

## Collective responsibility and collective obligations without collective moral agents

Gunnar Björnsson, Stockholm University  
gunnar.bjornsson@philosophy.su.se

Sometimes bad things are the aggregate result of actions and omissions of more than one agent. In such cases, we often blame that group of agents for the result. Some blame their divorced parents for not preserving a reasonable relationship. Others blame voters for electing corrupt leaders. Others yet blame the growing threat of climate catastrophe on people who have known about the risks of greenhouse gas emissions without taking appropriate measures.

As it is with blame, so it is with obligations. We often think that groups of agents have obligations that no individual could discharge on their own. The parents had an obligation to preserve their relationship, we might think, while the voters had an obligation not to elect corrupt leaders, and people aware of climate risks an obligation to significantly reduce those.

Attributions of blameworthiness to groups are not mere theoretical exercises. They direct our indignation over what has happened and indicate who should compensate those harmed or bear the burdens involved in preventing further negative consequences. Likewise, attributions of group obligations often seem relevant for what actions group members should take, for whether the group can be justly coerced to behave in a certain way, and for distributions of blame and compensatory or reparative burdens should the group shirk its duties.

Though common and seemingly important, attributions of group blameworthiness and group obligations can also seem metaphysically or ethically suspect. Often, no individual member of the group had control over the outcome for which they are blamed, and no individual member can make a difference as to whether the group discharges its obligation. This makes it difficult to understand group attributions in terms of attributions of corresponding individual blameworthiness and obligations. Moreover, the groups themselves often fall short of standard conditions of moral agency. They seem to lack many properties normally associated with agenthood, including beliefs about their circumstances, and they lack the sort of stable inner organization that might make it clear what capacities they have and what demands can be properly directed at them, other than those directed at their members.

In response to this agency challenge, philosophers who want to defend attributions of collective obligations to groups of these kinds have either (i) argued that the groups in question have the requisite capabilities to have obligations of their own or (ii) suggested ways in which the existence of related individual obligations can make it true that these groups have obligations. Philosophers who have defended attributions of collective responsibility and blameworthiness have suggested that members of the relevant collectives can *share* responsibility for an outcome in virtue of being causally or socially connected to that outcome.

This chapter details some cases where it is natural to attribute obligations or blameworthiness to groups that cannot be plausibly attributed to their individual members, and discusses the agency challenge mentioned above as well as proposed replies and problems and prospects for these. The most promising replies, I will argue, understands these groups' obligations and blameworthiness as grounded in demands on individual agents.

**Attributable to the group, but not to any individual**

Consider:

*Offshore wind:* Like most summer days, a large group of children are enjoying the beach, playing on their air mattresses close to land. This particular day, there are fifteen of them; three came with a parent, the other twelve live nearby. Without warning, the weak onshore wind quickly turns strongly offshore, and the children begin drifting out, beyond the range of their swimming capabilities. They need to be rescued. In response, each of the three adults can swim out and catch one child before the wind has carried it too far, provided that she starts swimming now. But there is also a lifeboat that could be dragged a few yards to the sea if at least two adults joined forces; with it, two of them could row out and pick up all the children. Each of the adults quickly realizes all this. (This case, like many discussed in the literature, is modelled on cases in Held 1970.)

About this situation, it is very natural to think that

THREE'S OBLIGATION: The three adults have an obligation to save all the children.

If the group lacks any weighty reasons not to save all the children, the obligation is naturally understood as an *all-things-considered* rather than a mere *pro tanto* obligation. (In what follows, when I talk about obligations, I will have the former kind in mind unless otherwise stated.)

Clearly, to attribute this obligation is not to attribute to each of the three adults an obligation to save all the children, as neither of them can do this on her own: the relevant reading of THREE'S OBLIGATION is not straightforwardly *distributive*. Instead, one might be tempted to understand it as saying that each individual has an obligation to *help* save all the children: to participate or do her part in or contribute to a joint action of saving all. But whereas THREE'S OBLIGATION is independent of the three adults' willingness to help, their individual obligations to help save all are not. Consider:

*Plenitude:* Each of the three adults is willing and ready to help save all the children. Because of this, if one of them didn't help, the other two would join forces and save all. (None has stronger normative reasons to help than the others.)

*Helplessness:* Each of the three adults is a closet racist. Because the other twelve children are of what she considers the wrong color, she is unwilling to help save all the children and would refuse if prompted.

*Scarcity:* One adult is ready to help save all the children and would do so together with whoever would be willing to help her. But the other two are closet racists, unwilling to help, and would refuse to help if prompted.

In *Plenitude*, no individual adult has an obligation to help save all, because the other two will gladly do what is required. In *Helplessness*, no individual has an obligation to help save all because the reluctance of the others makes it impossible for her to do so. Here, the individual might have an obligation to quickly try to get the others to join her in saving all, say, before turning attention to her own child. But because she *cannot* help save all, she has no more of an individual obligation to do so than the adult in

*Solitary helplessness*: Like *Helplessness*, but there is only one adult on the beach when the wind turns, and she can save only one child.

In *Scarcity*, finally, the willing individual has no obligation to help save all the children because the unwillingness of the two racists makes it impossible for her to do so. By contrast, each of the racists can help save all the children (thanks to the willing individual) and her help is needed (because of the unwillingness of the other racist). Each of the racists thus seems to have an individual obligation to contribute.

The plausibility of THREE'S OBLIGATION seems independent of these variations in the willingness of the members of the group and their resulting individual obligations. They have an obligation to save all regardless of whether more or fewer individuals than needed are willing to help.

