

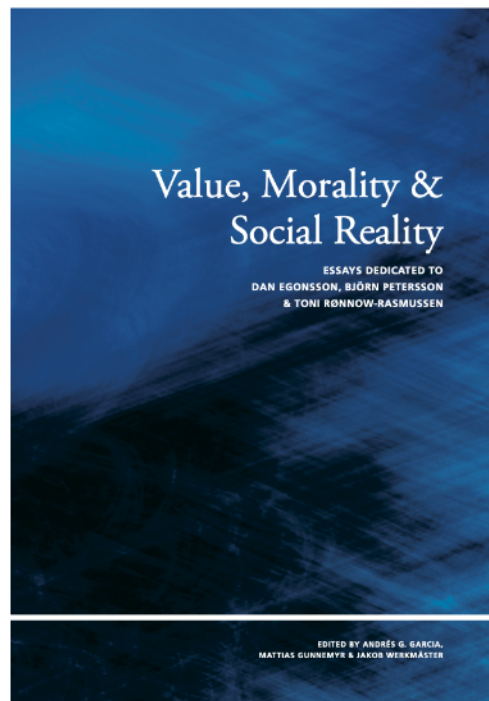
# Causal Involvement, Collectives, and Blame

## Replies to Petersson

*Matthew Talbert*

In: Garcia, A., Gunnemyr, M. & Werkmäster, J. (2023) *Value, Morality & Social Reality: Essays dedicated to Dan Egonsson, Björn Petersson & Toni Rønnow-Rasmussen*. Lund: Department of Philosophy, Lund University. DOI: <https://doi.org/10.37852/oblu.189>

ISBN: 978-91-89415-65-2 (print), 978-91-89415-66-9 (digital)



Published by the Department of Philosophy, Lund University.  
Edited by: Andrés García, Mattias Gunnemyr, and Jakob Werkmäster  
Cover image by Fabian Jones. Cover layout by Gunilla Albertén.

DOI: <https://doi.org/10.37852/oblu.189.c537>



This text is licensed under a Creative Commons Attribution-NonCommercial license. This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license does not allow for commercial use.

(License: <http://creativecommons.org/licenses/by-nc/4.0/>)

Text © Andrés García, Mattias Gunnemyr, and Jakob Werkmäster 2023.  
Copyright of individual chapters is maintained by the chapters' authors.

# Causal Involvement, Collectives, and Blame

## Replies to Petersson

Matthew Talbert

**Abstract.** This paper argues that there is reason to distinguish between moral responsibility and moral blameworthiness and, in particular, that we can acknowledge that a person is responsible for the negative outcomes of their behavior without this informing our judgments about the person's blameworthiness. This theme is elaborated in the context of a discussion of some of Björn Petersson's work on collective responsibility.

## 1. Introduction

For some time, I have been interested in how to think about responsibility and blameworthiness for the causal outcomes of behavior. In this paper, I consider how my own thinking on this topic interacts with the views that Björn Petersson has defended.<sup>1</sup> I begin with a sketch of my own (rather inchoate) thoughts on the topic, then I turn to consider aspects of Petersson's work on collective responsibility and their connection with the issues in which I am interested.

---

<sup>1</sup> I am delighted to contribute to this collection honoring the careers of Dan Egonsson, Björn Petersson, and Toni Rønnow-Rasmussen. Though my paper deals only with Björn's work, I would like to express my gratitude to Toni and Dan, along with Björn, for the hospitality and kindness that they showed me during my years at Lund University.

## 2. Outcomes, Responsibility, and Moral Blameworthiness

It is generally assumed by those who write about moral responsibility that we can be morally responsible not only for our actions but also for the consequences of our actions. Moral responsibility theorists also tend to assume a close fit—on the negative side of the moral ledger—between moral responsibility and moral blameworthiness. So, if you are morally responsible for some harmful consequence of a wrongful action, then you are also blameworthy on account of the harm that you caused. And, for those who are comfortable with talk of degrees of moral responsibility, if a person bears comparatively greater responsibility for some outcome, then she is, to that extent, more blameworthy on account of that outcome.

Contrary to the above perspective, I am going to argue that moral responsibility and moral blameworthiness come apart in certain ways and specifically with respect to the consequences of actions. I'll sketch the general lines of my thinking in this section and fill in some additional details below in my responses to Petersson.

