with the possibility of someone else *causing* the facts which ground A’s degree of responsibility for *o* on some occasion. Suppose B slips A a drug that causes uncontrolled violent outbursts, one of which ends up harming V. What grounds the fact that A isn’t responsible for V’s harm, according to Individualism, are facts pertaining only to A, such as the fact that he had no control over his actions; this fact merely happens to have been caused, on this occasion, by B. The point is we can always tell a *complete* story about what makes it the case that A is responsible to degree *x* for *o* in terms that don’t mention people other than A.

In this paper, I propose to revisit Zimmerman’s arguments against the pie fallacy. My motivations are twofold. Firstly, it seems to me that the pie fallacy is still pervasive in both ordinary and

academic discussions of responsibility.<sup>1</sup> It is presupposed by the common talk of ‘distributing’ or ‘apportioning’ responsibility, and of responsibility being ‘transferred’ or ‘shifted’ from one person to another, for example. It also colours much of our talk of responsibility in war. When a state or individual is described as ‘wholly’ or ‘fully’ responsible for a conflict, the unspoken implication is often that they are thereby *solely* responsible for it.

But secondly, though I share Zimmerman’s commitment to Individualism, I nevertheless think he is wrong about many of its consequences in cases where multiple people are responsible for the same outcome. In some causal structures, there are interesting functional dependencies between agents’ degrees of responsibility that are missed in Zimmerman’s discussion. The goal of this paper is therefore to explore these dependencies in more detail. I start (§2) with a brief explanation of what I take to be the three dimensions of responsibility for outcomes: the causal dimension, the agential dimension, and the epistemic dimension. I then discuss three different kinds of case in which two agents, A and B, are both responsible for the same outcome *o* – where A and B make independent contributions to bringing *o* about (§3); where A’s action is a cause of B’s causing of *o* (§4); and where *o* is caused by a *group*, B, of which A is a member (§5). My goal in each case will be to establish whether A is *less* responsible for *o* than he would have been if B hadn’t been involved and all else had been equal. Zimmerman appears to think the answer is ‘no’ in every case; I’ll show that the true picture is in fact more complex than this. I’ll end (§6) with some broader reflections on the consequences of my discussion for how to think about responsibility in war.

## 2. Preliminaries

### 2.1. Clarifying the Target

It’s widely recognised that the phrase ‘is responsible for’ has several different meanings. Consider the following sentence, for example:

- (1) A short-circuit was responsible for a signal failure on the Underground; so although I was responsible for making sure Julie got to school on time, I’m not responsible for Julie’s failure to get to school on time.

The first occurrence of ‘responsible’ in (1) is used basically as a synonym for ‘caused’. The second occurrence of ‘responsible’ is used to talk about the duties I had, perhaps as Julie’s parent.<sup>2</sup> The third occurrence, meanwhile, is used to denote a certain kind of relation between agents and token actions, omissions, or outcomes (in this case, between me and Julie’s failure to get to school on time). It is the third sense of ‘responsible’ that I will be concerned with in this paper.

Some philosophers will protest that I have still not specified my target narrowly enough. They will insist that even the third occurrence of ‘responsible’ in (1) is ambiguous between multiple senses, and until I specify which of these senses I’m talking about, it won’t be clear how my claims about ‘responsibility’ should be interpreted. Arguments for this conclusion usually proceed by first noting that responsibility is closely connected to certain kinds of *responses* to agents’ actions – we can

---

<sup>1</sup> One intriguing example of this is the common cognitive behavioural therapy exercise designed to help alleviate strong feelings of guilt or shame about past events (see, e.g., <https://www.psychologytools.com/resource/pie-charts/>). Participants are asked to divide up a pie chart into sections corresponding to how responsible various people were for that event, to help them realise how they might be less responsible for it than they initially thought they were. While I don’t presume to question the therapeutic effects of such exercises, I will be questioning the underlying assumptions about responsibility they appear to make.

<sup>2</sup> The concept in question here is sometimes described, including by Zimmerman (1985a: 115), as ‘forward-looking’ or ‘prospective’, rather than ‘backward-looking’ or ‘retrospective’, but this is a misnomer – as is evident from (1), this use of ‘responsible’ can be used to talk about what I *had* a duty to do in the past just as much as what I *have* a duty to do in the future.

*attribute* an action to an agent, hold them to *account* for that action, demand that they *answer* for that action, and so on – and different such responses seem to be apt in different situations. Therefore, it's claimed, we must distinguish a different sense of 'responsible' for each of these responses, such that to be 'responsible' for one's actions in such a sense *just is* for the corresponding response to one's actions to be apt.<sup>3</sup>

Let me briefly explain why I think these philosophers are mistaken. First, it's worth noting that there are well-established linguistic tests for ambiguity. Consider the following sentence, for example:

(2) The table has four legs and five columns.

(2) sounds distinctly zeugmatic. This is defeasible evidence that 'table' has at least two senses – one referring to items of furniture and another referring to ways of organising information. Now compare (2) with (3):

(3) In part *because* Katya is responsible for Evan's injury, she is an appropriate target for blame and contempt; it would be apt for her to feel guilt and regret; she has a duty to apologise to Evan, to compensate him, and to answer demands to explain or justify her actions; and she is liable to punishment, and to defensive harm to avert future threats to Evan brought about by his injury.

(3) does not sound at all zeugmatic, at least to me. This is defeasible evidence that 'responsible', as it occurs in sentences like 'Katya is responsible for Evan's injury', has just one sense.<sup>4,5</sup> But it also suggests that what 'responsible' picks out in such contexts is a certain relation between agents and outcomes the obtaining of which, in the right circumstances, partially *grounds* the aptness or permissibility praise and blame, punishment and reward, defensive harm and demands for an explanation, an apology or compensation, rather than the other way around.<sup>6</sup>

Notice that this is perfectly compatible with the observation that one can, for example, be liable to defensive harm to avert a threat without being blameworthy for it, or *vice versa*. Consider an analogy with causation: that the Black Death was a cause of both the Peasant's Revolt and the rise of Lollardy doesn't imply that revolts can't happen without the rise of new religions (or *vice versa*), since for one thing, the conditions under which mass disease outbreaks cause revolts are different to those under which they cause the rise of new religions. Similarly, that Katya's responsibility for Evan's injury on a given occasion is a ground of both her blameworthiness and her liability to defensive harm doesn't imply that one cannot be liable to defensive harm without being blameworthy (or *vice versa*), because the conditions under which responsibility for an injury partially

---

<sup>3</sup> See, in particular, Shoemaker (2011), drawing on an earlier distinction from Watson (1996) (though Watson prefers to talk of two 'faces' of responsibility, rather than two senses of 'responsible'); see also Pereboom (2014: 2-4).

