Does collective responsibility have a point?

[Katie Steele, LSE]

Abstract: This paper considers whether there is reason to appeal to a special notion of collective responsibility in order to fill ‘voids’ in moral responsibility for harmful outcomes. While most authors focus on the (in)coherence of the proposed solution of holding groups responsible qua groups in the given situations, here the focus is rather on the (in)coherence of the motivating problem itself, i.e., the appeal to ‘responsibility voids’. It is argued that blameworthiness does not straightforwardly attach to outcomes; thus, while group contexts exacerbate potential disconnects between blameworthy acts and harmful outcomes, the so-called responsibility voids are by no means peculiar to groups and do not warrant a special notion of collective responsibility.

0 Introduction

This paper is underpinned by a pragmatic or functionalist approach to collective responsibility. The pertinent question here is: Do we get a better overall moral picture, i.e. one that better allows us to navigate the world, when we allow that groups, as well as individuals, can be held responsible in some sense for choices/acts/outcomes?

The great pragmatic promise of extending moral judgment to groups, quite simply, is that it may fill voids or gaps in the moral fabric that one would like to see filled. The voids or gaps in question are situations where a significant harm is caused by the actions of agents, but no one is clearly morally to blame or is otherwise accountable for the harm, or else various agents who we intuitively have no

1 There may be other reasons to appeal to group responsibility. The focus here is on responsibility voids, but one might otherwise be concerned about the opposite phenomenon—cases of individuals being burdened with too much responsibility (over-demandingness)—and consider whether the appeal to group moral responsibility might help with this purported problem. It is well beyond the scope of this paper to also consider the over-demandingness motivation for group responsibility. In any case, I consider the problem of responsibility voids to be more compelling, and also less specific to particular moral accounts.
think are morally implicated in the harm are let off the hook. Note that what is implicit here is that responsibility voids revolve around outcomes. As in, it is not choices or acts that are the main target in these discussions of responsibility voids, but rather outcomes. Or so I would contend. Outcomes clearly have great pragmatic importance—they tell whether people’s lives are going well for them.

The focus on outcomes gives rise to the following ambiguity: Do the worries expressed by various authors concern voids in blameworthiness, or rather voids in remedial responsibility (to use David Miller’s 2007 term) when it comes to a harmful outcome? These two moral concepts are of course related. Remedial responsibility for a harm means responsibility for fixing or otherwise compensating for the harm. A prime candidate for remedial responsibility, but by no means the only candidate, is the agent who is suitably blameworthy with respect to the harm (if there is such an agent). Blameworthiness is harder to pin to particular outcomes, so it is not surprising that it is often difficult to work out who in fact is blameworthy when it comes to a particular outcome. This basic point will turn out to be a major theme of the paper. Roughly speaking, being blameworthy means being an appropriate target of blame, where blame arguably means ’holding an agent to account’, and may involve reactive attitudes like disgust. Some claim that blame has a role to play in managing relationships and/or expectations for the future. The legal analogue of this pair of concepts, i.e. remedial responsibility and blameworthiness, is civil liability versus criminal liability (and indeed these terms are used by Feinberg 1968 in his seminal paper ‘Collective Responsibility’).

As per many other authors, I will restrict my attention here to moral responsibility understood as blameworthiness. While remedial responsibility is of course very important (and admittedly may even be the main concern of those worried about responsibility voids), it is a complicated issue. Determining

---

2 Note that this paper focuses on harmful activity rather than beneficial activity, because harms make the issues more vivid.
3 Here I am siding with a Strawsonian-style account of blame (see Strawson 1962). Scanlon’s (2008) account of blame is in the same ballpark, in that blame has a pragmatic role, and is in a sense conceptually prior to blameworthiness.
remedial responsibility for a harmful outcome (whether it be the blameworthy agent, the causally responsibility agent, the main beneficiary or someone else with a special relationship to the outcome) is a messy business, dependent on the particularities of one’s moral theory as well as contextual details. Blameworthiness, on the other hand, promises to be a tidier theoretical concept, i.e. one that permits an analysis that cuts across different moral theories. Moreover, getting clear about who is blameworthy with respect to an outcome still has pragmatic relevance. For one, there is the link between blameworthiness and remedial responsibility for a harmful outcome that has already occurred. Beyond that, there is the future-oriented role of blame—the acknowledgment that something ought to change. If nobody can be blamed for a harmful outcome that has occurred, then the same situation could very well arise again in the future. For this reason, in particular, voids in blameworthiness where harmful outcomes are concerned at least deserve attention.

The paper is organized as follows: Section 1 introduces the classic cases that inspire discussion of collective responsibility qua blameworthiness: cases of responsibility voids due to so-called ‘non-reductive collective fault’. Section 2 gets clearer on the notion of blame, and casts doubt on the peculiarity of the voids in blameworthiness introduced in Section 1, and thus the motivation for a special notion of collective responsibility. Section 3 investigates how group contexts nonetheless exacerbate issues of individual blameworthiness that lead to voids. Section 4 concludes.