In my view, judgments about blameworthiness are properly independent of causal judgments. This is because: (i) judgments about blameworthiness are fundamentally judgments about the fittingness of the morally offended responses involved in moral blame; and (ii) the obtaining of a causal relation, or the occurrence of an unwelcome causal outcome, is never in itself morally offensive in a way that makes the responses involved in blame appropriate. It is only when an outcome is caused in a certain way—e.g., by independently morally objectionable intentions and motives—that moral offensiveness arises. Indeed, I am inclined to think that whatever genuine moral offensiveness arises in a given context is entirely accounted for by the moral quality of such things as an agent's intentions and motives, and not at all by the fact that these things happen to bear a causal relation to an unwelcome result.<sup>2</sup>

However, judgments about moral responsibility may well be dependent on causal judgments. Plausibly, a person is morally responsible for something, such as a causal outcome of her behavior, if she fulfills certain control conditions and epistemic conditions with respect to that outcome. And, particularly when it comes to outcomes, control is at least partly a causal or explanatory notion: an agent's having the relevant control over an outcome is partly a function of that agent being causally or explanatorily connected to the outcome. Of course, causing an outcome is not enough to be morally responsible for it, so the control condition is supplemented by, or it includes, an epistemic condition: one must have known or suspected—or one should have known or suspected—that a certain consequence was a likely result of one's action.

---

<sup>2</sup> There's an obvious connection between what I say in this paper and the usual things that people say when they argue against resultant moral luck. For arguments related to those in this paper, but with more discussion of moral luck, see Enoch & Marmor (2007), Graham (2017), Khoury (2018), and Zimmerman (2002). For my own views on moral luck, see Talbert (2019).

I conclude from the above that, to the degree that judgments about moral responsibility are dependent on causal judgments, and judgments about blameworthiness are not so dependent, judgments about responsibility and blameworthiness can come apart in ways that are not standardly recognized. It seems to me, in particular, that the fact that a person is morally responsible for an outcome may not tell us anything about the degree to which she is open to moral blame.

Instead of separating blameworthiness from moral responsibility, why not simply say that we are not morally responsible for the consequences of our actions, as Andrew Khoury (2018) suggests? I think this is too revisionary a use of “moral responsibility.” Thus, I grant that we can be morally responsible for the consequences of our actions insofar as we can intentionally or knowingly (or negligently or recklessly) bear the right causal or explanatory relation to these consequences. The responsibility here goes beyond mere causal responsibility: in virtue of satisfying relevant causal and epistemic conditions, an outcome can be attributed to an agent as an exercise of her powers of agency, and the outcome may be said to be the agent’s doing in a way that seems conceptually distinct from an attribution of causal responsibility. But as I’ve said, I think that moral responsibility for consequences need not affect blameworthiness—it need not affect the aptness of directing towards an agent the negative emotional responses involved in moral blame.

I admit, however, that it also seems an undue revision to normal speech to deny that people can be “to blame” for outcomes. I suggest that to say that someone is to blame for an outcome means no more than that that person is morally responsible for an unwelcome outcome, and that she is responsible for this outcome in virtue of some morally objectionable feature of her self that bears the right causal/explanatory relation to the outcome. These morally objectionable features might be such things as the agent’s morally bad motives and intentions, her bad desires and patterns of concern, and so on. Such things are not required for moral responsibility per se—one can be morally responsible for an outcome in virtue of praiseworthy motivations—but if her motives, patterns of concern, etc. hadn’t been morally deficient, then she wouldn’t be *to blame* for that outcome.

An agent who is to blame for an outcome is blameworthy in virtue of the factors (e.g., her bad intentions) that make it correct to say that she is *to blame* for the outcome. But these morally objectionable features of the agent *on their own* (and even if they hadn’t led to a bad outcome) are enough for her blameworthiness. So, an agent is open to blame, and so blameworthy, regardless of whether her objectionable intentions and motives are causally connected to an unwelcome outcome for which the agent is morally responsible and to blame.

Moreover, the fact that a person is morally responsible (and to blame) for an unwelcome outcome does not make her *more* blameworthy than she already is just in virtue of possessing the objectionable features that, together with the obtaining of certain causal connections, make her to blame for the outcome. This is because,

as I suggested above, the obtaining of the causal connections necessary for establishing the *morally responsible for* and *to blame for* relations is not in itself morally offensive. The moral offense is already there in the objectionable features of the agent that may or may not give rise to the unwelcome outcome. Since the actual occurrence of the outcome, and the obtaining of the *to blame for* relation between agent and outcome, does not affect the moral quality of the objectionable features of the agent, it does not amplify blameworthiness (though see Lang 2021, Chapter 2, for the contrary view). Of course, as others have noted (e.g., Enoch & Marmor, 2007; Lawson, 2013), the occurrence of a bad outcome may make an actor's objectionable motives (etc.) more conspicuous than they would otherwise be, but this explains only our readiness to blame in the face of such outcomes, and not the blameworthiness of those whom we blame.