In light of this, one might be tempted to interpret THREE'S OBLIGATION as saying that each individual has a *conditional* obligation to help: an obligation to help *if that would make a difference*. Each adult in each of *Plenitude*, *Helplessness*, and *Scarcity* plausibly has such an obligation. But the same seems true about the lonely adult in *Solitary helplessness*, and we don't want to conclude that she has an obligation to save all the children. Moreover, even when individual conditional obligations coincide with a corresponding group obligation, the proposal gets things wrong. Consider *Helplessness*: Since no individual adult could have made a difference to whether all children were saved, none of the conditional obligations is violated when each proceeds to save only their own child. But THREE'S OBLIGATION most definitely is.

Although THREE'S OBLIGATION is independent of the willingness of members to help, it clearly does depend on the group's ability to save all children, which in turn requires that there are things that the individuals can do that, together, results in the saving of all the children. Because the group's obligation is independent of directly corresponding individual obligations but is dependent of what the individuals can do together, it is natural to say that it is their *collective* obligation. And the same reasoning seems applicable to the other cases of group obligations mentioned earlier: the obligation of the parents to uphold a reasonable relationship, the obligation of eligible voters not to elect corrupt leaders, and the obligation of people aware of the risk of climate catastrophe to significantly reduce it. They all seem to be collective obligations. (For related arguments, see e.g. Aas 2015; Björnsson 2014; Forthcoming; Dietz 2016; Schwenkenbecher 2014; Wringe 2016; cf. Copp 2007, though he focuses mainly on institutional groups.)

*Offshore wind* can also illustrate the phenomenon of collective responsibility and collective blameworthiness. (For simplicity, I will focus on blameworthiness rather than responsibility more generally.) Specifically, consider the *Helplessness* version, where each of the three adults swims out to save her own child. Here it seems that the adults are responsible for not saving all the children, and deserve blame for not doing so. After all, they could easily have done so by joining forces in some constellation or other, and they didn't because of their lack of regard for the lives of the twelve. Corresponding to THREE'S OBLIGATION, we thus have:

THREE'S BLAMEWORTHINESS: The three adults deserve blame for not saving all the children.

But it is also clear that we cannot blame them individually for the failure to save all, or the failure to help save all. Given the racism of the other two, no one individual was the reason that not all the children were saved, and no individual could have helped save them all. The three seem to be to blame for the outcome collectively, not individually.

As with collective obligations, then, collective responsibility and blameworthiness seem independent of corresponding individual blameworthiness. The other cases mentioned above could illustrate the same phenomena. The failure of the two divorced parents to preserve a reasonable relationship could not be pinned individually on either of them if the recalcitrance of each would have foiled any attempt on part of the other; the failure would still seem to be *their* fault, due to *their* recalcitrance. Likewise, but on grander scales, for the election of corrupt leaders and the failure to significantly reduce the risk of global climate catastrophe: no one who voted for the leadership can be individually blamed for the outcome, and no one—at least no ordinary citizen—who understood the risk for climate catastrophe but failed to take significant measures can be blamed for the current grave risk. But *together* they can be. (See e.g. Björnsson 2011; Held 1970; Kutz 2000; May 1992; Miller 2006; Sadler 2006; Sartorio 2004; Sverdlik 1987.)

Or so it seems. In the remainder of the chapter, we look at how these natural attributions of collective obligations and collective blameworthiness give rise to philosophical puzzles, and at various attempts solve these.

### **Collective obligations: agency worries**

The central problem with the idea of collective obligations is that groups of the sort we have looked at rarely qualify as *moral agents* in any strict sense. This is a problem if we assume that:

AGENCY REQUIREMENT: Only moral agents have obligations.

For an individual to be subject to moral obligations, it seems, it needs to have some amount of self-control, including some control over its goals and some basic capacity for moral cognition, and be guided by its beliefs towards achieving its goals. Many think that some suitably organized groups, such as institutional groups, corporations, and states, satisfy this requirement (see e.g. Björnsson and Hess 2017; Copp 2007; French 1984; Hess 2014; List and Pettit 2011; cf. Tollefsen 2015). But the trio of adults on the beach, the divorced couple, the group of eligible voters, and the group of people aware of the risk of climate catastrophe seem to lack the relevant sort of organization. Though their members have beliefs, goals, self-control, and capacities for moral cognition, it is not clear how the groups themselves can be said to have any of these. (For an early statement of this problem, see Benjamin 1976; also see Collins 2013; Isaacs 2011; Lawford-Smith 2015.)

It is true, of course, that the groups that we have looked at are capable of the relevant states, or at least capable of acting together. Their members are moral agents that can act together with other agents in ways that involve shared goals, perspectives, and decision procedures, all possibly informed by the moral cognition of the members. For example, if the three adults in *Offshore wind* want to, they will quickly organize to save all the children based on coordinated assumptions about the situation and a shared plan. For some groups, such as the group of people alive today and currently aware of the risk of climate catastrophe, such organization might seem farfetched. In principle, however, each member of that group could operate on the assumption that other members take the risk to be real and morally very significant, take ways of mitigating the risk to carry substantial weight in deliberation, and act on this while seeing themselves as part of a global effort to minimize that risk, even if that effort would be wildly distributed.

That groups have these capacities suggests two possible ways of responding to the AGENCY REQUIREMENT, each represented in the literature. The first argues that the capacity is enough to make a group subject to moral obligations. The second instead understands collective obligations to

$\phi$  as grounded in members' individual obligations, obligations the discharging of which would (likely) lead the group to  $\phi$ . The challenge for this response is to identify the right individual obligations: we have already seen that they cannot merely be obligations to  $\phi$  or participate in  $\phi$ -ing. I discuss these responses in turn.

### **The capacity response to agency worries**

The first suggestion, pursued by Bill Wringe (2010), is that the AGENCY REQUIREMENT is too strong and that one can be the subject of moral obligations merely by having the capacity for the sorts of properties associated with moral agency (cf. Shockley 2007 and Aas (2015), though Aas (2015: 18) takes demands on groups to be entirely dependent on moral demands on members). This *capacity response* has some initial plausibility. Central requirements for moral agency, such as self-control and reasons-responsiveness, are themselves at least in part capacity requirements, and so is the most obvious precondition for the existence of an all-things-considered obligation: that the bearer of the obligation is able to discharge it. But the suggestion faces difficult problems.