1 Classic responsibility voids in the literature

The most striking and commonly cited candidates for collective responsibility in the literature involve responsibility voids that arise due to apparent ‘non-reductive collective fault’ (for this terminology, see, for instance, Feinberg 1968 or Narveson 2002). The idea is that the group as a whole perpetrates a harm that seemingly warrants considerable blame, but the individuals involved are relatively blameless. Curiously, the two key types of scenario that purportedly exemplify non-reductive collective fault or blameworthiness involve very
different types of groups: on the one hand, random collections of individuals that face coordination problems, and on the other hand, highly integrated ‘corporate agents’ with sophisticated procedures and voting rules. The two types of group scenario will be illustrated below, after which we turn to some prima facie difficulties with introducing a notion of group blameworthiness.

First let us consider the ‘corporate group’ case. This kind of group is one that has formal procedures and voting rules in place that allow the group to ‘act so as to pursue a collectively endorsed body of desires according to a collectively endorsed body of beliefs’ (List and Pettit 2011, p 158). The most stark responsibility voids arise when corporate groups confront a ‘discursive dilemma’ (for discussion of such cases, see List and Pettit, pp 166-167, Copp 2006, Braham and Van Hees 2011). The following example is lifted from Braham and Van Hees (2011, p 11):

*Corporate Discursive Dilemma:* A three-member committee of an employee-owned firm has to decide whether or not to impose a pay-cut in order to finance new workplace safety measures. It is agreed that the decision to impose the pay-cut and implement the safety measure is to be taken if an affirmative group verdict is reached on three issues, each of which is determined by simple majority.⁴ The committee members M1, M2 and M3 vote as follows, giving the stated group opinions:

<table>
<thead>
<tr>
<th></th>
<th>Serious danger?</th>
<th>Effective measure?</th>
<th>Bearable loss?</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>M2</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>M3</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Group</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

⁴ In other words, the agreed procedure is a *premise-based* approach to group decision-making, where each of the premises is decided by the simple-majority voting rule. Any conclusions entailed by the premises are automatically added to the group stock of opinions.
Accordingly, the safety measure goes ahead, financed by the pay-cut. Let us assume here that this outcome is harmful, by objective moral standards (just for simplicity, as our focus here is harmful outcomes).

It is clear why this is considered a case of non-reductive collective fault or blameworthiness. By logical implication, none of the individual group members, M1...M3, hold the opinion that the pay-cut should be implemented to finance the safety measure. So each individual appears relatively blameless. Yet the majority vote on each of the premises is affirmative, and so the group opts for the pay-cut. It looks like the group is to blame for this outcome, even if the individual group members are not.

We turn now to a rather different case—a random collection of individuals facing a coordination problem (see, for instance, Feinberg 1968, Held 1970, Jackson 1987, Parfit 1984, Lippert-Rasmussen 2011). This is to say that the ‘group’ is not organized via formal procedures and voting rules to act in accordance with collectively endorsed beliefs and desires. Rather, individuals act according to their own beliefs and desires, where this includes beliefs about what others in the group will do. There are various types of strategic situations where groups may fail to achieve a Pareto-efficient outcome and may furthermore deliver a harmful outcome (depending on what is taken to be the baseline for ‘harm’). One scenario of this sort is the well-known Prisoners’ Dilemma. The following example is lifted, with only slight modification, from Feinberg (1968). (I consider this an example of a so-called n-person Stag Hunt game):

*Stag Hunt Coordination Dilemma:* Consider the case of the Jesse James train robbery. An armed man holds up an entire car full of passengers, threatening to kill one hostage. If the passengers had risen up as one man and rushed at the robber, one or two of them, perhaps, would have been shot; but collectively they would have overwhelmed him, disarmed him, and saved their property and the hostage. Yet they all meekly submitted, opting for the safer option of keeping quiet, rather than the risky option of rushing at the man in the hope that others would also choose this option.
Again, it would be difficult to argue that the individuals in this scenario are to blame for choosing the safe option rather than the option with a high (subjective) probability of death (given justified beliefs about what others will do) and a small probability of avoiding a harm. Yet collectively the group allows a harm to happen (let’s say the hostage is killed) when it was in their power to prevent this harm. It is tempting to say that the group is blameworthy for this outcome, even if the individual members are not.

It is all very well to say that these are clear cases of a deficit in individual blameworthiness for harmful outcomes, where the missing blameworthiness looks to be at the group level, but this naturally raises the question: What does it mean to consider a group blameworthy at all? The next section will examine in detail the concept of blameworthiness with respect to outcomes; for now, let us consider two basic reasons to be sceptical of the notion of collective, as opposed to individual, blameworthiness.