### 3. Replies to Petersson

I turn now to consider some of Björn Petersson's reflections on collective agency and responsibility for outcomes and how the view outlined in the previous section might interact with Petersson's. I begin with Petersson's 2008 paper, "Collective Omissions and Responsibility" (which builds on Petersson, 2007). The title makes Petersson's focus on omissions clear, but I take my comments to apply to both actions and omissions.

Petersson is a realist about collective responsibility: "groups as such can be morally responsible for effects of their acts, in a sense that cannot be reduced to judgments about individual members' acts" (2008, 244). Collective responsibility, then, is not merely an aggregation of, nor does it collapse into, instances of individual responsibility. Moreover, for Petersson, while "[y]ou are individually responsible ... for your intentional marginal contribution to some harm" your "individual responsibility need not coincide with your co-responsibility for the overall harm produced by the collectively responsible group" (2008, 251). This last point is easiest to see in cases in which the collectively caused "harm is overdetermined, so that no individual member's act makes any difference to the occurrence of the event in question" (2008, 251).

As far as moral responsibility goes, I see no reason to disagree with what Petersson says. I may be part of a group, and it may be true that the group has caused some harm and that, as an appropriately situated member of this group, I share in its responsibility for doing so. As a member of the group, the harm it caused is partly attributable to me, but my responsibility need not be proportional to my difference-making contribution to the harm. For example, in overdetermination cases, my individual action may make no difference to the occurrence of a harm: had I not acted, the overdetermined harm would have been the same. So, we have at least one

way in which collective responsibility does not collapse into individual responsibility.

But I have trouble seeing how related conclusions about blameworthiness might follow. Blameworthiness seems to me individualized in a way that responsibility may not be. Unfortunately, it is difficult to make this case in the context of the current debate about responsibility since, as I noted above, most theorists assume a tight fit between moral responsibility and blameworthiness. For example, Petersson says that he is working with a “thick” sense of moral responsibility: “[t]o hold a collective morally responsible (in the thick sense) for some harm is to imply that moral sanctions are in place” (2008, 251). It is clear from context that the sanctions Petersson has in mind include moral blame. Indeed, once it is clear that we are considering a case of moral wrongdoing, it tends to make little difference—for Petersson and most other responsibility theorists—whether we speak of “moral responsibility” or of “blameworthiness.” This easy transition from *responsibility* to *blameworthiness* is one of the things I am arguing against.

Now, it might be thought that Petersson and I are simply using “moral responsibility” differently, but I don’t think this recognizes the disagreement between us. I mean to argue that the conditions that Petersson—rightly, in my view—claims are sufficient for group responsibility do not yield straightforward conclusions about blameworthiness in the way that I take Petersson to suppose.

### **3.1 Loosely Structured Groups and Individualized Blame**

The first point I want to make is that sometimes “blaming a collective” can be suitably analyzed in terms of blaming individuals even if a corresponding analysis of moral responsibility would not be appropriate. This observation is most applicable in cases of “loosely structured” collectives (May, 1990), which are Petersson’s central focus in “Collective Omissions and Responsibility.” Such loosely structured entities “need not have a common decision procedure, let alone be formally constituted as a group” (Petersson, 2008, 246).

Here is an example (from Petersson, but based on one in Held, 1970) of a loosely structured group and a case in which their collective omission seems blameworthy.

Suppose an injured person was trapped under a girder, and that the joint effort of two people would be needed to lift it. You and I were the only ones who knew about this accident, and we could have helped if we had acted together.... [but] no collective effort took place. (Petersson, 2008, 257)

Suppose the person trapped under the girder dies. In this case, our loosely structured group is plausibly responsible for a death, and we are plausibly open to blame (depending on what explains our failure to cooperate). But what does blame come to in the case of such a loosely structured collective? A more concretely structured group might be clearly targeted *as a collective entity*: a sports team can be disbanded

or made to forfeit a game, a manufacturing company can be fined, and so on. But with loosely structured entities, it is harder to say what blaming a group comes to—or at least it is harder to see how blaming the group comes to anything more than blaming individuals. In addition, the moral significance of blame (its point and aim) may be exhausted once *individuals* have felt blame’s sting, have repented, and so on. After individuals have been reached in this way, what is left over, in the loosely structured case, for “blaming the group” to do?