The general problem is to explain why properties that are usually taken to be necessary for moral agency and moral obligations are not in fact necessary.

On certain views about the source of moral obligations, the problem is particularly stark. Consider contractualism. On standard contractualist views, agents are subject to moral norms in virtue of being parties of a hypothetical social contract. But the group of adults who happen to be on the same beach is not itself plausibly such a party in addition to its three members. Or take versions of constitutivism according to which duties have their source in constitutive aims of moral agents: the aim of rational self-determination, say. It is not obvious how the kinds of groups that we have looked at could have constitutive aims of that sort.

Of course, some will reject both contractualism and constitutivism. But the problem is raised directly by a more uncontroversial requirement on moral agency: that of self-control. It is unclear how unstructured groups can exert such control when they lack anything corresponding to a controlling self, as these are commonly understood in the literature: a practical self-conception, or a set of values, principles, or higher-order plans or preferences. The problem is also raised by standard epistemic requirements on particular obligations. Generally speaking, an agent has no obligation to respond to morally relevant factors when, for no fault of her own, she lacks any beliefs about these factors. If the sort of unstructured groups that we have looked at lack beliefs of their own, it is thus unclear how they can have any obligations. It is true, of course, that each of the three adults in *Offshore wind* has beliefs about how all the children could be saved. But because, apparently for no fault of its own, the group is not yet organized such that individual beliefs are tied to the group's decision making and deciding in any systematic way, it is not clear how these individual beliefs can be beliefs of the group.

### **Grounding collective obligations in demands on individuals**

The problems for the capacity response all suggest that it is the moral agency of the *members* of non-agential groups that make it seem plausible that such groups have obligations. It is because each of the three adults is a bona fide moral agent and knows what is going on in *Offshore wind* that the trio has an obligation to save all the children. And this in turn suggests that group obligations depend

on or should be understood in terms of moral demands directed at the members, in some way or other. To further see this dependence, contrast *Offshore wind* with

*Uncertain success*: Like *Offshore wind*, but the boat looks leaky and difficult to handle and the adults are unsure whether an attempt to use it to save all would be successful. If they tried, they could find themselves struggling with a useless sinking boat, having wasted their chance to save their own children while they were within swimming distance. (In fact, a joint attempt to save all would have been difficult but ultimately successful.)

Assume that the *pro tanto* agent-relative duty of each parent to save her own child outweighs the possibility of saving all, such that a parent's participation in a joint attempt would be merely permissible, not required. Unlike in *Offshore wind*, it now seems false that the group has an obligation to save all the children. Based on this contrast, it seems that if it cannot be morally demanded of the members of the group that they react and act in ways that would result in a group's action, then that action cannot be demanded of the group.<sup>1</sup> Moreover, it is not clear how the agent-relative duties of the parents can provide the group, understood as an independent duty bearer, with reasons outweighing the prospect of saving many more of the children. This suggests that THREE'S OBLIGATION is grounded in individual obligations which are in turn grounded in the moral issue at hand and the group's capacity to handle that issue through some combination of individual actions.

This grounding suggestion conflicts with Wringer's (2016) claim that, phenomenologically, individual obligations in cases like *Offshore wind* depend on the collective obligation. But it allows that concrete individual obligations to contribute to a collective outcome are dependent on various collective matters. In *Offshore wind*, for example, individual obligations to act and react in various ways in light of what others do depend on the fact that it is important that two or more adults team up to save the child, which in turn depends on the fact that if two or more adults wanted to save all the children, they would; in other cases, individual obligations depend on such things as the members' implication in a collective harm or the importance of a joint promise being kept. It is not clear that phenomenology requires more dependence than this.

If a group's obligation to  $\phi$  are grounded in obligations of individuals, we need answers to these two questions:

What are the relevant individual obligations?

How must they be related to the group's  $\phi$ -ing?

Regarding the latter, one might think that if the individuals discharge their obligations, this should make it *sufficiently likely* that they  $\phi$  (Aas 2015; Collins 2013). Alternatively, one might think that in order for the group to have an obligation to  $\phi$  rather than merely an obligation to *try* to  $\phi$ , the satisfaction of their individual obligations must *ensure*, under the circumstances, that they  $\phi$  (Björnsson 2014; Forthcoming; cf. Pinkert 2014).

---

<sup>1</sup> This claim might seem problematic for what we might call "institutional" groups, such as committees or task forces created to fill a particular role in an organization or in society (cf. Copp 2007). But insofar as it is plausible that these are under moral obligations proper rather than merely subject to morally grounded institutional norms, I take this to be because they come close to satisfying standard requirements of moral agency.

But most work has been focused on the former question, regarding what individual obligations ground group obligations. Since, as we saw in section 1, individual obligations to actually participate in or help bringing about  $\phi$ -ing are dependent on what other members will do, the relevant individual obligations will have to be farther removed from contributions to actual  $\phi$ -ing. It has thus been suggested that a group's obligation to  $\phi$  is grounded in the fact that the group would  $\phi$ , or would be sufficiently likely to  $\phi$ , if members discharged their individual obligations to

*take steps to collectivize*; to transform the group into a group agent that has as its own obligation to  $\phi$  (Collins 2013; cf. Isaacs 2011: 144–54 on “putative obligations”; Hindriks 2019; Collins 2019),<sup>2</sup>

*we-reason*; to identify  $\phi$ -ing as the optimal solution to a problem that group members cannot solve individually and to deduce their own individual actions based on this (Schwenkenbecher 2018; 2019),<sup>3</sup>

*be prepared* to do their part in  $\phi$ -ing should they be sufficiently certain that others would as well (Aas 2015),<sup>4</sup>