The first reason for scepticism concerns what many hold to be a necessary condition for blameworthiness: that the agent concerned is *agentially responsible* for acting in the way that they did. Depending on how strictly one takes the relevant notion of agential responsibility to be, only one or else neither of the above groups satisfies this condition. The random collection of individuals in a strategic environment hardly acts as a group agent. As noted already, there are no collectively endorsed beliefs, desires or choices to speak of in this case. Various authors argue that it is possible for corporate groups, on the other hand, to meet the standards of agency required for agential responsibility. Even so, the corporate group described above—the committee that convenes to decide on pay-cuts and associated safety measures—does not necessarily satisfy these standards, depending on how they are interpreted. If agency merely requires intentional action in line with beliefs and desires, then the pay-cut committee looks to meet the agency condition. If, on the other hand, agency necessarily has a reflective aspect, as is commonly held (see List and Pettit 2011, Braham and Van Hees 2011, Gilbert 2006 and numerous authors on individual agential responsibility), then the pay-cut committee does not look to satisfy this
condition. The idea is that agency is not just about the ability to act on one’s beliefs and desires, but also about the ability to revise these beliefs and desires in response to evidential and higher-order considerations. Indeed, many authors stipulate that the agent must be able to distinguish good from bad outcomes, and ultimately right from wrong, and assess their own choices accordingly. The pay-cut committee qua group can form collective opinions that underpin an intentional collective decision to act, but there is no facility at the group level for reflecting on the resultant group opinions. In order to do this, the group would need to explicitly entertain further propositions like ‘implementing a pay-cut for the reasons stated aligns with general moral principles’, or ‘majoritarian voting is a morally adequate standard of evidence for the premises underlying the pay-cut decision’.

The fact that the groups described above do not look to satisfy standards of agency qua groups that are widely regarded as necessary for holding an agent blameworthy is not in itself reason to foreclose the possibility of collective blameworthiness in these scenarios. If one takes a pragmatic approach that makes blaming itself more primary than blameworthiness or agential responsibility (a la Strawson), then the key question is whether it makes sense to blame groups; if so, in those cases groups are blameworthy. And this may well imply that a different notion of agential responsibility is appropriate for groups as compared to individuals. Having said that, we would want to have rather compelling cases of blaming targeted at groups to warrant introducing a new notion of agential responsibility applicable to groups. And it is not clear that we can explain how blaming groups would work at all, let alone describe compelling cases of this phenomenon.

The trouble with blaming groups like those described above, qua groups, is, quite frankly, that there is no obvious point to doing so. Unless there is a reflective or at least intentional agent to whom the blame is addressed (as per the traditional

---

requirement of agential responsibility) the blame would seem to have no target.\(^6\) It would be a waste of anyone’s time and moral energy to do any such blaming because there would be nobody to take it on, so to speak. To get around this, one might propose that the blame be directed at the individual group members, even though the overall collective blameworthiness is ‘non-reductive’ in the sense discussed above. Cases of blaming groups would be unusual cases where we end up blaming individual members of a group for an outcome, even though they have each acted in a way that would not ordinarily provoke (the same extent of) blame, i.e. they are not individually blameworthy to the same extent. But this account of blaming groups is no more satisfying. Indeed, some authors have argued (e.g. Mark Reiff, as reported in Smiley 2011) that such a move effectively severs the link between blameworthiness and control over one’s actions, and this threatens to encourage apathy in the face of blame. The more direct criticism is that the move is simply inappropriate or even conceptually confused. If the situation calls for blaming the group, then the whole point is surely that this blame cannot be distributed to individuals (cf. Narvesan 2002). In short, holding groups *qua* groups responsible (i.e. blaming them) does not look to be a workable solution to the problem at hand.

2 On blameworthiness, and the prevalence of voids

This section argues that blaming groups may not, in any case, be well motivated, and that we should cast a more critical eye on the supposed problem: While responsibility voids are problematic, they are not in fact rare or unusual, and not at all peculiar to group scenarios. This is essentially because the concept of blameworthiness does not straightforwardly attach to *outcomes*. Blameworthiness is really a matter of an agent’s subjective attitudes—her beliefs and desires—and these attitudes do not always track the true causal structure of the world. This alone means that there is much scope for blameworthiness voids when it comes to ‘real-world’ outcomes. Moreover, we will see that the notion of ‘non-reductive collective fault’, introduced in the last section as a feature of the

\(^6\) This is to say that blame may have a point if the target is the corporate group described, but certainly not if the target is the random collection of individuals.
classic motivating cases for collective responsibility, is suspect from the outset, as blameworthiness is not the sort of thing that should be expected to add up to some predetermined total.

Much hinges, then, on how we should understand blameworthiness. Some preliminary considerations were introduced in the previous section: Recall the general view that agential responsibility is necessary for blameworthiness. As noted above, agential responsibility is cashed out in terms of whether an agent chooses in accordance with her beliefs and desires, subject to due reflection. So agential responsibility is a matter of how an agent deliberates about what to do. This would suggest that agential responsibility is, strictly speaking, a property of an agent’s acts, or rather her choices, depending on what is taken to be the relationship between choice and act.\(^7\) (From now on, I will refer just to acts.) To say that a person is agentially responsible for her act is to say that she \textit{arrived} at this choice of act via an appropriate method of deliberation, or at least that the act \textit{reflects} her well-considered beliefs and desires. Of course, it is natural to talk also of agential responsibility for a particular outcome or event in the world. For example, one might say that ‘the young man is agentially responsible for the car being written-off’. But this is apparently just loose talk, and can be decomposed as follows: ‘the young man is agentially responsible for, say, choosing to drive recklessly, and this reckless driving was causally implicated in a foreseeable way in the car being written-off’. The assumption here is that responsibility for something means having full control over that thing in accordance with one’s will. While this control extends to our well-reflected acts, it does not, for obvious reasons, extend to wider-world outcomes.