As Petersson notes, “[w]hen we blame someone, we want the wrongdoer to care about our negative attitudes towards him, as well as about our reasons for having this attitude...” (2008, 249). Again, a concretely structured group—the sports team or the commercial enterprise—might issue a corporate apology that registers the group’s acceptance of blame.<sup>3</sup> But how does our desire for such a response apply to a loosely structured group? I suspect that our concern will have been met when the individuals composing such a group—or perhaps a specific subset of these individuals—have taken up an appropriately repentant stance. This stance *might* be expressed in a thought that has, so to speak, collective content. Each individual in Petersson’s example might think: “*we* acted wrongly, *we* should have done something.”<sup>4</sup> But would anything be missing, would our blame not have achieved its aim, if each individual, or at least those who do have something to apologize for, merely thought: “*I* should have done something different”?<sup>5</sup>

Petersson considers a few variations of the case of the person trapped under the girder. First, it might be that you and I are faultlessly unable to communicate and to coordinate our efforts. Second, one or both of us might be a committed non-cooperator, which explains why our efforts were not effectively coordinated.<sup>6</sup> Finally, it might be the case “that we both considered which options the group had as one unit of causal agency, and that we both agreed” on what was needed to save

---

<sup>3</sup> Of course, we are liable to think that something is left undone by such an apology if we are uncertain whether members—perhaps specific members—of the group feel personal moral guilt.

<sup>4</sup> Olle Blomberg observes in written comments that Petersson “might prefer to say that each has a thought from the ‘we-perspective,’ where this is distinct from having a thought with ‘merely’ collective content. ... [the collective perspective would be] part of the thought’s mode rather than its content.” Blomberg directs readers to Petersson (2015; 2017)

<sup>5</sup> Perhaps this thought won’t be apt in every case. As Mattias Gunnemyr points out in written comments, the case might be one in which I am stuck with a non-cooperator (and that is why we do not act) or the case might be such that if I act alone (and others fail to join in), then I will simply make things worse. In such cases, Gunnemyr suggests that the appropriate thought might be, “*we* should have done something different.” I’m inclined to think, though, that the thought might be, “*he* (or *they*) should have done something different.”

<sup>6</sup> In comments, Blomberg notes that if both of us are non-cooperators, then something would be missing if each of us thought only, “I should have done something different” since individual action would have changed nothing; so, the relevant thought must be, “*we* should have done something different.” See Blomberg and Petersson (2023).

the trapped person, but we could not agree on how to time our effort, and “[w]hile we argued ... the victim died” (Petersson, 2008, 258).

Here is how Petersson assesses responsibility and blameworthiness in these cases (note the easy shift between talk of “blame” and of “moral responsibility”):

Firstly, we would not blame people for not acting together if they were unable to form joint attitudes to begin with. Secondly, if the attitudes of individual members of a group hinder joint efforts, these individuals are the ones to be blamed for the effects of not acting jointly. Finally, in the third type of case, where a harm is caused by a failure to act that can be explained by members’ attitudes concerning a group’s options regarding that harm, we find it less inappropriate [to] hold the group as such morally responsible. (2008, 259)

In the third case, what does holding the group responsible—blaming the group—come to? I suspect that the blame applied in the third case will be quite similar to the individual blame applied in the second case. That is, we will blame individuals in both cases, and the blame might take a similar form in both cases. And this will be so even if there is a genuine difference in moral responsibility between the two cases: individual responsibility in the second case, and group responsibility in the third.<sup>7</sup>

Petersson is aware, of course, that “collective sanctions inevitably strike individual members” of a group (2008, 251); I am suggesting that, in some cases, there may not be anything to blaming a group over and above individually felt sanctions. And this indicates a sort of disanalogy with collective moral responsibility. One can have the thought about blame that I have described while still supposing that groups cause things and are morally responsible for outcomes in a way that it is not reducible to individual causal contributions. It may be that collective responsibility does not collapse into individual responsibility, but I suspect that collective blame—or the point of blaming a collective—can collapse in this way, at least in cases of loosely structured groups.

### **3.2 Groups and “Moral Taint”**

I turn now to individual blameworthiness and its relationship to group responsibility.

---

<sup>7</sup> In comments, Gunnemyr suggests an alternative that I admit has some appeal for me, though I don’t have space to develop the idea here. The basic thought is that, in the third case, the members of the group are plausibly engaged in (objectionable) “we-mode thinking” as they fruitlessly (and callously) debate what to do. So, perhaps I could say that it is solely these thoughts—undertaken from a collective perspective—and the objectionable (and, in some sense, collective) quality of will that the thoughts manifest, which grounds blame. This preserves my resistance to allowing outcomes to affect blameworthiness but would allow for blameworthiness to be more inherently collective, in the relevant cases, than I have suggested.



Petersson says that “our method for delimiting the collective agent that is morally responsible for a specific harm should be such that it picks individuals that justifiably can be blamed for what the group has done” (2008, 251). Thus, we should establish “a link between the individual and the group’s act” to avoid holding responsible “innocent bystanders who may have been causally involved [in the production of harm] through no fault of their own” (2008, 251).