*care* to the right extent about what is morally at stake, in the sense of being disposed to (i) pick up information about what reactions and actions tend to promote what is morally important and (ii) be moved by such information when opportunity arises (Björnsson 2014).<sup>5</sup>

Though these proposals all violate the AGENCY REQUIREMENT, they suggest a principled way of explaining why collective obligations are not subject to this requirement while individual obligations are: it is specifically *basic* obligations—obligations to we-reason or care, say—that require rich forms of agency. Moreover, each proposal seems to get cases like *Offshore wind* right. At least at a first glance, it seems plausible that each of the three adults has an obligation to care about the predicament of the children, to be prepared to collaborate with the others, to think about how they can together best solve the problem at hand, and, assuming that the others discharge these obligations, to take steps to come together with the others in some constellation or other in a coordinated effort to save all children based on a shared understanding of the situation. And it seems likely that they if they all discharged these duties, they would indeed save all the children. Each proposal also seems to capture the fact that what concrete actions can be demanded of the

---

<sup>2</sup> I'm ignoring a complication of Collins' view here. Though she suggests that people can be charitably interpreted as conveying the sort of fact listed here when they attribute obligations to non-agential groups, she also denies that the group is, strictly speaking, an obligation bearer.

<sup>3</sup> Strictly speaking, Schwenkenbecher talks of non-overridden reasons to we-reason, rather than obligations to do so, and restricts her account to cases where the problem cannot be solved by a single individual.

<sup>4</sup> In a footnote, Aas (2015: 6n13) says that the preparedness might come to nothing more than good character. So understood, the individual obligations he appeals to strongly resemble the caring proposal of Björnsson 2014. But in other places (pp. 13, 16, 19, 21), Aas spells out the preparedness in terms of conditional intentions. Moreover, it is central to his argument that if each individual is prepared to do their part in  $\phi$ -ing if others do theirs, this counts as the group's *trying* to  $\phi$  (p. 16), and this looks very implausible on a merely dispositional understanding of “preparedness”.

<sup>5</sup> Strictly speaking, the proposal is that group obligations as well as individual obligations are grounded in moral demands that individuals care. Whether these demands are best thought of as obligations is a further question.

individuals will depend on what the others do, as well as the fact that the obligation to save all disappears in *Uncertain success*.

As we shall see, however, the first three of these proposals fail to account for at least two kinds of intuitively compelling cases of group obligations. In the first, members of a group are obviously uninterested in helping, rendering futile any individual attempt to bring about a collective effort as well as any preparedness for or deliberation about such an effort. Consider a version of *Helplessness* where

*Visible helplessness*: It is true of each adult that were she to consider ways of saving all the children, she would immediately see that the other two would not be interested in a joint effort and that any sufficient effort to convince them would risk her own child.

The additional features of *Visible helplessness* seem to do nothing to undermine THREE'S OBLIGATION: together they can save all, and they would if they all wanted to. This is in line with the *caring* proposal, as it seems plausible that the individuals can be required to care about survival of all the children in the relevant dispositional sense, and that this would ensure that they saved all. But contrary to the collectivization, we-reasoning, and preparedness proposals, the individuals have no obligations to we-reason, take steps to collectivize, or be prepared to act with the others should they be willing to contribute.

In the second kind of problem case, the discharging of a collective obligation requires no collaborative coordination in relation to a joint goal. In discharging paradigmatic individual duties, such as duties not to steal, lie, or kill, we rarely have this as a goal around which we coordinate our actions: for most of us, the idea of stealing, lying, or killing never comes up. The same seems true about various group obligations. The divorced parents might discharge their obligation to preserve a reasonable relationship by being sensible and respectful without even considering the possibility that their relationship would turn sour. Similarly for

*Ferry ride*: A small ferry would capsize if many of its passengers moved laterally in sync. On this ride, as on most, almost all passengers are sitting quietly after a hard day's work.

Even if the possibility is too remote to consider, they could easily capsize the ferry if they wanted to. In light of this, it is plausibly their obligation not to. This is very much in line with the *caring* proposal, as it seems both (i) that it can be demanded of the passengers that they care about what is bad about capsizing, and (ii) that if the passengers do care, that this would ensure that they don't capsize the ferry (cf. Björnsson 2014; Forthcoming). But the group's obligation not to capsize the ferry conflicts with the *preparedness*, *we-reasoning*, and *collectivization* proposals. Absent any present danger, the passengers have no obligations to try to form a group agent with all the other passengers, or to think about how the group is best to avoid capsizing the ferry, or even to be prepared to join the others to prevent it from capsizing.

In light of *Visible helplessness* and *Ferry ride*, it thus seems that any understanding of a collective obligation to  $\phi$  as grounded in individual obligations would require us to understand the latter as concerned with dispositions of some sort to contribute towards  $\phi$ -ing, rather than as some overt action, mental activity, or even active state of preparedness.



### What makes these things *obligations*?

Even if an account of this sort captures our intuitive attributions of group obligations, it leaves the question of why the fact that a group would  $\phi$  if members cared appropriately would constitute an *obligation* on part of the group to  $\phi$ . This question is particularly pressing as all the proposals in this family seems to suggest that what looks like a straightforward attribution of *an obligation to  $\phi$  to a group* is in fact something quite different and much more complex, such as an attribution of *individual obligations to members of the group to have a certain disposition such that if all members had that disposition they would  $\phi$* .

Elsewhere, I have suggested that the relation postulated in the *caring* proposal constitutes obligation because it is the very same relation that constitutes individual obligation, i.e. an obligation of a group of only one individual. If this is correct, there are no hidden structural differences other than the number of obligation bearers. To give a sense of this sort of reply, I will briefly recount the proposal in question and explain how it applies to the individual case, and why the relation constitutes obligation. This is the proposal:

OBLIGATION: A group has a moral obligation to  $\phi$  if and only if the group's  $\phi$ -ing (i) is morally important and (ii) would be ensured, in a normal way, if members cared as can be morally demanded of them (Björnsson 2014; Forthcoming; cf. Björnsson and Brülde 2017).

