1 Introduction

Bringing about good outcomes typically makes agents worthy of praise, just as bringing about bad outcomes typically makes agents worthy of blame. Blameworthy agents are obliged to apologize and make amends for the bad outcomes that they bring about. Agents that fulfil their obligations in this regard may become worthy of forgiveness and perhaps even reconciliation with those negatively affected by their actions. The issue whether these patterns of evaluation are as applicable to collectives as they are to individuals is a controversial one. Part of the controversy stems from doubts about the existence of collective agents. While complex synergies may arise between the intentions and actions of individual group members, perhaps collectives as such cannot form intentions and react appropriately to being praised or blamed. The assumption here is that a praiseworthy or blameworthy agent must have causally produced an outcome of value for which they can become the fitting targets of attitudes, most notably the attitudes of praise or blame. The question whether there are additional requirements that determine the extent to which an agent is praiseworthy or blameworthy is a difficult one. It seems plausible to suggest, as Peter Strawson does, that factors pertaining to the agent's quality of will are also relevant in this context (Peter Strawson, “Freedom and resentment”, Proceedings of the British Academy, vol.48 (1962), pp. 1–25). For other potential conditions involving the agent’s epistemic profile and awareness of value, respectively, see, e.g., Virginia Held, “Can a random collection of individuals be responsible?”, Journal of Philosophy, vol.67, no.14 (1970), pp. 471–481; Philip Pettit, “Responsibility incorporated”, Ethics, vol.117, no.2 (2007), pp. 171–201. Finally, it should be noted that some controversies must be set aside here regarding the issue whether moral responsibility pertains to outcomes or actions. For a discussion of the outcome conception and its flaws, see, e.g., Thomas Nagel, “Moral luck”, Proceedings of the Aristotelian Society, Supplementary vol.50 (1976), pp. 137–51; Steven Sverdlik, “Collective responsibility”, Philosophical Studies, vol. 51, no. 1 (1987), pp. 67–76; David Enoch & Andrei Marmor, “The case against moral luck”, Law and Philosophy, vol. 26, no. 4 (2007), pp. 405–36.
blamed. That collectives should be so capable is arguably a requirement for them to be agents and thus eligible for assignments of moral responsibility in the first place.\(^2\)

The practice of holding collectives responsible by blaming and praising them is also made controversial by its perceived risk of unfairly affecting individual group members. While the unfairness objection is rarely articulated very clearly in the philosophical literature, expressions of its underlying worries are manifold.\(^3\) One such worry is that holding collectives morally responsible by blaming them harms their individual group members. This could be considered very unfair, specifically in cases where the bad outcomes produced by the collectives are of a sort to lie beyond the intentional control of any of their individual group members. If there are reasons