This is correct, as far as it goes, but in turning to collectively caused, overdetermined harm Petersson says:

In such cases, we may not be able to assign ordinary individual moral responsibility to any member, while we still find the group’s behavior reprehensible. To hold you co-responsible, then, is to hold you to account for the group’s act in virtue of the features that [make] you a member of the collective agent, regardless of your individual intentional marginal contribution. The direction of the relation between collective responsibility and co-responsibility is supposed to be top-down—members will be morally tainted by the worth of the collective action. (2008, 252)

Again, I find this a plausible thing to say about responsibility. Depending on our views about causation, we may have difficulty assigning individual moral responsibility in overdetermination cases because it may be unclear whether certain causal links hold between an individual’s choices and a harmful outcome. Yet it may still make sense to say that the group of which the individual is a member caused the harm and is responsible for it, and so it may make sense to assign moral responsibility to the individual in virtue of his membership in the group (where this membership is partly established by locating the relevant links “between the individual and the group’s act” that Petersson mentions).

But what should we say about blameworthiness? Here I am interested in Petersson’s reference to moral taint and to the reprehensible nature of the group’s behavior. Behavior, I take it, may be unwelcome, hurtful, and even wrong (though not, I think, wrong in a sense relevant for blameworthiness), independently of the motives and aims that explain the behavior. But I do not see that behavior, or the consequences of behavior, taken independently of facts about motives and aims, can count as morally reprehensible in a way that is relevant to concerns about moral taint and moral blameworthiness. So, I do not see how a group’s actions can be reprehensible in this way, and in a way that could plausibly reflect on the group’s individual members, except that these individuals’ motives, aims, and intentions are morally reprehensible. But in this case, moral taint, which I take to be the grounds for a judgment of blameworthiness, is not top-down; it is, rather, bottom-up.<sup>8</sup>

---

<sup>8</sup> In comments, Blomberg suggests that a set of micro-aggressions that are individually not particularly morally serious might aggregate in such a way that the target of these slights has grounds for significant moral offense directed at the group that has collectively slighted him. I must admit that this is an appealing proposal, and I am not sure that what I say in the text can accommodate it.

Perhaps moral responsibility can reasonably be conceived of as “top-down,” if this means that a person’s responsibility is not necessarily explained by their individual causal contribution to an outcome but rather may be explained by their membership in a group and by what the group has done. But blameworthiness, as far as I can see, goes the other way: an individual’s openness to blame is explained by facts about that individual, facts beyond their membership in a group. And each individual member of a group may show that they are not open to blame by showing that there is nothing reprehensible in their motives and intentions—importantly, this may not be the same thing as the individual showing that they are not morally responsible for a collectively caused outcome. In other words, it’s possible that an individual fulfills relevant causal, epistemic, and group membership conditions such that they count as co-responsible for a harmful outcome and yet their individual motives and intentions may show that they are not blameworthy. (Similarly, you can be morally responsible for a harm that you knowingly brought about, and yet you will tend to avoid blame if your intentions and motives are unobjectionable.)

Of course, it may be that, for certain groups with certain aims, there is no way that an individual can be a willing and informed member of that group without this indicating something reprehensible and blame-grounding about that person. But, again, the individual moral taint here will be a function of the fact that the individual willingly, and with relevant knowledge, joined a group of that sort. There is, I think, no top-level description of a group and its aims, no matter how distasteful, that suggests that individual members of the group are morally tainted without this taint being explicable in terms of reprehensible features of the individual agents.

### **3.3 Causal Influence and Blameworthiness**

In “Co-responsibility and Causal Involvement” (2013), Petersson responds to Christopher Kutz’s (2000) argument for rejecting causal involvement as a necessary condition on co-responsibility.<sup>9</sup> Kutz’s argument against the causal involvement condition focuses on cases of overdetermination, particularly the Allied bombing of Dresden at the end of WWII.

Given the number of aircraft and bomber crews involved in the raid on Dresden, the destruction of the city was overdetermined: “Each of the 8000 crewmen’s causal contribution was ... ‘marginal to the point of insignificance’” (Petersson, 2013, 848; quoting Kutz, 2000, 118). We might wonder, then, how such minor causal contributions can ground ascriptions of co-responsibility for the fact that Dresden was destroyed. How could such small contributions ground the relatively high degree of moral blame that we might think apt in the case of the intentional destruction of a city populated largely by civilians? It’s better, Kutz argues, to think of responsibility and blame in such cases as depending not on individual causal

---

<sup>9</sup> Again, Petersson is “talking about responsibility in a thick sense, in which moral responsibility is essentially connected to the justifiability of blame and other moral sanctions” (2013, 850).

contributions, but rather on the presence of morally objectionable intentions, such as the intention to destroy civilian-populated Dresden.