The concept of blameworthiness will inherit these features of agential responsibility. My own view of the relationship between the two is along the lines of Strawson (1962): an agent is blameworthy just when it makes sense to blame the agent, where this amounts to calling the agent to account in some (active/expressive) way. Strictly speaking, it only makes sense to call an agent to

\(^7\) If acts are by definition things that an agent has full control over (as per Joyce 1999, p57–61), then an agent’s choices coincide with her acts. Here acts are limited to mental willings, or something along these lines.
account in this way for that which is under her control and which furthermore reflects her character. As such, it only makes sense to call an agent to account for that which she is agentially responsible, and it was noted above that this is limited to the agent’s well-reflected acts. The upshot, then, is that an agent can be blameworthy only for her well-reflected acts.

Of course, as per agential responsibility, we also talk of agents being blameworthy for outcomes. To give a dramatic example, we blame Stalin, amongst others perhaps, for millions of civilian deaths in Soviet Russia in the mid twentieth century. Moreover, here it seems that the blame should be directly associated with the horrific outcome, and that anything less would somehow not respect the gravity of the situation. That is, limiting the blame to Stalin’s decision-making psychology seems intuitively too weak or evasive. Notwithstanding these considerations, I contend that the more careful analysis of ‘Stalin is blameworthy for the deaths of millions of civilians’ is as follows: ‘Stalin is blameworthy for a series of choices/acts that were causally implicated, in a significant way, in the deaths of millions of civilians.’

We need not get too hung up on whether it is right to speak of an agent being agentially responsible and/or blameworthy for an outcome. The relationship between blameworthy acts and outcomes is, in any case, an important one, and worth investigating. What we want to know is when a blameworthy act and some outcome are connected in a morally significant way, whether we speak of blameworthiness with respect to the outcome or of blameworthiness for the outcome or whether we use some other locution. The former expression better reflects the analysis just given, so let us stick with that terminology from now on. By way of response to the key question, a proposal in the spirit of Feinberg (1968, p. 674), with minor modifications, is not a bad starting point:

For an individual to be blameworthy with respect to an outcome, the following three conditions must be satisfied:
1. **Causal relevance**: It must be true that the agent did the harmful thing in question, or at least that her act or omission made a substantial causal contribution to it.

2. **Blameworthiness**: The causal contributory act must have been (in some way) blameworthy.

3. **Blameworthiness relevance**: The perceived possibility of the outcome resulting from the act must be the reason the act was blameworthy (in the way specified in 2).

These three conditions each raise a host of interpretative issues. We have attended already to the concept of blameworthiness (and related *agential responsibility*); while there is more to say on this, the finer details do not matter for our purposes.⁸

The remaining two conditions are no less complicated. Causation is of course a highly contested concept, and so there are many plausible ways to understand the causal relevance condition. For now, note just that this condition concerns singular causation rather than general or average causation. For instance, it is not the fact that Betty’s drink-driving raised the probability of the death of a pedestrian that is at issue. For this condition we are interested, rather, in the causal connection between a particular act and a particular outcome: Was Betty’s drink-driving a contributing cause to the outcome that actually happened, let’s say the death of a pedestrian? The third ‘blameworthiness relevance’ condition, on the other hand, does implicate average causation, because it involves causal prediction. This condition is arguably the most complicated part of the whole package; it concerns the agent’s own reasons for acting and the way or extent to which her reasoning was blameworthy because it involved or should have

---

⁸ For instance, does agential responsibility for an act require that the agent was able to choose/do otherwise, or does it suffice that the agent chose/acted in accordance with her beliefs and desires? Does blameworthiness require that the agent reflect specifically on the moral status of her choices/acts, or does it suffice that the agent endorses, in some other sense, her relevant beliefs and desires?
involved some degree of belief that the outcome in question would result from her act. So, for instance, in deliberating about whether to drink-drive, Betty believed or should have believed that this act would significantly raise the probability of a pedestrian dying (from being hit by her car). In this case, it seems that the blameworthiness-relevance condition is satisfied with respect to Betty's drink-driving and the actual outcome of a pedestrian dying having been hit by Betty's car.

We see that blameworthiness with respect to an outcome is a highly qualified notion. Blameworthiness for acts arguably already comes in degrees—one can be more or less blameworthy for the sorts of acts one chooses. Things get much fuzzier, however, when we try to establish the strength of the link between the blameworthy act and some outcome. For starters, the actual causal connection between act and outcome may (on some accounts) be more or less strong. Beyond that, there is the related but nonetheless different issue that the outcome itself may be more or less implicated in the blameworthiness of the act. There is clearly much more to be said on the topic of blameworthiness with respect to outcomes, and we could test our intuitions against a host of subtly different cases. But let us not get embroiled in the details here. Note, however, that many of the issues will depend on what one ultimately thinks is the moral significance of the connection between a blameworthy act and an outcome. If the real concern is the harmful outcome itself (perhaps for the purposes of determining remedial responsibility), then the metaphysical connection between blameworthy act and outcome is presumably what is at issue. This seems to be the presumption of the Feinberg-inspired conditions that we have been discussing. If, on the other hand, the real concern is justified blame and blameworthiness, then it is rather the evidential connection between blameworthy act and outcome that is relevant. Indeed, the outcomes of a person’s acts, or in other words, the traces that a person leaves in the world, are often the most reliable evidence of her psychology, as well as its causal significance or power to effect change in the world. It is therefore no wonder that outcomes trigger blame, and we naturally talk of blameworthiness for outcomes,
even if, on closer inspection, it turns out that, for blame to have a point, it should be reserved for that which is a direct expression of the agent’s character.