As before, the concern here is with all-things-considered, not pro tanto obligations. The *moral importance* of the group's  $\phi$ -ing can be instrumental or non-instrumental. On some views, the only relevant moral importance is the importance of not acting on certain kinds of motives, but most acknowledge the moral importance of bringing about or not preventing certain kinds of actions and outcomes, regardless of specific motives. Substantive accounts of obligations can differ with respect to what is morally important and what makes it so, but also with respect to what it can be *morally demanded* that people care about, and to what extent. On some narrow Kantian views, morality only requires that we care about acting from maxims that can be willed as universal law, and on some narrow utilitarian views, it only requires that one cares about the surplus of pleasure over pain. But most will acknowledge demands to care about a variety of values, including staples of both deontological and consequentialist moral theories: telling the truth, keeping promises, not using others as mere means, human flourishing, happiness, relationships, knowledge, and so forth. Because OBLIGATION leaves open what can be morally demanded of members of groups, it is compatible with a variety of normative views, and because it leaves open what ultimately grounds such demands, it is also compatible with a variety of views about their nature and ultimate sources, including intuitionism, rationalism, contractualism, and constitutivism.<sup>6</sup>

---

<sup>6</sup> While allowing for a variety of normative views, the concept captured by OBLIGATION is not the only one that we might be concerned with. It is not a concept of objective, information-independent obligations of the sort that concern at least some normative theorists, but rather a concept of what we are required to do in light of available evidence. More generally, it makes agents' morally substandard caring the only thing that can prevent them from discharging their obligations, and is thus not the one we have in mind when thinking that a failure to discharge an obligation is morally innocent, due to excusing factors. I focus on this particular concept because I take it to be both central to everyday thinking about obligations and relatively simple. For its relation to some other notions of obligation, see Björnsson Forthcoming.

Implicitly, OBLIGATION accounts for the common idea that an obligation to  $\phi$  presupposes an *ability* to  $\phi$ : if all members of the group care about what makes  $\phi$ -ing bad, this would ensure, in a normal way, that they would  $\phi$ .<sup>7</sup> What, though, is it for sufficient caring to *ensure* an outcome, and to ensure it in a *normal* way?

For X to *ensure* Y is for X to rule out all relevant possibilities alternative to Y, not necessarily for X to cause Y. The lock on the door might ensure that no one enters even if no one tries to enter, and if the passengers on the ferry care about staying afloat, this might ensure that they don't capsize the ferry even if the reason that the ferry remains steady is they are all sitting down to rest after a hard day's work. For obligations that can only be discharged through coordinated action, sufficient caring will often ensure such action by prompting intentions to join forces if others are similarly prepared, and act on the perception of such preparedness among other members. Even here, though, individuals might participate out of habit, group conformity, or economic interest, and morally required caring kick in only if the joint endeavor is threatened.<sup>8</sup>

As noted, OBLIGATION requires that the way in which appropriate caring ensures the group's  $\phi$ -ing is relevantly *normal*: a matter of skill or ability rather than mere luck. This might require that members' appropriate caring would block at least some relevant threats to the group's  $\phi$ -ing by at

---

<sup>7</sup> Aas (2015: 17–19) provides a structurally similar account, but in terms of the group's "trying" to discharge its obligation, without any normality requirement, and only requiring that such trying should make the outcome sufficiently likely. Lawford-Smith 2015 and Hindriks 2019 deny that non-agential groups can have abilities required for irreducible group obligations, but they don't consider the sort of proposal offered here. For other discussions of group abilities, see e.g. Pinkert 2014 and Collins 2019: 74–85.

<sup>8</sup> Just as circumstances might be such that the appropriate caring of all three adults ensures that all children are saved, they can be such that *the appropriate caring of all moral agents in the world* ensures the outcome, because it ensures that the three adults on the beach care. But, intuitively, the group of every moral agent in the world does not plausibly have an obligation to save all the children. Why is that? According to OBLIGATION, something can be a group's obligation only if it is something that the group would do if caring appropriately. In this case, saving the children is not something that the group of every moral agent in the world would have done if all had cared appropriately. Only the people on the beach would have been involved.

What constrains attributions of  $\phi$ -ing to groups? A rough rule seems to be that acceptable attributions pick out groups whose members would be relevantly involved in producing or allowing for the activity or outcome constitutive of  $\phi$ -ing if they all cared appropriately. (Doing actual work might not be required if one is relevantly intentionally related to the  $\phi$ -ing: in *Plenitude*, it might make sense to say that the adults saved all the children even if one of them stands by letting the other two do the actual work. This is not to deny that differences in kind or degree of participation can matter. It will make sense to say of the two adults who did the actual work that they saved the children, but would be at least misleading to say of a pair consisting of one of those two and the adult standing by that they did.) Moreover, attributions might be more coarse-grained if a precise identification of this group is inconvenient. For example, if the saving effort involved a large number of people on the beach, it might be natural to say that "the people on the beach" saved those adrift even if a few individuals on the beach did not participate due to bad motives or lack of capacity.

It often makes sense to attribute obligations to groups that include only some of those that would be involved if everyone cared appropriately. In a version of *Offshore wind* where one of the adults refuses to help save all, it might make good sense to say not only that the three adults have an obligation to save all, but also that the two willing adults have such an obligation: they would do so if they cared appropriately. For *Plenitude*, such a restriction would at least be misleading, as one of the two might end up standing idly by while the other does all the work together with the third adult. (I thank Saba Bazargan-Forward and Deborah Tollefsen for pressing me on these issues.)

least marginally *coordinated* behavior, rather than by individual behavior completely insensitive to that of other members. It also rules out cases where group actions would ensure  $\phi$ -ing only in ways deviating significantly from the intended:

*Serendipity*: The adults' only chance of saving all the drowning children is by using a lifeboat mounted on a support structure. But an arm holding the boat in place is rusted stuck and it would seem that the only way of releasing it is by hitting that arm with an iron rod lying nearby. As it happens, doing so would fail to release the lifeboat, but would make a loud sound that would catch the attention of a coast guard boat that is anchored behind a rock and would come and rescue the children.