In response, Petersson notes that even if an action “made a small, negligible or imperceptible difference to the occurrence of a great harm, it made a difference to the occurrence of that harm” (2013, 858). As such, the case at hand does not threaten “the idea that being co-responsible for something requires making a difference to its occurrence” (2013, 858). “At most,” such cases “show that *degree* of co-responsibility for a specific event need not correspond to the size of the causal contribution” (2013, 858, emphasis in original).

The suggestion is that one’s degree of responsibility—which will, for Petersson, entail conclusions about blameworthiness—might outstrip one’s degree of causal contribution. This is a noteworthy feature of Petersson’s account for my purposes. For, as Petersson observes, “it might seem odd to insist that causal involvement is an essential ... condition for co-responsibility, while admitting that there is not always any straightforward ... relation between causal contribution to an event and degree of co-responsibility for that event” (2013, 858). In fact, I think there is something odd here, and I’ll try to draw it out.

Petersson’s first response to the oddness he mentions is to note that, on the causal-involvement account, even a small contribution to a horrible outcome may warrant significant blame. Even if a person is blamed in proportion to their causal contribution, “[w]hat she should be blamed for would still vary not only with her share of the total event but also with the value of that event. 1/8000th of an atrocity could be an atrocity” (Petersson, 2013, 858).

It is true that a harm that is small according to some scale of measurement may still be a morally significant harm. Still, I think the above suggests an implausible representation of our blaming practices. Blame and the feelings that express it are not generally divided up to neatly correspond to fine-grained judgements about causal contributions. These responses do not, for example, automatically realign to reflect revised judgments about relative causal contributions when we learn that there were a few more contributors to a harm than we had originally supposed.

More helpful is Petersson’s elaboration of the thought that “the causal involvement condition does not imply proportionality between blame and causal contribution” (2013, 858). In this context, Petersson notes that a “justification of an assignment of blame and responsibility will have to appeal to a variety of factors *in addition to* the claim that something bad has happened and that the recipient [of these assignments] was involved in it” (2013, 859; emphasis added).

In addition to attending to a participant’s causal involvement, we will also be sensitive when assigning blame “to the agent’s type of involvement, the agent’s mental capacities, beliefs and intentions,” and perhaps also to features of “the social context” in which the agent acted (Petersson 2013, 859). If we are thinking of a collective action, like the bombing of Dresden, it will be relevant to assessments of individual blameworthiness that “participation in a collective project signals a certain kind of commitment” to that project: “[s]uch considerations may explain

why we might consider you highly blameworthy even for a very small contribution to a collectively produced effect” (Petersson 2013, 859). As Petersson observes, “the relative weight of ... non-causal considerations typically becomes greater in cases of participation [in a collectively caused outcome] than in cases of single individual actions” (2013, 859).

My suspicion is that, at least in collective cases, the weight of these non-causal considerations can be so great as to swamp our interest in individual causal contributions, at least when it comes to blameworthiness. Petersson disagrees. He says that the above “admission [about the importance of non-causal factors] does not undermine the causal involvement condition for co-responsibility” since “what makes the [individual Dresden] bomber co-responsible for the event in the first place is that his act contributed to its occurrence” (2013, 859).

I agree that a causal contribution is required to establish an individual bomber’s partial moral responsibility for the destruction of Dresden. But I suggest that it is much less clear why establishing such a causal relation is necessary for a bomber to be *blameworthy*, particularly when we note how large a role noncausal factors play in our judgments about the appropriateness of blame. If, as Petersson says, we can “explain why we might consider you highly blameworthy even for a very small contribution to a collectively produced effect” by referring to your commitment to an objectionable group project (2013, 859), why should we not regard you as worthy of blame on account of that commitment even if you made *no* contribution to the collectively produced effect?

Suppose that a bomber pilot is deeply and objectionably committed to the goal of bringing about the fiery deaths of Dresden civilians, yet he fails to make even a small contribution to this outcome because the bombs that he drops fail to detonate or because the bomb bay doors of his aircraft malfunction. I can see why this might make a difference to the bomber’s causal contribution to an outcome and so also to his moral responsibility for that outcome, but I fail to see how this makes the morally offended attitudes involved in blame any less appropriate than they would be had the bomber been successful in achieving his aim.<sup>10</sup> And this is because, had the bomber been successful, his openness to morally offended responses seems to me to already have a secure footing just in virtue of the morally offensive commitments and attitudes that explain the bomber’s choices and actions.