From the above condensed account of blameworthiness, it should already be obvious that there are going to be many ‘responsibility voids’ associated with blameworthiness with respect to an outcome, that is, cases where an intentional act is apparently causally implicated in an outcome, but where we cannot establish ‘full blameworthiness’ with respect to the outcome (whether we pursue the metaphysical or the evidential route, as outlined above). The significance of this, given our purposes here, is that it should give us pause in thinking that responsibility voids associated with groups deserve special attention. We will return to the group context in the next section, but for now, it helps to catalogue five major reasons for responsibility voids or rather ‘disconnects’ when it comes to blameworthiness with respect to an outcome. These can be labelled as i) causal anomalies ii) constrained choice, iii) bad luck and iv) good luck and v) causal ignorance. Below, a brief account of each is given in turn.

Causal anomalies. The causal relevance condition for blameworthiness with respect to an outcome raises problems of a technical nature that may well be overcome by a suitably sophisticated notion of causation. (Note that this condition is more pertinent to the metaphysical account of blameworthiness with respect to an outcome.) Classic problem-cases for any account of causation include (but are not limited to) omissions or failures-to-act that we intuitively think caused a particular outcome, and cases of over-determination.\(^9\) Consider an example of the latter case: Imagine that Sarah intended to prevent her brother from getting the great job that he wanted, and to this end, she hid the car keys so that her brother would be unable to get to the interview on time. Unbeknownst to Sarah, the car had unforeseen engine problems that day, and so Sarah’s brother would have been unable, in any case, to get to the interview on time, due to the car not starting. Did Sarah cause the bad outcome that her brother was unable to get to his interview on time? Many would say that Sarah is a cause of this outcome, at least in the sense relevant for blameworthiness. The problem

\(^9\) For discussion of these problem cases, see Sartorio (2007).
here is that, according to all the standard, deterministic notions of causation, i.e. the sufficiency, necessity, necessity & sufficiency and but-for varieties of causation, Sarah does not cause her brother to be late for his interview. This is because Sarah’s act is not sufficient for her brother being late (it depended also on when he decided to get ready), and nor is Sarah’s act necessary for her brother being late. It is not even the case, due to the over-determination, that ‘but for Sarah’s act in this particular circumstance, her brother would not have been unable to get to his interview on time’.

Note that Braham and Van Hees (2009) appeal to a notion of cause as a ‘necessary element of a sufficient set’, the so-called NESS-test for causation (similar to Mackie’s INUS condition approach) to account for over-determination cases. For our example, Sarah’s act satisfies the NESS-test for causal relevance because her hiding the keys is an element of at least one minimally sufficient set of elements for her brother being unable to make his interview on time. This interpretation of the causal relevance condition is promising, but it is not clear, without further investigation, that it will also account for the other problematic cases, like omissions that are intuitively causally relevant. The standard difference-making or but-for test of causal relevance may be more promising in the case of omissions. Furthermore, one may not be convinced that over-determining ‘causes’ really ought to count as genuine cases of causal relevance. Perhaps these are simply cases where we should acknowledge a gap in blameworthiness with respect to an outcome: a blameworthy act was performed, a harmful outcome occurred that is related to the act, and yet nobody is blameworthy with respect to the outcome.

---

10 This list of deterministic notions of causation is from Braham and Van Hees (2009). Like these authors, I restrict my attention to deterministic notions of singular causation. I will not consider the possibility that the laws of nature are such that single chains of events are best described in terms of an indeterministic notion of causation, i.e. causal claims involving objective chances, as in Vallentyne (2008).

11 An INUS condition is an ‘insufficient but non-redundant part of a condition which is itself unnecessary but sufficient for the occurrence of the effect’.

12 Roughly speaking, a ‘minimally sufficient causal set’ is one that has no causally extraneous events—each event in the set is necessary for the sufficiency of the entire set for bringing about the outcome in question.

13 For instance, the omission or failure of Sarah to feed her dog is intuitively causally relevant to the dog dying, but can we say that not-doing-something is an element in a sufficient set for the occurrence of some effect?
**Constrained choice.** These are cases where a person acts and causes foreseeable harm, but the act was nonetheless not blameworthy due to the constraints of the choice situation—‘damned if you do and damned if you don’t’ kinds of scenarios. For instance, consider the predicament faced by the mountaineer Joe Simpson during his expedition in the Peruvian Andes: He could choose not to cut the rope from which his injured friend was dangling, and likely cause them both to die of cold and exhaustion. Or he could choose to cut the rope, which would likely kill his friend, but save his own life. Joe cut the rope. In the event, his friend lived, but had he died, this would surely be a horrible outcome that was the foreseeable result of a deliberate yet (by most accounts) blameless act. Thus Joe is not blameworthy with respect to the supposed death.