Suppose that it is the group's duty to try to save the child, and to try to get the lifeboat unstuck using the iron rod, and that these are the group's duties because of the importance of the child's survival. Still, if the only way that they would bring about the child's survival would be by accident, they do not plausibly have an obligation to save the child in this case. By contrast, if they had known about the nearby coast guard boat and would have employed the iron rod to catch their attention, it would have been their obligation to save the child by doing so.

A final aspect of OBLIGATION worth mentioning is the connection between obligations and blameworthiness. If a group fails to discharge an obligation, it must be because one or more members' caring fell short at some point, at least. But if something morally bad happens as a result of people not caring about the values at stake as can be demanded of them, they are plausibly morally to blame for it: at least that seemed to be the case in *Helplessness*. (We return to the issue of collective blameworthiness in the next section.)

The claim now is that OBLIGATION is an equally plausible account of obligations of individuals, understood as "groups" of one agent. The demands on caring are the same, and what is of moral importance is the same. What is different is merely that it is the appropriate caring of a single agent that ensures the  $\phi$ -ing of that one agent. Moreover, the account applies as straightforwardly to the case where one agent can either save her own child or use a life boat to save all children, to individual obligations to  $\phi$  that are discharged without the aim of  $\phi$ -ing, and to a one-person version of *Serendipity*. But if OBLIGATION is an adequate account of both individual and group obligations, it is clear why group obligations to  $\phi$  are *obligations*: they are constituted by the very same relation between the obligation bearer and the  $\phi$ -ing as are individual obligations.

### **Collective responsibility: agency worries**

Turn now from the issue of collective obligation to that of collective responsibility or blameworthiness. THREE'S BLAMEWORTHINESS—the idea that the three adults are to blame for not saving all the children in *Helplessness*—can seem problematic in much the same way as did THREE'S OBLIGATION.

On the one hand, we cannot plausibly understand THREE'S BLAMEWORTHINESS distributively, as attributing individual blameworthiness for not saving all the children to each of the three adults. No individual in *Helplessness* can be individually blamed for not saving all the children, or for the fact that the trio did not save all the children: the individual had no say in that matter given that the other two individuals were unwilling to participate. (This is not to deny that they might be to individually to blame for other things, such as lack of concern for the other children, or failure to think about how all might be saved.)

On the other hand, the group itself seems to lack the sort of properties that are required for individual responsibility and blameworthiness. If an entity cannot be blameworthy without violating an obligation, and if only moral agents can be subject to obligations, it follows immediately that non-agential groups cannot be blameworthy. Things are a little less straightforward if one can be to blame even without strictly speaking violating any obligations. (An example might be not giving up one's seat on the bus to someone more need of it. For discussion, see Driver 1992.) Still, the underlying worries from section 2 would remain, as blameworthiness is standardly assumed to require both some degree of self-control and beliefs about morally relevant features of the situation. As already noted, the group seems to lack anything corresponding to the self that can exert relevant control over its actions as well as anything corresponding to beliefs. Furthermore, neither the group nor its members can plausibly be blamed for that lack of a self or lack of belief: prior to the unexpected offshore wind, there was no reason to form a collective agent.

### **Collective blameworthiness as *shared* blameworthiness**

Our discussion of collective obligations suggested that these are grounded in demands on individuals. Likewise, collective blameworthiness seems to depend on criticizable input from group members. As in the case of obligations, we can contrast *Helplessness* with *Uncertain success*. In the former case, where the group of adults seems to deserve blame for not saving all the children, the failure to save all the children was due to the adults' callous, racist attitudes. In the latter, where the group does not seem to blame for the failure to save all, the outcome was not due to any criticizable attitude on their part, as evidence suggested to each of them that the prospects for saving all children were bleak. Here, even morally conscientious individuals could be expected to save their own children instead.

If group blameworthiness is grounded in the moral shortcomings of members, that could straightforwardly explain why the group does not have to satisfy standard requirements on blameworthiness for individual agents, again in line with what we said about collective obligations. It can also explain why individuals that are part of blameworthy non-agential groups are individually *implicated* in the collective blameworthiness. When the group is blamed for their failure to save all the children in *Helplessness*, an individual racist cannot plausibly reply, "Sure, many of the kids died because *we* didn't want to save them, but what does that have to do with *me*?" And at least partly because group blame seems to implicate members and relate intimately to individual moral shortcomings of this sort, it is natural to think, as many have, that there is such a thing as *shared* or *joint* responsibility for group actions and outcomes (e.g. Björnsson 2011; Kutz 2000; May 1992; Miller 2006; Sadler 2006; Sartorio 2004; Sverdlik 1987).

In virtue of what, though, do individuals in a group share responsibility for some action or omission of the group, or something that it has brought about or let happen? On some influential suggestions, one can share responsibility for a harm in virtue of intentional or knowing participation in group activity that stands in some relevant causal relation to that harm even if one does not stand in that relation oneself (Kutz 2000; Miller 2006; Sadler 2006; Sverdlik 1987), or in virtue of participation in a culture or way of life giving rise to the harm (Kutz 2000; Silver 2002), or in virtue of endorsing the sorts of actions or omission that directly explained the outcome (May 1992). Suggestions like these seem to make sense of some central cases of collective responsibility, explaining how participation in or endorsement of the machinery of an unjust war might make one partly responsible for its resulting atrocities, even though one could have done nothing to prevent

them. But as illustrated by *Helplessness*, it seems that a group of individuals can be to blame, together, without intentional or knowing participation in anything meaningfully called a group activity, and without the harm being the upshot of a culture or way of life. Moreover, since no action or omission by any one agent directly explained or caused the outcome, other agents cannot be responsible for that outcome in virtue of endorsing or engaging in actions or omissions of the type as that which explained it. It is clear, then, that even if these suggestions identify relations in virtue of which one might share blameworthiness, they fail to provide completely general accounts of collective responsibility (Björnsson 2011).