<sup>4</sup> Notice, by contrast, that 'I was responsible for getting Julie to school on time and for her failure to get to school on time' *does* sound fairly zeugmatic, not unlike 'He took his hat and his leave' or Dickens' famous 'She went home in floods of tears and a sedan chair'. This corroborates the hypothesis that the second and third occurrences of 'responsible' in (1) do indeed have different meanings.

<sup>5</sup> As Tadros points out (in 'Responsibility as Personal Significance', unpublished manuscript), the claim that 'responsible' has just one sense is perfectly compatible with the claim that there are different *kinds* of responsibility, just as the fact that 'toothbrush' has just one sense is compatible with there being different kinds of toothbrush. But this needn't prevent one from theorizing about responsibility (or, indeed, toothbrushes) in general.

<sup>6</sup> As Tadros (*supra* note 5) also points out, if being responsible for Evan's injury *just was* for, say, blame to be an apt response, then (3) would presumably entail that it is apt to blame Katya for Evan's injury in part because it is apt to blame her for it, which is absurd. Yet (3) is not absurd – indeed, it is arguably one of those platitudes about responsibility that any successful account of it should be able to accommodate, or at the very least explain away.

grounds blameworthiness might be different from those under which it partially grounds liability to defensive harm.

There is no doubt more to say here. Nevertheless, I will assume for the remainder of the paper that there is a single, univocal relation at issue in talk of responsibility for outcomes. Those who continue to suspect that such talk is ambiguous are free to resolve that ambiguity in whatever way they see fit.

## 2.2. *Responsibility, Blameworthiness and Liability*

An important platitude about responsibility is that it comes in *degrees*. Someone can be more or less responsible for an outcome, depending on their degree of agential involvement in bringing the outcome about. But we should be careful to distinguish degrees of responsibility from other things which also come in degrees. Firstly, an agent's degree of responsibility for an outcome is not the same as the amount of praise or blame they deserve as a result. For one thing, one can be fully responsible for normatively neutral outcomes (the position of one's hand, for example), for which no praise or blame is due. But arguably, one can also be blameworthy for being responsible for a good outcome, or praiseworthy for being responsible for a bad outcome. Suppose a trolley network operator, fully informed and in full control of her faculties, decides to switch a runaway trolley into the path of one person to save the hundred people it would otherwise have killed. The operator is responsible for the victim's death, in my view (see §4, below); and of course a death is, by any standards, a bad outcome, despite it being the inevitable result in this instance of preventing an even worse outcome. Nevertheless the trolley operator is presumably to be *praised*, not blamed, for doing the right thing in switching the trolley; all the more so in light of the fact that doing the right thing in this case required making the incredibly difficult, and no doubt traumatising, decision to foreseeably bring about someone's death. How much agents should be praised or blamed for an outcome depends on the character and strength of their *duties* with respect to the outcome, as well as, perhaps, their degree of responsibility for it.

Secondly, an agent's degree of responsibility for an outcome is not the same as the range of *liabilities* they acquire as a result. This is important because what one is liable to suffer to avert a threat arguably *does* depend on who else is responsible for that threat and to what degree. If a threat can be averted only by either killing A or by killing B, and A is more responsible for the threat than B, then arguably, all else being equal, A is liable to be killed and B is not, regardless of how responsible B is (McMahan 2011: 551-5; c.f. Frowe 2014: ch.4).<sup>7</sup> It's also widely assumed that what one is liable to pay a victim in compensation depends on how responsible others are for the victim's injury, because the total amount of compensation paid has to ultimately add up to what the victim lost. One can dispute these claims;<sup>8</sup> but the important point is that responsibility and liability are different things, and there is no straightforward relationship between one's degree of responsibility for an outcome and what one is liable to suffer or pay as a result.

## 2.3. *The Dimensions of Responsibility*

Giving a proper account of responsibility, and of the ways in which it comes in degrees, is of course beyond the scope of this paper. But it will help for what follows to have the broad contours of the concept in view. As I see it, there are three dimensions to responsibility for outcomes. The first is the strength of the *causal* connection between the outcome and the agent's actions – the extent to which those actions contributed to bringing the outcome about. Some people deny that there is any sense to be made of the notion of degrees of causal contribution – Zimmerman

---

<sup>7</sup> Conversely, Tadros (2016) has argued that agents can in some cases be liable to be killed to avert threats for which they don't bear *any* responsibility whatsoever, so long as they are sufficiently culpable with respect to the threat.

<sup>8</sup> I have argued against the latter claim, in particular, elsewhere; see Kaiserman (2017b: §4).

himself, for example, explicitly says that causal contribution is “a factor which admits of no degrees” (Zimmerman 1985a: 117). I have sought to challenge this contention in previous work by articulating a novel account of degrees of causal contribution (Kaiserman 2016). The account takes as its starting point the familiar idea that causes are non-redundant members of sets of events that were *jointly sufficient*, in the circumstances, for their effects. More precisely:  $X_1, \dots, X_n$  collectively caused  $Y$  if and only if  $X_1, \dots, X_n$  were jointly sufficient in the circumstances for  $Y$ , and no proper sub-plurality of  $X_1, \dots, X_n$  were jointly sufficient in the circumstances for  $Y$ . To be a *cause* of  $Y$  is then to be one of a plurality of events that collectively caused  $Y$  (just as to be an *author* of a book is to be one a plurality of people who collectively authored it).

Causation itself is a non-scalar relation, on this view, since sufficiency is all-or-nothing. But just as multiple authors can contribute to different extents to the authoring of a book, so too can multiple causes contribute to different extents to the causing of an effect, by contributing to different extents to making the plurality as a whole jointly sufficient for the effect. In Kaiserman (2016) I cash this out in probabilistic terms to arrive at a precise measure of one event’s degree of contribution to a causing of some effect. But the details won’t be necessary for our purposes – the important point is just that we can make sense of the idea of an action’s degree of contribution to bringing a certain outcome about, and this is plausibly one of the factors that go into determining an agent’s degree of responsibility for that outcome.

The second dimension of responsibility is the strength of the *agential* connection between the agent and her action. If A, a kleptomaniac, compulsively steals B’s watch, then A is not responsible for B’s loss, even though his actions caused the loss, because his action wasn’t *free*. Although it was in some sense because of A’s *action* that B is now deprived of her watch, it can’t properly be said to be because of A that she is deprived of it – stealing B’s watch was not something over which A had any control. There are different views, of course, on what it is for an action to be free in this sense, and how, if at all, free will comes in degrees. My own view<sup>9</sup> is that an agent acts freely to the extent to which they were sensitive, in acting, to *reasons* for or against acting that way, where degrees of reasons-sensitivity are themselves to be understood in causal, rather than modal, terms. But again the details won’t matter for our purposes here.