**Bad luck.** This category is similar to the previous one, but where the constrained aspect of the choice situation is compounded by bad luck. Here the agent acts in a way that is blameless, but not obviously so, as the selected act has a calculated risk of significant harm, according to the best available evidence. Consider, for instance, a medical scenario where an eye-doctor has to make a treatment decision for a patient who for special reasons cannot make the decision herself. There may be a range of possible treatments with different risk profiles as regards prospective quality of sight and other health factors. The doctor chooses a treatment that she justifiably considers to be in the interests of her patient despite a small probability of it leading to permanent blindness. Unfortunately, the bad outcome does occur, and the patient is made permanently blind by the treatment. Here again there is no blameworthiness with respect to the patient’s blindness, even though the doctor chose the treatment deliberately and foresaw the possibility of the harmful outcome; the act was simply not blameworthy because the choice situation was constrained—the best act was a risky one.

**Good luck.** This is the flipside of the previous case: here the agent performs a blameworthy act but the outcome that occurs has relatively little to do with this blameworthiness. For instance, assume that Betty believes (or should believe, given the evidence accessible to her) that drink-driving has a very small
probability of resulting in the death of a pedestrian, a significant probability of resulting in a minor injury to a pedestrian, and the remaining probability of resulting in no harm. Betty deliberates and chooses to drink-drive. The actual outcome is minor injury to a pedestrian. Is Betty blameworthy with respect to the outcome of minor injury to a pedestrian? While Betty's act is seriously blameworthy, only a relatively small amount of this blameworthiness has to do with the outcome of minor injury (because, let's say, most of the blameworthiness of the act of drink-driving is not attributable to the risk of this outcome but rather to the risk of a pedestrian dying). So one might say, as awkward as it sounds, that Betty is highly blameworthy for her act but comparatively less blameworthy with respect to the minor injury of a pedestrian.

_Causal ignorance._ Here the agent's causal beliefs do not even reflect the best available evidence (unlike the previous two cases), and yet the ignorance is excusable.\(^{14}\) For example, consider the scene in Arthur Miller's 'The Crucible' where Elizabeth Proctor tells the court that her husband was always faithful to her. Unbeknownst to Elizabeth, the court already knows that John Proctor was not always faithful, and her telling this white lie contradicts her husband's word about her honesty and results in her husband being sentenced to death. Elizabeth is clearly not blameworthy with respect to this horrible outcome, because, due to her ignorance about the likely causal effect of her testimony, her act was not blameworthy to begin with. A slightly more complicated case is one where an act that is in fact blameworthy is misguided with respect to the causal facts but nonetheless results in some harm. To give another literary example, consider Prince Hamlet's attempt to avenge his father's death by murdering his uncle King Claudius. Hamlet strikes and kills the man lurking behind the curtain in his mother's chamber, but it turns out not to be the King but rather Polonius. Hamlet's act is arguably blameworthy, at least by modern standards, and yet the

\(^{14}\) Of course there are huge questions surrounding the notion of 'non-culpable ignorance'—whether it is limited to factual opinions, or may apply to moral/evaluative opinions as well, whether the question is rather when the agent was at fault (perhaps at some earlier time when they should have sought the relevant evidence) as opposed to whether the agent is at fault, and so on. To a large extent, it depends on one's basic views about blameworthiness and blaming. Here we simply assume that sometimes ignorance of the causal facts (based on best available evidence) is excusable; the examples given are taken to be compelling cases.
outcome that he intended and which made his act blameworthy (the death of the
King) is not the outcome that actually occurred; one can thus argue that despite
doing something blameworthy and thereby causing a harmful outcome, Hamlet
is not blameworthy with respect to Polonius’ death.

We see that there are many opportunities for ‘responsibility voids’ when it}
comes to blameworthiness with respect to harmful outcomes, or in other words,
there are many somewhat puzzling mismatches between blameworthy acts and
harmful outcomes. For some of these cases, the agent is more blameworthy for
the act than they are blameworthy with respect to the outcome in question. In
other cases, it is rather the other way around: the agent is less blameworthy for
the act than we might expect given the harmfulness of the resulting outcome.
Either way, when we go looking for blameworthiness with respect to the harmful
outcome, things are not necessarily as we expect. But none of this should come
as a surprise. Indeed, blameworthiness with respect to an outcome should not
generally be expected to ‘add up’ to anything in particular. There is no such
thing, really, as the appropriate amount of blame for any particular kind of
outcome, because the same kind of outcome can be produced in all sorts of ways.
Moreover, there is no such thing as the appropriate amount of blameworthiness
that an act should transmit to the outcome it actually yields. This is because
agents’ beliefs do not necessarily match the true causal structure of the world,
even when these beliefs are fully justified from the subjective point of view.

It follows from these observations about ‘responsibility voids’ that we also have
good reason to be sceptical of the notion of ‘non-reductive collective
fault/blameworthiness’, especially when cast as for, or with respect to, an
outcome. In that case, even our best motivating cases for blaming groups qua
groups, from the pragmatic point of view, are seriously undermined. The
problem is that there is no obvious sense in summing the blameworthiness of
individuals, and the question arises: With respect to what total blameworthiness
might we calculate a shortfall? Indeed, it may well be argued (see Zimmerman
1985) that the very idea of calculating collective blameworthiness for an
outcome is dangerous because it leads to thinking that there is some fixed
amount of blameworthiness with respect to an outcome that is simply shared amongst the identified perpetrators. Perceived individual blameworthiness may in this way be improperly diminished.