To capture the cases that are left unaccounted for, it is natural to generalize: perhaps one can share blameworthiness for something by being to blame for part of a cause of it, regardless of relations to other agents. In *Helplessness*, for example, it might seem that each adult is to blame for part of the fact that not enough able adults were willing to save all children and looking to see if the others were too. Similarly, each person voting for the corrupt leadership might be to blame for part of the voting result that brought them to power, and each of the divorced parents might be to blame for part of the lack of cooperativeness that caused the deterioration of their relationship (cf. Sartorio 2004).

But mere blameworthiness for part of a cause does not seem sufficient. Consider:

*Partly excused helplessness:* Like *Helplessness*, but only one of the adults is racist; the other two are unwilling to help because each thinks, based on very convincing but misleading evidence, that the boat is leaky and difficult to handle and that an attempt to rescue all would risk the life of their own child, who is among those farthest from the beach.

Here, like in *Helplessness*, the group fails to save all children because not enough adults are willing to attempt the rescue. Moreover, the racist seems as much to blame for her part in this fact as she was in *Helplessness*. However, we cannot here blame the trio for the failure to save all: two of them have perfectly good excuses for being unwilling to try. And although we can criticize the racist for not caring enough about some of the children and blame her for not trying to get the others to help, we cannot blame her for the failure to help save all, as she would have been unable to get the others to join. Judging from the contrast between this case and *Helplessness*, then, the way in which she is implicated in the outcome in the latter is *irreducibly collective*, as part of a blameworthy group, not just as responsible for part of a cause of the failure to save all (Björnsson 2011: 185–6).

### **How is shared blameworthiness possible?**

The conclusion of the previous section was that group blameworthiness is irreducibly collective, but not in a way that requires substantive social relations. Even if this is correct, it remains unclear *why* the involvement of other blameworthy agents' in an outcome should matter for how the individual is implicated. The racist has no more control over the outcome in *Helplessness* than in *Partly excused helplessness*, and though it might be tempting to say that *the group* has more control in the former case because all members have access to correct information about the situation, the group seems to lack the agential properties, including self-control and beliefs about its surroundings, that we standardly take to be necessary for responsibility.

My own explanation for this lies in the nature of blameworthiness, whether individual or collective. According to an influential understanding, often employed in the collective responsibility literature (e.g. Feinberg 1968 and May 1992), blameworthiness for something bad does not require

that one was in full control of it, but rather that it happened partly because of some of one's *faults*. Faults are often understood as blameworthy or shameful actions or omissions, but a tradition from P.F. Strawson (1962) understands basic faults as instances of substandard responsiveness to values, or substandard caring. In line with this, I have defended the following account of the blameworthiness of both groups and individuals:

BLAMEWORTHINESS: For a group to be morally to blame for Y is for Y to (i) be morally bad and (ii) be explained in a normal way by the group's members' not caring, in ways that can be morally demanded of them (Björnsson 2011; 2017).

Like any account that takes collective blameworthiness to be grounded in demands on individual moral agents, BLAMEWORTHINESS explains why groups can be to blame even if they do not themselves satisfy standard requirements on moral agency.

But BLAMEWORTHINESS has a number of more specific explanatory virtues: It straightforwardly explains THREE'S BLAMEWORTHINESS in *Helplessness*, as well as the trio's lack of blameworthiness for not saving all the children in *Uncertain success* and the racists lack of blameworthiness for the outcome in *Partly excused helplessness*: only in *Helplessness* was the failure to save all due to the substandard caring of the members of the group.

Importantly, the claim that the failure was due to *their* substandard caring is not distributive: it does not imply that it is true about any one of them that the failure was due to *her* substandard caring. The latter might seem implausible given that her substandard caring made no difference to the outcome. Nevertheless, BLAMEWORTHINESS explains how members are implicated in collective blameworthiness for something: the individual member's substandard caring *is part of* why that thing happened. And it explains why failures to discharge obligations, understood in line with OBLIGATION, imply that at least some agent is thus implicated in blameworthiness for this failure: because obligations are morally important things that are ensured by appropriate caring, failures to discharge an obligation must be both morally bad and due to one or more instances of substandard caring.

In addition to this, BLAMEWORTHINESS is not designed specifically to handle collective blameworthiness: it applies equally to cases of individual blameworthiness, where bad things have happened because one agent failed to care enough on one or more occasions. Because it takes individual and collective blameworthiness to consist in the very same relation, it offers the same kind of benefit as OBLIGATIONS: it explains why collective blameworthiness is *blameworthiness*.

### **Concluding remarks**

Attributions of collective obligations and collective responsibility or blameworthiness to non-agential groups form an important part of moral cognition. But they raise difficult problems of interpretation. What I have hoped to do in this chapter is primarily to make clear what the central problems are, outline some of the phenomena that need to be accounted for, and point to problems and prospects for some existing proposals. In my view, there is most promise in accounts of collective obligations and blameworthiness that understands these as involving the very same relations as their individual counterparts and ground both in demands on individual agents to care about morally important matters. But whether or not such accounts are ultimately successful, I hope that they help to further clarify both the phenomena and the problems they raise.

## Acknowledgments

The paper has greatly benefited from comments on earlier versions at the Moral and Practical Philosophy Seminar at UCSD, the Seminar in Practical Philosophy and Political Theory at the University of Gothenburg, and the Higher Seminar in Practical Philosophy at Stockholm University, as well as from comments by Dick Arneson, Bill Wringe, Anne Schwenkenbecher, Benjamin Matheson, Stephanie Collins, Frank Hindriks, Niels de Haan, David Copp, Olle Blomberg, Saba Bazargan-Forward, and Deborah Tollefsen.