The third and final dimension of responsibility is the strength of the *epistemic* connection between the agent and the outcome. If A decides to take her dog for a walk around the park on a bright day, and the dog is later struck by lightning, then A is not responsible for the dog’s death, even though it was in part the result of decisions A freely made, because A couldn’t reasonably have foreseen what the consequences of those decisions would be. Arguably A need not have foreseen *o* itself in order to be responsible for it, so long as she knew, or *should* have known, that her action posed a sufficiently large risk of a certain kind of outcome occurring, of which *o* was an instance. But beyond that, the nature of the epistemic connection required for responsibility is contested.<sup>10</sup> I won’t have any more to say on the matter here.

With these preliminaries in place, we are now ready to return to the question at hand: in what ways might one person’s degree of responsibility for some outcome depend on how responsible others are for that same outcome?

### 3. Overdetermination and Joint Causation

The first and simplest way for two people to both be responsible for the same outcome is for each of them to make an independent contribution to bringing the outcome about. But it’s important

---

<sup>9</sup> See A. Kaiserman, ‘Reasons-Sensitivity and Degrees of Free Will’, unpublished manuscript.

<sup>10</sup> For the current state of play, see Robichaud and Wieland (2017).

to further distinguish two quite different ways this might happen. Consider the following three cases:<sup>11</sup>

*One Assassin*: A shoots V. The bullet strikes V with deadly force. V dies.

*Two Assassins*: A and B both shoot V. Each bullet strikes V with deadly force at the same time. V dies. All else is the same as in *One Assassin*.

*Two Assassins (Hardy Victim)*: A and B both shoot V. Both bullets strike V at the same time, and V dies. But unbeknownst to A and B, V was *hardy* – two bullets were required to kill him. All else is the same as in *One Assassin*.

*Two Assassins* is a case of *overdetermination*. A's and B's shots were *individually* sufficient, in the circumstances, for V's death; the death was therefore caused *twice over*. *Two Assassins (Hardy Victim)*, meanwhile, is a case of *joint causation*. A's and B's shots were *jointly*, but not individually sufficient for V's death; the death was therefore caused only once, albeit *by* two events. Zimmerman considers both kinds of case in his paper (Zimmerman 1985a: 116-9). His conclusion is that A is no less responsible for V's death in either *Two Assassins* or *Two Assassins (Hardy Victim)* than he is in *One Assassin*. This section will reconsider Zimmerman's arguments for that conclusion.

There is clearly no difference in A's degree of responsibility between the three cases along either the agential or epistemic dimensions. The question is therefore whether there is any difference between the cases in the strength of the *causal* connection between A's shot and V's death. I agree with Zimmerman that there is no such difference between *One Assassin* and *Two Assassins*. In both of these cases, V's death is *fully* caused by A's shot; the only difference is that in *Two Assassins* the death was *also* (fully) caused by B's shot. The mere addition of B does nothing to weaken the strength of the causal connection between A's action and V's death. It is as much *because of A* that V's death occurred *Two Assassins* as it is in *One Assassin*.

Note that this is already enough to show that the pie fallacy is a fallacy. If there were some fixed 'amount' of responsibility for V's death to be distributed among those who are responsible for it, then the fact that A is solely responsible for V's death in *One Assassin* but not in *Two Assassins* would have to imply that he is less responsible in *Two Assassins*. Yet this is not the case. Indeed talk of A and B 'sharing' anything in *Two Assassins* seems quite inappropriate – they each, as it were, have their own pie.

The situation seems to me quite different, however, with *Two Assassins (Hardy Victim)*. Whereas in *One Assassin* A's shot *caused* V's death, in *Two Assassins (Hardy Victim)* it merely *contributed*, along with B's shot, to the causing of V's death. While it was *partially* because of A that V's death occurred, therefore, it wasn't *wholly* because of him. It follows that A is less responsible for V's death in *Two Assassins (Hardy Victim)* than he is in the other two cases; how much less will depend, moreover, on the relative contributions of A's shot and B's shot to bringing the death about. The more responsible along the causal dimension B is for the death, the less responsible A will be.<sup>12</sup>

Note however that this is all perfectly compatible with Individualism. What grounds A's degree of responsibility in *Two Assassins (Hardy Victim)* is the degree of contribution he made to the causing

---

<sup>11</sup> I borrow these cases from Bernstein (2017).

<sup>12</sup> Recall that A in *Two Assassins (Hardy Victim)* didn't *know* that V was hardy; presumably, therefore, he believed his shot to be individually sufficient in the circumstances for V's death, though in fact it wasn't. Moreover, whether V turned out to be hardy is something over which A had no control. All this might imply that A is as blameworthy in *Two Assassins (Hardy Victim)* as he is in *One Assassin*. (Then again, it might not – as several authors have noted, it's not clear whether the grounds of blameworthiness are ever truly 'within one's control'; see Sartorio (2015) for more on this kind of 'moral luck'.) As explained in §2, however, my concern here is with responsibility for outcomes, not blameworthiness.

of V's death; that the other cause happens to the action of another responsible agent is irrelevant. We can stipulate that B isn't at all responsible for V's death – because he reasonably believed that the gun was filled with blanks, perhaps – or indeed replace B's action with some non-agential event – perhaps the second gun was on a shelf and happened to go off spontaneously in V's direction at the same time as A pulled his trigger – and this would make no difference to A's degree of responsibility for V's death. That one agent was partially responsible for some outcome doesn't imply that anyone else was also partially responsible for it.

Zimmerman, however, denies that there is any such difference between *Two Assassins (Hardy Victim)* and *One Assassin*. He provides the following argument in support of this claim. For A's shot to be a cause of V's death, in both cases, “requires what may be called the co-operation of other occurrences and conditions” (Zimmerman 1985a: 116). For example, it requires that the laws of physics be true, that there is no obstacle between A and V, that V is not wearing a bulletproof vest, and so on. The only difference is that in *Two Assassins (Hardy Victim)*, the conditions from which co-operation is required include facts about the actions of B. But, Zimmerman argues, there is no “reason for thinking that this difference between the cases in what occurrences are involved in the co-operation at issue is of any relevance to the assessment of the degree to which A is responsible for the outcome” (Zimmerman 1985a: 117).