The upshot of the discussion thus far, then, is that the notion of collective blameworthiness has little pragmatic point. We see that neither the problem of responsibility voids due to collective blameworthiness (discussed in this section), nor the solution of blaming groups (discussed in the previous section), is convincing. The next section, however, shows why groups are nonetheless special in considerations of responsibility/blameworthiness.

3 Individual blameworthiness and the group

The main point of this section is to highlight how group contexts make the assessment of individual responsibility difficult, sometimes extremely so; but instead of seeking a quick fix at the level of the collective, it is better just to acknowledge the difficulties. Admittedly, nobody who defends an account of collective responsibility (including Gilbert 2006, Copp 2006, List and Pettit 2011) argues that this should in any way supplant individual responsibility, or that it somehow irons out the problems of individual responsibility. (Indeed, the authors just mentioned defend collective responsibility on the basis that it is possible for groups to satisfy the conditions of agency that apply to individuals, and any such groups are therefore fit to be held responsible.15) There is nonetheless the danger that too much attention is paid to the spectre of collective responsibility when it comes to group contexts, and too little attention paid to the hard problems of individual responsibility. By way of making a small contribution towards addressing this imbalance, this section considers why groups exacerbate issues of individual responsibility with respect to outcomes.

15 In other words, as mentioned in Section 1, the main defenders of collective responsibility defend their position by appeal to bottom-up considerations of agency. They effectively argue that some groups are, in the relevant respects, just like individual persons, and can therefore be held to account. But note that the cases in which groups fulfill ordinary conditions of agency will be few and far between (depending on the strictness of said conditions). Moreover, the main question in this paper is rather whether it is worth appealing to some alternative, special notion of blameworthiness/agental responsibility at the group level, in order to address apparent ‘responsibility voids’.
To that end, I return to the five major reasons for ‘responsibility voids’ introduced in the previous section.

*Causal anomalies.* Groups add to the stock of cases that present problems for defining an appropriate causal relevance condition for blameworthiness with respect to outcomes. In particular, groups permit many cases of causal over-determination that are arguably more compelling instances of causal relevance than non-group cases of causal over-determination. Consider the following example (where the group is merely a random collection of individuals): a number of employees bully one particular colleague until that person resigns under duress. It may be that only three bullying colleagues would have produced the same effect; in this case, however, there were five people involved. Intuitively, it seems right to say that all five were blameworthy for the bullying, and moreover, all five causally contributed to, and thus were blameworthy with respect to, the resignation outcome. So we had better design our causal relevance condition to handle such a case. When it comes to corporate groups with voting procedures for deciding on group acts, the individual members’ causal contributions to the group act/resultant outcome are even more difficult to work out, due to the variety of ways in which a vote can influence the group outcome (dependent on the voting rules). Braham and Van Hees’ (2009) proposed *NESS-test* for causal relevance looks promising here, as well as for cases of over-determination. A further issue that is not settled by the *NESS-test* interpretation of causal relevance, however, is how ‘direct’ a causal contribution an act must make to an outcome. For instance, what should we say about the sixth employee who merely encouraged the others to bully their colleague? Is this person’s act causally relevant to the outcome? We cannot settle these issues here;¹⁶ the overall point is simply that ignoring group contexts when formulating

¹⁶Feinberg (1968, p. 684) makes the following comment about how the law in fact distinguishes various causal roles:

‘The common law, therefore, divides guilty felons into four categories: “perpetrators,” “abettors,” “inciters” (all three of these are “accomplices”), and “criminal protectors,” so that one may be guilty of a given crime either as its principal perpetrator (and even perpetration is a matter of degree, abettors counting as “principals in the second degree”) or as accessories, that is, inciters or protectors.’
the causal relevance condition(s) risks overlooking important cases that help refine intuitions about blameworthiness with respect to outcomes.

*Constrained choice.* The group context can constrain an individual’s choice situation in subtle ways that are relevant to blameworthiness. Of course, the mere presence of other individuals interacting in one’s environment influences the choice set because it affects what possible consequences are thought to result from one’s acts. The Jesse James train hold-up in Section 1 is a cinematic example of people acting in light of how others are expected to act. Beyond sharing an environment, persons are involved, more or less voluntarily, in various sorts of ‘corporate groups’. It is clear how *involuntary* membership in a group (e.g. participation in the family group or citizenship in a particular national group) constrains one’s choice set. Membership of such groups may directly limit one’s options through coercion (think of a national legal system), or else simply due to expectations surrounding the social roles that one ends up occupying within, and in virtue of being a member of, the group. What I want to emphasise is that voluntary membership of groups, too, comes with constraints. Opting to be part of a group *extends* a single person’s agency by allowing opportunities for achieving things that are beyond the reach of one person alone; but we should not forget that the group option generally involves some sacrifice relative to alternative options, as a person’s own objectives and the negotiated group objectives are not likely to perfectly coincide. ‘Going with the group’ is generally a compromise option that may be difficult to assess relative to other options in one’s choice set. By way of a stylised example, consider the decision of Aung San Suu Kyi to accept membership of the partially reformed Burmese Parliament. From her perspective, this move evidently has the potential to yield good and otherwise unachievable political outcomes, but it also risks bad outcomes; the question is how, all up, the expected consequences balance out, relative to other possible options that might have been chosen (such as continuing to boycott the national elections). Chosen acts may appear blameworthy because there is the potential for resulting harm, but the problem may well lie with the constraints of the choice situation, made complicated by the group context, rather than with the act itself.
Bad luck. I will not say too much on this (or the good luck case) because they are variations on the theme of constrained choice sets, as was noted in the previous section. Recall that this is the case where a risky option is selected (in light of the best available evidence) but it goes badly. In the Burma case, a possible scenario is where the greater but still limited parliamentary participation of non-military candidates in fact backfires because it lends credibility to a regime that continues to repress citizens. Despite this being a bad outcome, there is no blameworthiness with respect to this outcome, at least on the part of the reformist candidates, if the acts in question were not blameworthy in the first place, due to the constraints of the choice situation.