---

<sup>10</sup> One could opt for a sufficiently fine-grained account of “the bombing of Dresden” such that if the ineffective bomber had been absent, then the fine-grained version of the event would not have occurred. On this account, our bomber’s presence (and his failed efforts) would play a role in bringing about (the fine-grained version of) “the bombing of Dresden.” One can establish a causal link in this way if one is sufficiently motivated and suitably flexible in their account of causation. But I don’t see how establishing a causal link of this sort makes it any clearer that our bomber is a candidate for blame—this seems no more clear than it already was in virtue of the bomber’s objectionable moral orientation, which guided his objectionable efforts. For Petersson’s take on this fine-grained approach, see (2013, 852-856).

Relatedly, and as I suggested in the previous section, if we had a bomber pilot who made a clear contribution to the destruction of Dresden, and yet we somehow became convinced that there was *nothing* independently objectionable in his motives and intentions, then, while we might assign him causal responsibility, and even moral responsibility if he satisfied certain epistemic conditions, we would have no grounds for finding him worthy of blame. In the absence of such things as independently criticizable motives and intentions, even the most significant causal contribution will not be enough for blameworthiness.<sup>11</sup>

Certainly, Petersson and I disagree about much of the foregoing. But occasionally, the disagreement seems to me less stark than it might initially appear. Consider the following passage, which seems to express straightforward disagreement with one of my central claims above:

If we have evidence of an agent being committed to contributing to an outcome along with others, but it is clear to us that the agent completely fails to contribute to that outcome, I would regard it as absurd to blame the agent *for* that outcome. (Petersson, 2013, 864; italics in original)

Again, this may seem straightforwardly at odds with the account I gave of the unsuccessful Dresden bomber. But note that Petersson speaks here of blame *for an outcome*. I agree that if a Dresden pilot made no causal contribution to the fact that Dresden was bombed, then it would be absurd to blame him, even partially, *for that outcome*. But I claim that the unsuccessful bomber is blameworthy—that is, he is open to the responses involved in blame—on account of his bad motives and intentions, and that he is blameworthy to the same extent that he would be had he been partly responsible for (and so, partly *to blame for*) the fact that Dresden was bombed.

Petersson seems willing to meet me at least part of the way here. Consider the following example of his, inspired by a television comedy:

the thoughtful mother of a blind young neo-Nazi regularly swaps the son's swastika-badges ... for completely innocent symbols, without his knowledge. Suppose this son joins a neo-Nazi-demonstration and that the demonstration to some minor extent is successful in creating conflict and violence. At the same time, unintentionally the blind son radically diminishes this harmful effect of the collective behaviour, just by being visible in the crowd with cute symbols on his clothes. (2013, 864)

Petersson says of the son:

---

<sup>11</sup> As I noted in the previous section, it may be difficult to see how one could fulfill relevant epistemic conditions, and engage in certain courses of conduct, without this evincing an independently criticizable moral orientation, but still, the attribution of such an orientation seems to be necessary for blameworthiness.

He fully shares the participatory intentions of his fellows, but he obstructs the fulfillment of those intentions. We should blame him for trying but surely we cannot blame him for the small raise in conflict and violence that is an effect of the demonstration. It turned out that his mother's well-intentioned deception successfully prevented him from making himself responsible for that sort of harm. (2013, 864)

Again, I agree. The son is not morally responsible for the outcome in question because he does not contribute to it, and so we cannot say that he is to blame for the outcome. But, as Petersson allows, the son is open to blame on account of *what he tried to do*, and presumably this has to do with the fact that his trying was explained by his sharing the objectionable motives and intentions of his (more causally effective) fellow neo-Nazis.

And suppose that the son had succeeded in helping to bring about the sort of unwelcome outcome at which he aimed. This would not make him more worthy of (or worthy of more) blame because it would not make his bad motives and bad intentions more morally offensive—more blame grounding—than they already were. Of course, if the son were to have caused some bad outcome, this might draw our attention to his objectionable aims and intentions and give us additional reason to regret their presence in him. In this way, the occurrence of the outcome might serve to explain why we blame without providing further moral grounds for blame.

## Conclusion

As I just observed, the occurrence of an unwelcome outcome can draw attention to blame-grounding motives and intentions. However, Petersson says that “[t]he idea that causal links are relevant merely as indicators of intentions gets things the wrong way around” (2013, 864; Petersson directs this comment toward Lawson 2013).