It seems to me, however, that there is a very important difference between B's shot in *Two Assassins (Hardy Victim)* and, say, the lack of any obstacle between A and V; namely, that relative to the explanatory perspective relevant to attributions of responsibility, B's shot is a *cause* of V's death, whereas the lack of obstacles is a mere *background condition*. There is, of course, a venerable tradition going back at least to Mill (1868) of refusing to recognise such a distinction between causes and background conditions – between “the real cause [and] that without which the cause would not be able to act as a cause” (Plato, *Phaedo* 99a-b). But the fact is we do rely on these so-called “principles of invidious discrimination” (Lewis 1973: 559) in much of our causal theorizing, including our theorizing about responsibility. This was clearly appreciated by J. L. Mackie in his influential work on causation. “[C]ausal statements”, Mackie argued, “are commonly made in some context, against a background which includes the assumption of some *causal field*” (Mackie 1974: 34); this ‘field’ is not “part of a cause, but is rather a background against which the causing goes on” (Mackie 1974: 63). Which facts are ‘held fixed’ as part of the causal field is, of course, sensitive to context – it depends on our explanatory purposes. But “in determining responsibility we do not choose the causal field quite arbitrarily” (Mackie 1955: 145) – there is a particular kind of causal field which is relevant to attributions of responsibility for outcomes. Though it's hard to characterise this causal field in general terms, it's intuitively clear, in many cases, what for the purposes of attributing responsibility ought to be regarded as causes rather than mere background conditions. Thus while there certainly is a context relative to which V's death in the three cases above was caused only by a massive plurality of events, including V's lack of a bulletproof vest, the absence of any obstacles, the laws of physics, and so on, the context of relevance to responsibility seems rather to be one in which facts about the laws of physics, the relative positions of A and V, and so on, are held fixed; and relative to this context, V's death was individually caused by A's shot in *Two Assassins*, but only *jointly* caused by A's shot and B's shot in *Two Assassins (Hardy Victim)*.

Chockler and Halpern (2004; see also Halpern 2015, 2016) have recently proposed a measure of ‘degrees of responsibility’ which has quite different implications about the three assassin cases. On their view, A's degree of responsibility for an outcome *o* is equal to  $1/(N+1)$ , where *N* is “the minimal number of changes that have to be made to obtain a contingency where [*o*] counterfactually depends on [A's action]” (Halpern 2015: 92). However exactly ‘changes’ are

individuated,<sup>13</sup> it seems clear that A's 'degree of responsibility' for V's death on this view will be equal to 1 in *One Assassin* and *Two Assassins (Hardy Victim)* (since V's death actually counterfactually depends on A's action in these cases, so  $N = 0$ ), but less than 1 in *Two Assassins* (since V's death doesn't counterfactually depend on A's action in this case, so  $N > 0$ ) – the opposite result to the one I have argued for above.

It seems to me, however, that Chockler and Halpern here commit exactly the fallacy Zimmerman is concerned with exposing. They mainly focus on voting cases, where every cause is the action of a responsible agent – thus, they argue, someone who votes for an outcome is more responsible for it when the motion passes 5-4 than when it passes 9-0, because in the latter case “responsibility is more diffuse” (Chockler and Halpern 2004: 108). But later in the same paper, Chockler and Halpern acknowledge that responsibility might also seem more ‘diffuse’ if a motion is carried 100-99 than if it is carried 1-0, even though in both cases their measure assigns each positive voter degree of responsibility 1. This suggests that what is really driving the intuitions Chockler and Halpern are relying on here is a version of the pie fallacy, the fallacy of thinking that an agent's degree of responsibility for some outcome is diminished merely in virtue of others also being responsible.

## 4. Manipulation and Intervening Agency

A second way for two people, A and B, to both bear some responsibility for the same outcome,  $o$ , is for A to contribute to the causing of  $o$  *via* the actions of B. Cases like these raise two distinct questions: (i) Is A's responsibility for  $o$  diminished by virtue of the presence of B in the causal chain between A's action and  $o$ ?; and (ii) Is B's responsibility for  $o$  diminished by virtue of the presence of A in the causal history of B's action? I'll take each question in turn.

### 4.1. Intervening Agency

As Zimmerman (1985b) notes, there is a long-established doctrine in private and criminal law according to which the actions of ‘intervening agents’ do indeed diminish defendants' responsibility for outcomes. Consider the following two cases, for example:

*Bomb*: A plants a bomb outside V's house and lights the fuse. The bomb goes off, killing V.

*Bomb (Intervening Agent)*: A plants a bomb outside V's house and lights the fuse. Luckily, the fuse fizzles out before reaching the bomb. But it is then spotted by B, another enemy of V's, who happened to be walking by. B opportunistically decides to relight the fuse, and the bomb goes off, killing V. All else is the same as in *Bomb*.

Intuitively A is substantially less responsible for V's death in *Bomb (Intervening Agent)* than he is in *Bomb*, even though it's true in both cases that the death wouldn't have occurred but for A's planting of the bomb. It's tempting to explain this in terms of the pie fallacy – B's responsibility for V's death in *Bomb (Intervening Agent)* ‘relieves’ A of responsibility for it. But it would be a mistake to think in this way. To see this, consider the following case:

*Bomb (Lightning)*: A plants a bomb outside V's house and lights the fuse. The fuse fizzles out before reaching the bomb. However, it is then relit by a lightning strike. The bomb goes off, killing V. All else is as in *Bomb*.

---

<sup>13</sup> For Chockler and Halpern this will ultimately depend on the causal model used to represent the situation of interest.



A's responsibility for V's death also seems diminished in *Bomb (Lightning)*, and for similar reasons.<sup>14</sup> Thus it cannot be *because* of the fact that another agent is also responsible for V's death in *Bomb (Intervening Agent)* that A is less responsible.

Zimmerman similarly concludes that there is "nothing special about intervening agents with respect to the allocation of [responsibility]" (Zimmerman 1985b: 357). But this nevertheless leaves open the question of *why* A is less responsible in *Bomb (Intervening Agent)* than in *Bomb*, if not because of the fact that B is also responsible in the former case. Zimmerman (1985b: 356) tentatively suggests that "the actions of intervening agents are [often] difficult to foresee and control, and this renders many of the . . . consequences of our own actions [in cases with intervening agents] difficult to foresee and control". But V's death was just as foreseeable in *Bomb (Intervening Agent)* as it was in *Bomb*. Of course, A couldn't reasonably have foreseen the *mechanism* by which V's death was eventually brought about in *Bomb (Intervening Agent)*. But requiring an agent to have reasonable foresight of the mechanism by which an outcome is brought about seems too strong as a condition on responsibility. If I kill someone using a gun which, unbeknownst to me, works in a very different way to traditional guns (by electrically charging the bullets and accelerating them through a potential difference across the two ends of the barrel, perhaps), I seem no less responsible for the death in virtue of my ignorance of the workings of the gun.