Good luck. This is the other case where an act is in fact blameworthy but the resultant outcome is, happily for others, not the worst that could arise. Again, the inherent uncertainties involved in strategic group scenarios mean that such cases are potentially very prevalent. For instance, we can well imagine it often happening that sinister political connivances fail to yield their intended outcomes due to other players acting unpredictably. In such cases, well-informed and blameworthy acts are performed, yet harmful outcomes do not materialise.

Causal ignorance. Finally, there is much potential for causal ignorance in the group context, and to make matters worse, it will often be unclear whether the ignorance is excusable or not, as it is not clear what one should be able to predict about others’ behaviour. For instance, Held (2002) discusses cases where a number of people together participate in hate speech, thus affirming each others’, say, reprehensible ethnic prejudices. Some persons, however, take the hatred further and commit physical acts of violence against the targeted ethnic group. The question is whether the other ‘group members’ (those who share the prejudice but did not participate directly in the violence) are blameworthy with respect to the violent outcome. Of course, there is the issue raised earlier as to how direct one’s causal contribution to an outcome must be in order to satisfy
the causal relevance condition for such an outcome. But there are also more basic questions about the extent of blameworthiness—should those participating in hate speech have predicted that this would likely lead to violent outcomes? If so, their blameworthiness goes beyond ‘mere’ hate speech. This case also indirectly highlights further issues of social epistemology—people (justifiably) use others’ testimony as evidence by which to assess their own factual (and moral) opinions. It may happen that a group of people mutually reinforce each others’ mistaken beliefs, so that there is resulting widespread ignorance that is the fault of nobody. Here again the group context may well exacerbate ‘responsibility voids’ with respect to harmful outcomes. Intentional acts are performed which result in harmful outcomes, but due to excusable ignorance, nobody is blameworthy for the acts, let alone for the outcomes.

4 Concluding remarks

This paper has focussed on so-called ‘responsibility voids’ with respect to outcomes, or in other words, puzzling ‘disconnects’ in blameworthiness between intentional acts and harmful outcomes. The main claim is that such disconnects are exacerbated by, but are not unique to, group contexts. While groups certainly make assessments of individual blameworthiness difficult, especially with respect to outcomes, there is no reason to appeal to a special notion of collective blameworthiness, even if this were a workable notion.

It is also worth recalling why we should be interested in blameworthiness with respect to outcomes in the first place, and what is so worrying about responsibility voids, whether in group contexts or not. The metaphysical connection between blameworthy acts and harmful outcomes has some importance for determining remedial responsibility. As mentioned earlier, however, remedial responsibility is an eclectic notion, and any voids in

---

17 We noted earlier (see footnote 15) that it might make sense to have several causal relevance conditions, to distinguish, say, “perpetrators,” “abettors,” “inciters” and “criminal protectors”.

18 Note that the possibility of non-culpable factual ignorance is less controversial than the possibility of non-culpable moral ignorance. Many people would resist the idea that it is excusable to falsely believe that ethnic prejudice is morally permissible (but see Zimmerman 1997 for the contrary view).
blameworthiness as a basis for remedial responsibility can be filled in other ways (by appeal to other special relationships between actors and outcomes). My own view (which I did not so much push in this paper) is that the more interesting connection between acts and harmful outcomes is an evidential one. The real concern should be to identify blameworthy acts; outcomes simply signal what kinds of acts are blameworthy and whether we have an instance of one at any given time. On this view, responsibility voids are worrying from a future-oriented perspective. If a harmful outcome does not point, evidentially, to a blameworthy act, what is to prevent such an outcome from occurring again?

This last consideration should give us pause. While no one may be blameworthy with respect to a particular harmful outcome, its occurrence is not irrelevant for the future choices of individuals. The harm may still serve as evidence of the consequences of acts in a particular context that effectively reduces the scope of non-culpable ignorance in the future. Moreover, there is something ‘non-reductive’ about the kind of response that is often demanded in light of this evidence, even if the notion of ‘non-reductive collective fault’ is not itself compelling. Agents in a group context must, in some cases, appeal to each other to improve their group organisation, in order to better coordinate their acts so as to avoid harmful outcomes.
Bibliography