For Petersson, reference to causal involvement is essential to our responsibility practices. When we blame people, we aim, he says,

to make them react with corresponding feelings of guilt or remorse, not over past states of mind as such, but over what these states of mind have led to.... We want to make them realize that their choices had an impact—that it was no coincidence that bad things happened when they made those choices. Their choices *mattered*, not by themselves, but because of their connections to events that made the world worse. (Petersson 2013, 864; emphasis in original)

There's some tension, I think, between these claims and Petersson's observations about the blind neo-Nazi, which are immediately prior in the text. As Petersson suggested, we blame the neo-Nazi for what he tried to do, and we presumably do so on account of the intentions that moved him to so try, and independently of the fact that his attempt failed.

Still, there's something right in what Petersson says here. It is important that it is "no coincidence" that intentions lead to choices, that choices regularly lead to actions, and that actions regularly have effects in the world. If bad intentions and objectionable strivings never led to consequences that concern us, then I assume that we would not have our habitual concern about people's bad intentions. But because there is a fairly reliable connection between bad intentions and bad outcomes, we do care about people's intentions, and certainly about their attempts, independently of whether their bad intentions lead to bad consequences on a particular occasion. Indeed, and more generally, we expend a great deal of energy thinking about how we stand in other people's estimations, about what they *really* think about us, even if these inner orientations are not revealed in their actions.

The centrality of internal factors, such as intentions, for our responsibility practices is, I think, most prominent when we consider excuses. As I've suggested, no matter how unwelcome the consequences of someone's actions, if we are convinced that the action is explained by morally faultless motives and intentions, then we have no adequate grounds for blame since, by our own lights, there is no moral affront on which our morally offended blaming responses might reasonably be founded. And if an agent's action turns out to be harmless, and yet we become convinced that he had the most objectionable motives and intentions, then it is easy to understand why moral offense is aroused on the part of those who have luckily not been exposed to harm. I concede, again, that we would not have this interest in others' bare intentions and motives if these things did not regularly lead to happy and unhappy outcomes. But given this regular connection, our moral interest in mere intentions (and other internal states of agents) is perfectly intelligible even when these things do not give rise to external outcomes.<sup>12</sup>

## References

- Blomberg, Olle & Björn Petersson (2023) "Team reasoning and collective moral obligation". *Social theory and practice*. Advance online publication. <https://doi.org/10.5840/soctheorpract2023120177>
- Enoch, David & Andrei Marmor (2007) "The case against moral luck". *Law and philosophy*, 26(4): 405-36.
- Graham, Peter (2017) "The epistemic condition on moral blameworthiness: A theoretical epiphenomenon" in P. Robichaud & J. W. Wieland (Eds.) *Responsibility: The epistemic dimension* (163-79). Oxford, Oxford University Press.

---

<sup>12</sup> I presented some of the ideas in the second section of this paper at seminars at Lund University and Gothenburg University. I'd like to thank all the participants at these seminars for their thoughts. Particular gratitude is owed to Olle Blomberg, Gunnar Björnsson, Mattias Gunnemyr, Björn Petersson, Toni Rønnow-Rasmussen, András Szígeti, and Caroline Torpe Touborg. Blomberg and Gunnemyr are owed additional thanks for their generous and helpful written comments.

*Causal Involvement, Collectives, and Blame*

- Held, Virginia (1970) "Can a random collection of individuals be morally responsible?". *Journal of philosophy*, 67(14): 471-81.
- Khoury, Andrew (2018) "The objects of moral responsibility". *Philosophical studies*, 175(6): 1357-81.
- Kutz, Christopher (2000) *Complicity: Ethics and law for a collective age*. Cambridge: Cambridge University Press.
- Lang, Gerald (2021) *Strokes of luck: A study in moral and political philosophy*. Oxford: Oxford University Press.
- Lawson, Brian (2013) "Individual complicity in collective wrongdoing". *Ethical theory and moral practice*, 16(2): 227-43.
- May, Larry (1990) "Collective inaction and shared responsibility". *Nous*, 24(2): 269-77.
- Petersson, Björn (2007) "Collectivity and circularity". *Journal of philosophy*, 104(3): 138-56.
- Petersson, Björn (2008) "Collective omissions and responsibility". *Philosophical papers*, 37(2): 243-61.
- Petersson, Björn (2013) "Co-responsibility and causal involvement". *Philosophia*, 41(3): 847-66.
- Petersson, Björn (2015) "Bratman, Searle, and simplicity. A comment on Bratman, *Shared agency: A planning theory of acting together*". *Journal of social ontology*, 1(1): 27-37.
- Petersson, Björn (2017) "Team reasoning and collective intentionality". *Review of philosophy and psychology*, 8(2): 199-218.
- Talbert, Matthew (2019) "The Attributionist Approach to Moral Luck". *Midwest studies in philosophy*, 43(1): 24-41.
- Zimmerman, Michael (2002) "Taking luck seriously". *Journal of philosophy*, 99(11): 553-76.