Instead I think the right place to locate the difference between *Bomb (Intervening Agent)* and *Bomb* is once again along the causal dimension of responsibility (see Kaiserman 2017b: §3). A's act was sufficient in the circumstances for V's death in *Bomb*. But it wasn't sufficient in the circumstances for V's death in *Bomb (Intervening Agent)* and *Bomb (Lightning Strike)*, as evidenced by the fact that the fuse fizzled out after being lit. So whereas A's act caused the death in *Bomb*, it merely contributed to a causing of the death in *Bomb (Intervening Agent)* and *Bomb (Lightning Strike)*. Of course, I'm assuming here that facts about the fuse, the location of V, etc., should be held fixed as background conditions, but facts about the lightning strike or B's murderous intentions should not. This illustrates another important feature of the causal field relevant to responsibility, of which Mackie was also well aware. "[I]n determining responsibility . . . our choice [of causal field] is determined by our moral expectations, our views about what is normal and proper" (Mackie 1955: 145). The word 'normal' has descriptive and teleological, as well as normative flavours – we can use it to describe something as statistically typical, functioning as designed, or permitted by some set of norms. The facts we hold fixed in determining responsibility tend to be those which are normal in one or more of these respects. Thus if B is injured while chasing a horse that has bolted from the paddock that A negligently left open;<sup>15</sup> or dies from minor injuries caused by A after refusing a blood transfusion on religious grounds;<sup>16</sup> or breaks his leg after jumping, in fear of his life, from a careering car recklessly driven by A;<sup>17</sup> our willingness to hold A fully responsible for B's harm seems to turn, not on how foreseeable B's intervening actions were by A, but rather on whether those actions were 'normal' or 'proper' things for B to do in the circumstances. If we judge that, given A's actions, B *had* to act as he did, in the sense of having no reasonable alternative *but* to do so, then we will typically conclude that A's actions were sufficient for B's harm, in the sense of relevance to responsibility.

It might be objected that causal facts are not the sorts of things that should turn out to depend on facts about what is 'normal and proper'. But it's important to distinguish two different kinds of

---

<sup>14</sup> See Hart and Honoré (1985: 162-3), who similarly note that "an act of God, in the sense of an event whose occurrence at the place in question is extraordinary" can "relieve a wrongdoer of responsibility" in the same way that acts of intervening agents can.

<sup>15</sup> See, e.g., *Cutler v. United Dairies (London) Ltd.* [1933] 2 KB 297 (CA).

<sup>16</sup> See, e.g., *R v. Blaue* [1975] 1 WLR 1411.

<sup>17</sup> See, e.g., *Jones v. Boyce* [1816] Stark 1, p. 493.

dependence here. On the view I am describing, what a causal claim *means* – which proposition is expressed by means of it – depends on which facts we’re holding fixed in a given context. But what *grounds* the truth of the proposition thereby expressed itself needn’t have anything to do with the normative facts (see Kaiserman 2017a: §6). Two people may agree on all the causal facts, but nevertheless disagree on whether a particular causal claim is true in a given context, *because* they disagree on which proposition is being expressed by that claim in that context (just as two people may agree on all the non-indexical facts, but nevertheless disagree about the truth of ‘It’s raining here’, because they disagree about which location ‘here’ picks out). This is the kind of disagreement at issue when two people disagree about, say, whether a pedestrian’s injuries were caused by the motorist’s speeding or the pedestrian’s inattentiveness, in a context where attributions of responsibility are salient.

#### 4.2. Manipulation

Let us now turn to question (ii), the question of whether an agent’s responsibility for some outcome is diminished in virtue of having been manipulated into acting as they did by another agent. Once again, it’s important to distinguish between two different ways in which this might happen. Consider the following three cases:

*Trolley*: A runaway trolley is headed towards ten people. B can switch the trolley onto a different track where it will kill just one person, V. He decides to switch the trolley, and V dies.

*Trolley (Moral Coercion)*: A wants to kill V, but lacks the means to do so directly. So she hatches a plan to send a trolley towards ten people, knowing that B will switch the trolley towards V. All else is the same as in *Trolley*.

*Trolley (Direct Manipulation)*: A wants to kill V, but lacks the means to do so directly. But she does have the ability to make B do whatever she wants him to do, having succeeded in hypnotising him earlier. So A hatches a plan to send a trolley down a track and force B, via hypnotic suggestion, to switch it towards V. All else is the same as in *Trolley*. (In particular, B would have switched the trolley independently, had he been given the opportunity to do so.)

In both *Trolley* and *Trolley (Moral Coercion)*, B acts for a particular reason – namely, the fact that ten people will die if he doesn’t switch and only one will die if he does. The only difference is that A happens to have caused that reason to obtain in *Trolley (Moral Coercion)* but not in *Trolley*. In *Trolley (Direct Manipulation)*, by contrast, B does not act for this reason, or indeed for any reason – rather, A’s action *pre-empts* the reasons that would have motivated B to switch the trolley had A’s action not occurred.

It’s commonly claimed that B is not responsible for V’s death in cases like *Trolley (Moral Coercion)*.<sup>18</sup> When one agent coerces another into acting a certain way, the thought goes, responsibility for the consequences of that action “gets shifted” (Kamm 2007: 312) from coercee to coercer (i.e. from B to A, in this case). But as Frowe (this volume) convincingly argues, there seems to be no difference between *Trolley (Moral Coercion)* and *Trolley* which could ground a difference in how responsible B is for V’s death.<sup>19</sup> B is responding, in the same way, to the same reasons to switch the trolley in both cases – how those reasons came about seems irrelevant to the strength of the agential connection between B and V’s death. It seems to follow, therefore, that B isn’t responsible

---

<sup>18</sup> “[P]ersons who are coerced into performing specified actions are commonly regarded as not morally responsible for those actions” (Arnold 2001: 53).

<sup>19</sup> See also Bazargan (2014: 11).

for V's death in cases like *Trolley* either, despite the fact that there is no-one else in *Trolley* for B's responsibility to be 'shifted' to.