## Bibliography

- Aas, Sean (2015) "Distributing Collective Obligation," *Journal of Ethics and Social Philosophy*, 9 1-23.
- Benjamin, Martin (1976) "Can Moral Responsibility Be Collective and Nondistributive?," *Social Theory and Practice*, 4 93.
- Björnsson, Gunnar (2011) "Joint Responsibility without Individual Control: Applying the Explanation Hypothesis," in Vincent, Nicole, van de Poel, Ibo and van den Hoven, Jeroen (eds) *Moral Responsibility: Beyond Free Will and Determinism*. Dordrecht: Springer pp. 181–99.
- Björnsson, Gunnar (2014) "Essentially Shared Obligations," *Midwest Studies in Philosophy*, 38 103–20.
- Björnsson, Gunnar (2017) "Explaining (Away) the Epistemic Condition on Moral Responsibility," in Robichaud, Philip and Wieland, Jan Willem (eds) *Responsibility: The Epistemic Condition*. New York: Oxford University Press pp. 146–62.
- Björnsson, Gunnar (Forthcoming) "Individual and Shared Obligations: In Defense of the Activist's Perspective," in Budolfson, Mark, McPherson, Tristram and Plunkett, David (eds) *Philosophy and Climate Change*. Oxford University Press.
- Björnsson, Gunnar and Brülde, Bengt (2017) "Normative Responsibilities: Structure and Sources," in Hens, Kristien, Cutas, Daniela and Horstkötter, Dorothee (eds) *Parental Responsibility in the Context of Neuroscience and Genetics*. Cham: Springer International Publishing pp. 13-33.
- Björnsson, Gunnar and Hess, Kendy (2017) "Corporate Crocodile Tears? On the Reactive Attitudes of Corporate Agents," *Philosophy and Phenomenological Research*, 94 273–98.
- Collins, Stephanie (2013) "Collectives' Duties and Collectivization Duties," *Australasian Journal of Philosophy*, 91 231–48.
- Collins, Stephanie (2019) *Group Duties: Their Existence and Their Implications for Individuals*, Oxford: Oxford University Press.
- Copp, David (2007) "The Collective Moral Autonomy Thesis," *Journal of Social Philosophy*, 38 369-88.
- Dietz, Alexander (2016) "What We Together Ought to Do," *Ethics*, 126 955-82.
- Driver, Julia (1992) "The Suberogatory," *Australasian Journal of Philosophy*, 70 286-95.
- Feinberg, Joel (1968) "Collective Responsibility," *The Journal of Philosophy*, 65 674-88.
- French, Peter A. (1984) *Collective and Corporate Responsibility*, New York: Columbia University Press.
- Held, Virginia (1970) "Can a Random Collection of Individuals Be Morally Responsible?," *The Journal of Philosophy*, 67 471-81.
- Hess, Kendy M. (2014) "The Free Will of Corporations (and Other Collectives)," *Philosophical Studies*, 168 241–60.
- Hindriks, Frank 2019: 'The Duty to Join Forces: When Individuals Lack Control'. *The Monist*, 102, pp. 204-20.
- Isaacs, Tracy (2011) *Moral Responsibility in Collective Contexts*, Oxford: Oxford University Press.
- Kutz, Christopher (2000) *Complicity*, Cambridge: Cambridge University Press.

- Lawford-Smith, Holly (2015) "What 'We'?", *Journal of Social Ontology*, 1 225–49.
- List, Christian and Pettit, Philip (2011) *Group Agency: The Possibility, Design, and Status of Corporate Agents*, Oxford: Oxford University Press.
- May, Larry (1992) *Sharing Responsibility*, University of Chicago Press.
- Miller, Seumas (2006) "Collective Moral Responsibility: An Individualist Account," *Midwest Studies in Philosophy*, 30 176-93.
- Pinkert, Felix (2014) "What We Together Can (Be Required to) Do," *Midwest Studies in Philosophy*, 38 187-202.
- Sadler, Brook J. (2006) "Shared Intentions and Shared Responsibility," *Midwest Studies in Philosophy*, 30 115-44.
- Sartorio, Carolina (2004) "How to Be Responsible for Something without Causing It," *Philosophical Perspectives*, 18 315-36.
- Schwenkenbecher, Anne (2014) "Joint Moral Duties," *Midwest Studies in Philosophy*, 38 58-74.
- Schwenkenbecher, Anne (2018) "Making Sense of Collective Moral Obligations: A Comparison of Existing Approaches," in Hess, Kendy, Ignieski, Violetta and Isaacs, Tracy (eds) *Collectivity: Ontology, Ethics, and Social Justice*. Rowman and Littlefield pp. 109-32.
- Schwenkenbecher, Anne (2019) "Collective Moral Obligations: 'We-Reasoning' and the Perspective of the Deliberating Agent," *The Monist*, 102.
- Shockley, Kenneth (2007) "Programming Collective Control," *Journal of Social Philosophy*, 38 442-55.
- Silver, David (2002) "Collective Responsibility and the Ownership of Actions," *Public Affairs Quarterly*, 16 287-304.
- Strawson, Peter F. (1962) "Freedom and Resentment," *Proceedings of the British Academy*, 48 187–211.
- Sverdlik, Steven (1987) "Collective Responsibility," *Philosophical Studies*, 51 61-76.
- Tollefsen, Deborah Perron (2015) *Groups as Agents*, John Wiley & Sons.
- Wringe, Bill (2010) "Global Obligations and the Agency Objection," *Ratio*, 23 217-31.
- Wringe, Bill (2016) "Collective Obligations: Their Existence, Their Explanatory Power, and Their Supervenience on the Obligations of Individuals," *European Journal of Philosophy*, 24 472-97