This form of reasoning – which proceeds by first articulating a case of manipulation, soliciting the intuition that the subject is not responsible for his actions, before arguing that the manipulation case does not differ in any morally important respects from a different case in which no manipulation occurs – is familiar from the literature on free will, and has been used by incompatibilists and responsibility sceptics to argue that we are responsible for our actions in a much smaller class of cases than is typically assumed.<sup>20</sup> Yet it seems to me that this process of reasoning goes wrong at the outset. B *is* responsible for V's death in *Trolley (Moral Coercion)*, as well as in *Trolley*. Admittedly he might not be as responsible as A, since A intentionally caused V's death whereas B merely caused it as a side-effect of intentionally saving the ten people on the other track. But to conclude from the fact that A is fully responsible for V's death in *Trolley (Moral Coercion)* that B is thereby not at all responsible – because all of B's responsibility has been 'shifted' to A – is to commit the pie fallacy.<sup>21</sup> Responsibility for an outcome is not some fixed quantity transferable from one person to another; that one agent is fully responsible for some outcome doesn't imply that he was solely responsible for it.

To reiterate, in saying that B is responsible for V's death in *Trolley* and *Trolley (Moral Coercion)*, I don't mean to suggest that B is blameworthy as result of his decision to switch the trolley; nor that he is liable *ex ante* to be harmed to avert the threat he poses to V, or liable *ex post* to cover the costs of V's death.<sup>22</sup> Whether B is any of these things depends on much more than his degree of responsibility for V's death; it depends, in particular, on whether he was morally required to switch the trolley. Indeed B may very well be *praiseworthy* as a result of his actions, despite their unfortunate consequences – it was *because of B*, we might say, that only one person died rather than ten.

The situation seems to me quite different, however, with *Trolley (Direct Manipulation)*. It wasn't *because of B*, in this case, that V died rather than the ten people on the other track; B himself had nothing to do what transpired, even though his body was used as a tool by A for bringing her desired outcome about. B certainly had a reason to switch the trolley in *Trolley (Direct Manipulation)*, which he would have acted on had he been given the chance. But he wasn't suitably sensitive, in acting, to this reason, or indeed to any reason – his actions were entirely under the control of A.

It's possible, I think, to get cases in between *Trolley* and *Trolley (Direct Manipulation)*, cases where B acts *partly* on the basis of a reason to act and partly as a result of a (non-reason-conferring) action of A's. Suppose for example that B is neither particularly ethical nor particularly susceptible to hypnotic suggestion. As such, neither the moral reason to switch the trolley (that it would save ten lives at a cost of just one) nor A's act of hypnosis were sufficient by themselves for B to switch the trolley, but both together were jointly sufficient (in the circumstances). In such cases, I would argue, B is *partially* reasons-sensitive in acting; how reasons-sensitive will depend, moreover, on how much that reason contributed, compared to A's hypnosis, to bringing B's action about.<sup>23</sup> There is therefore an interesting dependency between A's degree of responsibility for V's death along the causal dimension and B's degree of responsibility for the death along the agential dimension – the more A's action contributed, via B's action, to bringing about V's death, the less

---

<sup>20</sup> See especially Pereboom (2001).

<sup>21</sup> See also Sartorio (2016: 167-8), who similarly speculates that “a psychological ‘dilution of responsibility’ effect” may be behind “the appearance that responsibility is diluted or absent in certain cases where it's really not”.

<sup>22</sup> Though see Bazargan (2014) for arguments to the effect that B may indeed still be liable in such cases (basically, because it doesn't follow from that fact that B is required to kill V on lesser-evil grounds that he doesn't *wrong* V in doing so).

<sup>23</sup> See A. Kaiserman, *supra* note 9.

reasons-sensitive B was in acting the way he did. Again, though, this is all perfectly compatible with Individualism. It's not *because* A is more responsible that B is less responsible; rather, B's diminished responsibility is grounded in the fact that his action was only partly motivated by reasons to act (that it was an action of A's which combined with those reasons to bring about B's action is irrelevant).

It's often not easy to tell which of these categories a particular case belongs to. Suppose for example that A threatens B with some misfortune unless he kills V, which B then does. Was B's act entirely motivated by a reason for him to kill V, namely that it would allow him to avert the threat from A? Or was it simply that A's threat so overwhelmed B that he killed V effectively on impulse, without any sensitivity to the reasons for or against this course of action at all? Or was B's action in fact brought about by some combination of both these forces (and if so, how much did each contribute)? How we answer these questions will make a big difference to how responsible we should judge B to be for V's death, on the view sketched above, though we may well lack the evidence to properly adjudicate between these different possibilities.<sup>24</sup>

## 5. Groups and Collective Responsibility

In all the cases we have considered so far, A and B act *independently* in bringing about the outcome – they act for their own reasons, without any shared intention. But of course not all cases are like this. Many of the cases Zimmerman considers are ones in which individuals act as a *group*, each intending to play their part in realising a common purpose. Some philosophers have argued that the only way to capture the moral significance of such cases is by postulating the existence of group *agents* – agents constituted by individuals but whose beliefs, intentions and actions are in some important sense 'over and above' the beliefs, intentions and actions of their members. Let's suppose, for the sake of argument, that group agents exist and can be held responsible for the consequences of their actions just as individuals can.<sup>25</sup> This raises the following question: might an agent A be less responsible for some outcome *o* in virtue of being a member of a group agent B which is also responsible for *o*?

Start by considering the following pair of cases:

*Terrorist*: A, who is a member of a terrorist organisation B, kills V, intending in doing so to further the aims of B. A's act was not directed, encouraged, or otherwise caused by any of the other members of B.

*Lone Wolf*: A kills V, intending in doing so to further the aims of a terrorist organisation, B. But unbeknownst to A, his membership of B has recently been revoked. All else is the same as in *Terrorist*.

I take it that B is not responsible for V's death in *Lone Wolf* (assuming, again, that B played no causal role in bringing about A's decision to kill V, for example by radicalising him while he was a member). On some views, though, B *would* be responsible for V's death in *Terrorist*, because in that case A's action *constitutes* an action of B's.<sup>26</sup> But even if there is such a difference in B's responsibility

<sup>24</sup> On this point, see also Frankfurt (1973).

<sup>25</sup> For arguments to this effect see Pettit (2007).

<sup>26</sup> There are tricky questions here about what it is for an agent to act *qua* member of a group, in such a way that his action constitutes an action of the group. Clearly not every action of a member of a group constitutes an action of that group – Ronaldo's brushing his teeth in the morning doesn't constitute an action of Real Madrid's, for example. Nevertheless, it seems plausible that an individual's action can constitute an action of a group of which they are a member, even if that individual's action wasn't directed, encouraged, or otherwise caused by any other member of the group. Consider for example what Lawford-Smith (2019: 152) calls 'Latitude-Collective' behaviour, where a

between the two cases, it's not clear how this difference could make a difference to *A*'s degree of responsibility for *V*'s death. *A* acts with the same intention, in the same way, in both cases – that his membership of an organisation happens to have been revoked in one case but not the other seems to have little bearing on the nature of the connection between *A* and *V*'s death.

Despite this, however, some authors have defended theories of group responsibility which appear to imply that there could be a difference in *A*'s degree of responsibility for *V*'s death between the two cases. Advocates of such views are principally interested in alleged cases of 'responsibility voids', where a group is responsible for some outcome without any of its members being responsible for it (at least not in virtue of their actions *qua* individuals). In such cases, it is argued, there is a need to ensure that the group's responsibility for the outcome somehow 'shows up' at the individual level. Hence it is proposed that when a group is responsible for some outcome, the group's responsibility 'distributes' among its members. Through such a process of distribution, "collective responsibility changes form and becomes something closer to personal responsibility, albeit personal responsibility that exists only because one's collective is responsible for harm" (Smiley 2017).

It's not entirely clear whether this distributed share of the group's responsibility is supposed to be *added* to any individual responsibility for the outcome the members already have, or to *replace* it. In cases of responsibility voids, of course, these amount to the same thing, which is probably why the question has not received much attention. If we take the latter option, though, *A* may well end up being less responsible for *V*'s death in *Terrorist* than in *Lone Wolf*, since the share of *B*'s responsibility *A* receives may well be smaller than the degree of responsibility he has for it in virtue of his actions *qua* individual.

It will come as no surprise by now, however, that I reject the idea that a group's responsibility for an outcome 'distributes' among its members, insofar as this assumes that responsibility is some fixed quantity that can be transferred from one agent to another, rather than something entirely grounded in the relationship between an agent and an outcome. Moreover, nothing like this view is actually needed to meet the challenge posed by responsibility voids. For one thing, it's not at all clear that there really are any *genuine* cases of responsibility voids, by which I mean cases where a group is responsible for an outcome without any individual *ever* having been responsible for it.<sup>27</sup> The kinds of cases that are typically offered are ones in which a harm occurs partly as a result of procedural and structural failures within a group, which themselves suffice to ensure that no individual member was in a position to foresee or control whether the harm would occur. Consider Pettit's (2007) example of the *Herald of Free Enterprise*, a ferry which sank with the loss of nearly two hundred lives after leaving Dover with its bow doors open. The subsequent enquiry identified multiple failings in monitoring, design, and safety procedures, as well as a string of similar incidents involving doors being left open in the run-up to the disaster.<sup>28</sup> After the coroner returned a verdict of unlawful killing, seven senior employees of the operating company were charged with

---

subordinate encounters unforeseen difficulties requiring quick decisions to be made, and as a result exercises latitude by acting against the explicit commands of her superiors, but still in what she takes to be the best interests of the group, and in a way her superiors retrospectively approve of – Lawford-Smith suggests, and I agree, that the subordinate's action would still constitute an action of the group in such circumstances. In any case, I'll assume for the sake of argument that one can reasonably fill in the details of *Terrorist* in such a way that *A*'s action does constitute an action of *B*'s.

<sup>27</sup> Of course there may be cases where a group is responsible for an outcome without any of its *current* members being responsible for it, because all the individuals involved in bringing it about have since left the group. But these cases don't seem problematic in the same way.

<sup>28</sup> *Mv Herald of Free Enterprise*, Report of Court No. 8074 Formal Investigation, URL: <[www.gov.uk/maib-reports/flooding-and-subsequent-capsize-of-ro-ro-passenger-ferry-herald-of-free-enterprise-off-the-port-of-zeebrugge-belgium-with-loss-of-193-lives](http://www.gov.uk/maib-reports/flooding-and-subsequent-capsize-of-ro-ro-passenger-ferry-herald-of-free-enterprise-off-the-port-of-zeebrugge-belgium-with-loss-of-193-lives)>

manslaughter. Yet the cases against the defendants were ultimately thrown out by the judge in the High Court. Ironically, the company's shortcomings appear to have helped their defence: the case turned on whether it was obvious to the defendants that the ferry might set sail with its doors open, but given the poor internal communications in the company it couldn't be proven that the defendants were aware of the previous incidents of this having occurred.

In many such cases, however, it may be possible to find individual members responsible for the outcome via a version of the 'tracing strategy'.<sup>29</sup> We hold drunk drivers responsible for the harms they cause, even though they had no control over their actions at the time, because the harms were also caused by earlier actions of theirs over which they *did* have control (namely, their drinking). If a doctor fails to turn up to a mandatory training session on the dangers of a particular drug, then she is responsible for the harms caused by administering it, even though she couldn't have foreseen at the time that doing so would have those consequences, because the harms were also caused by her earlier failure to turn up to the training session, and it *was* reasonably foreseeable that not turning up to a training session might lead to patients being harmed. Similarly, one might argue, members of a structurally deficient group may bear responsibility for the consequences of a group action, even if as a result of those deficiencies they weren't in a position to foresee or control those consequences at the time, in virtue of their earlier failures to put in place measures which would have resolved those deficiencies or prevented them from developing. If the law is unable to reflect this, that's a problem for the law, but not for our practice of attributing responsibility.

Even if there are genuine cases of responsibility voids, though, it's not clear why a group's responsibility would need to 'show up' at the individual level. It's true, of course, that it's not typically possible to punish a group, or force it to pay compensation, without imposing costs on at least some of its members. But this phenomenon is not unique to cases of group responsibility.<sup>30</sup> It's true in the individual case too that punishing an agent imposes costs on other people, such as the agent's family or friends, who bear no responsibility for the outcome at all; yet this is not usually considered a moral barrier to just punishment. So long as the *object* of punishment is the group itself, and any individuals are harmed only as a *side-effect*, then punishing a group in cases of responsibility voids needn't violate the moral prohibition on punishing people for wrongs for which they are not responsible. Of course, there may be better or worse ways of distributing the costs to individuals of punishing a group. But first, we're *not* distributing the group's responsibility itself when we do this; and second, the principles underlying such a distributive exercise are just the principles underlying the just distribution of costs more generally – there is nothing morally exceptional about cases where the agent being punished happens to be a group.

## 6. Conclusion

In this paper I have sought to articulate and defend a form of individualism about responsibility, according to which how responsible an agent is for some outcome is fully grounded in facts about the agent, the outcome and the relations between them. I take this to be the fundamental insight of Zimmerman's original paper. But I've also sought to go beyond his discussion by considering different ways in which two agents might both bear responsibility for the same outcome, and drawing attention to certain important distinctions – overdetermination vs. joint causation, causes vs. background conditions, coercion vs. direct manipulation – which I think are conflated in Zimmerman's account.

Though I haven't addressed the topic explicitly, I think there are nevertheless some important lessons that can be drawn from the preceding arguments for those working on the ethics of war.

---

<sup>29</sup> See Vargas (2005) and Fischer and Tognazzini (2009) for two different perspectives on the tracing strategy.

<sup>30</sup> On this point, see Pasternak (2011) and Lawford-Smith (2018).

In particular, they should make one suspicious of any attempts to infer facts about one agent's degree of responsibility from facts about another's degree of responsibility. It's tempting to conclude from the fact that one individual or group is *wholly* or *fully* responsible for some outcome that they are thereby *solely* responsible for it; indeed, describing someone else as wholly responsible for some outcome is often simply interpreted as a way of denying any responsibility for it oneself. Yet this, as we've seen, is a mistake – two agents A and B can *both* be fully responsible for some outcome *o*, if for example A's and B's actions overdetermined *o*, or if A's action caused B's causing of *o*, or if A's action constitutes a group action of B's.

Similarly, our discussion should make us suspicious of talk of states or individuals 'sharing' responsibility, or of responsibility being 'distributed', 'apportioned', 'transferred' or 'shifted' between them. Such talk may be more appropriate in some cases than in others – for example, talk of A and B 'sharing' responsibility might be more appropriate in cases of joint causation, where their actions contribute to a single causing of an outcome, than in cases of overdetermination, where the outcome is caused by both actions individually. But it would be better to avoid such talk entirely, insofar as it suggests that responsibility for an outcome is some sort of quantifiable and divisible substance, rather than a property of agents grounded entirely in the relationship between them and the outcome. Responsibility, in other words, is not a pie, and we shouldn't talk about it as if it is.

**Acknowledgements:** I'm indebted to the organisers and participants of the second *Conversations on War* project in Montenegro for valuable comments on a distant ancestor of this paper.

## References

- Arnold, D. G. (2001). Coercion and moral responsibility. *American Philosophical Quarterly*, 38(1), 53-67.
- Bazargan, S. (2014). Moral coercion. *Philosophers' Imprint*, 14(11), 1-18.
- Bernstein, S. (2017). Causal proportions and moral responsibility. In D. Shoemaker (ed.), *Oxford studies in agency and responsibility* (vol. 4), (pp. 165-182). Oxford: OUP.
- Chockler, H. & Halpern, J. Y. (2004). Responsibility and blame: A structural-model approach. *Journal of A.I. Research*, 20, 93-115.
- Fischer, J. M. & Tognazzini, N. A. (2009). The truth about tracing. *Noûs*, 43(3), 531-556.
- Frankfurt, H. (1973). Coercion and moral responsibility. In T. Honderich (ed.), *Essays on freedom or action* (pp. 63-86). New York: Routledge.
- Frowe, H. (2014). *Defensive killing*. Oxford: OUP.
- Halpern, J. Y. (2015). Cause, responsibility and blame: A structural-model approach. *Law, Probability and Risk*, 14(2), 91-118.
- Halpern, J. Y. (2016). *Actual causality*. Cambridge, MA: The MIT Press.
- Hart, H. L. A. and Honoré, T. (1985). *Causation in the Law* (2<sup>nd</sup> ed.). Oxford: Clarendon Press.
- Kaiserman, A. (2016). Causal contribution. *Proceedings of the Aristotelian Society*, 116(3), 387-394.
- Kaiserman, A. (2017a). Necessary connections in context. *Erkenntnis*, 82(1), 45-64.
- Kaiserman, A. (2017b). Partial liability. *Legal Theory*, 23(1), 1-26.

- Kamm, F. M. (2007). *Intricate ethics: Rights, responsibilities, and permissible harm*. New York: OUP.
- Lawford-Smith, H. (2018). What's wrong with collective punishment?. *Proceedings of the Aristotelian Society*, 118(3), 327-345.
- Lawford-Smith, H. (2019). *Not in their name: Are citizens culpable for their states' actions?*. Oxford: OUP.
- Lewis, D. (1973). Causation. *Journal of Philosophy*, 70(17), 556-567.
- Mackie, J. L. (1955). Responsibility and language. *Australasian Journal of Philosophy*, 33(3), 143-159.
- Mackie, J. L. (1974). *The cement of the universe*. Oxford: Clarendon Press.
- McMahan, J. (2011). Who is morally liable to be killed in war. *Analysis*, 71(3), 544-559.
- Mill, J. S. ([1843] 1868). *A system of logic, ratiocinative and inductive (vol. 1)*. London: Longmans, Green, Reader, and Dyer.
- Pasternak, A. (2011). The distributive effect of collective punishment. In T. Isaacs and R. Vernon (eds.), *Accountability for collective wrongdoing* (pp. 210-230). Cambridge: CUP.
- Pereboom, D. (2001). *Living without free will*. Cambridge: CUP.
- Pereboom, D. (2014). *Free will, agency, and meaning in life*. Oxford: OUP.
- Pettit, P. (2007). Responsibility incorporated. *Ethics*, 117(2), 171-201.
- Robichaud, P. & Wieland, J. W. (eds.) (2017). *Responsibility: The epistemic condition*. Oxford: OUP.
- Sartorio, C. (2015). A new form of moral luck?. In A. Buckareff, C. Moya & S. Rosell (eds.), *Agency, freedom, and moral responsibility*. New York: Palgrave-Macmillan.
- Sartorio, C. (2016). *Causation and free will*. Oxford: OUP.
- Shoemaker, D. (2011). Attributability, answerability, and accountability: Toward a wider theory of moral responsibility. *Ethics*, 121(3), 602-632.
- Smiley, M. (2017). Collective responsibility. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Summer 2017 Edition)*, URL = <https://plato.stanford.edu/archives/sum2017/entries/collective-responsibility/>.
- Tadros, V. (2016). Causation, culpability, and liability. In C. Coons & M. Weber (eds.), *The ethics of self-defense* (pp. 110-130). Oxford: OUP.
- Vargas, M. (2005). The trouble with tracing. *Midwest Studies in Philosophy*, 29(1), 269-291.
- Watson, G. (1996). Two faces of responsibility. *Philosophical Topics*, 24(2), 227-248.
- Weinryb, E. (1980). Omissions and responsibility. *Philosophical Quarterly*, 30(118), 1-18.
- Zimmerman, M. J. (1985a). Sharing responsibility. *American Philosophical Quarterly*, 22(2), 115-122.
- Zimmerman, M. J. (1985b). Intervening agents and moral responsibility. *Philosophical Quarterly*, 35(141), 347-358.