\(^{2}\) The philosophers that question whether collectives as such can display moral agency in the sense of intending and acting include Max Weber, “Economy and society vol. 1” (Berkeley: University of California Press, 1914); Hywel David Lewis, “Collective responsibility”, *Philosophy*, vol. 23, no. 84, (1948), pp. 3–18; John W. N. Watkins, “Methodological individualism & social tendencies”, *British Journal for the Philosophy of Science*, vol. 8, pp. 104–117; Robin S. Downie, “Collective responsibility”, *Philosophy*, vol. 44, no. 167 (1969), pp. 66–69; Manuel G. Velasquez, “Why corporations are not responsible for anything they do”, in Larry May & Stacey Hoffman, eds., *Collective Responsibility: Five Decades of Debate in Theoretical and Applied Ethics* (Rowman & Littlefield publishers inc, 1992); Jan Narveson, “Collective responsibility”, *Journal of Ethics*, vol. 6, no. 2 (2002), pp. 179–198. While most philosophers agree that collectives do not have minds in the same sense that individuals have them, many maintain that there are derivative senses in which collectives can act and form intentions. For discussions, see e.g. Margaret Gilbert, *On Social Facts* (New York: Routledge, 1989), *Sociality and Responsibility* (Lanham, MD: Rowman and Littlefield, 2000), “Who’s to blame? Collective moral responsibility and its implications for group members”, *Midwest Studies in Philosophy*, vol. 30 (2006), pp. 94–114, *Joint Commitment* (Oxford: Oxford University Press, 2013); Raimo Tuomela, “Actions by collectives”, *Philosophical Perspectives*, vol. 3 (1989), pp. 471–496, “We-intentions revisited”, *Philosophical Studies*, vol. 125 (2005), pp. 327–269, “Joint intention, we-mode and I-mode”, *Midwest Studies in Philosophy*, XXX (2006), pp. 35–58, *Social Ontology* (New York: Oxford University Press, 2013); Marion Smiley, *Moral Responsibility and the Boundaries of Community* (Chicago: University of Chicago Press, 1992); David Velleman, “How to share an intention”, *Philosophy and Phenomenological Research*, vol. 57, no.1 (1997), pp. 29–50; Michael Bratman, *Faces of Intention: Selected Essays on Intention and Agency* (New York: Cambridge University Press, 1999), *Shared Agency: A Planning Theory of Acting Together* (Oxford: Oxford University Press, 2013); Christopher Kutz, *Complicity: Ethics and Law for a Collective age* (Cambridge: Cambridge University Press, 2007); Neta C. Crawford, “Individual and collective moral responsibility for systematic military atrocity”, *Journal of Political Philosophy*, vol. 15, no. 2 (2007), pp. 187–212; Kenneth Shockley, “Programming collective control”, *Journal of Social Philosophy*, vol. 36 (2007), pp. 442–445; David Sosa, “What is it like to be a group?”, *Social Philosophy and Policy*, vol. 26, no. 1 (2009), pp. 212–226; Christian List & Philip Pettit, *Group Agency: The POSSIBILITY, Design, and Status of Corporate Agents* (Oxford: Oxford University Press, 2011); Raimo Tuomela & Pekka Mäkelä, “Group agents and their responsibility”, *The Journal of Ethics*, vol. 20, no. 1–3 (2016), pp. 299–316. Note also that while it is controversial whether derivative senses of collective intentionality are enough for moral agency and moral responsibility as traditionally understood, it is also controversial whether traditional understandings of these concepts are the only ones that should matter in collective cases in the first place.

\(^{3}\) For some variations on this theme, see e.g. Lewis, op. cit., pp. 3–6; Steven Sverdlik, op. cit., p. 68; Velasquez, op. cit., p. 124; Cassie Striblen, “Guilt, shame, and shared responsibility”, *Journal of Social Philosophy*, vol. 38, no. 3 (2007), p. 473; Govert den Hartog, “Comment: collective criminal responsibility: Unfair or redundant”, *Netherlands Journal of Legal Philosophy*, vol. 38, no. 2 (2009), p. 119; Mark Reiff, “Terrorism, retribution, and collective responsibility”, *Social Theory and Practice*, vol. 28, no .3 (2008), p. 234; Thomas H. Smith, “Non-distributive blameworthiness”, *Proceedings of the Aristotelian Society*, vol. 109, No. 1pt1 (2009), p. 55; Björn Petersson, “Co-responsibility and causal involvement”, *Philosophia*, vol. 41, no.3 (2013), p. 850; Johannes Himmelreich & Holly Lawford-Smith, “Punishing groups: When external justice takes priority over internal justice”, *The Monist*, vol. 102, no. 2 (2019), pp. 134–150
to suspect that the practice of collective responsibility will often end up being unfair in this way, then this might recommend a rejection of the practice altogether—or so the argument goes.

The following paper distinguishes between different versions of the unfairness objection and argues that the worries that crop up for the practice of collective responsibility will also be relevant to the practice of individual responsibility. It is therefore unclear what reasons there are to think that there is something uniquely morally problematic about the former practice but not the latter. To make the point, a distinction is also drawn between two kinds of fittingness conditions: On the one hand, there are conditions that determine whether the attitudes of blame and praise are fitting in respect of a given target; On the other hand, there are conditions that determine whether and by what means it is fitting to express attitudes. The distinction is generalized and is shown, along with several other theoretical considerations, to undermine the different versions of the unfairness objection.

2 Holding and Expressing Attitudes in Individual Cases

Individuals that are responsible for good outcomes are typically praiseworthy (it may be fitting to praise them) for those outcomes and individuals that are responsible for bad outcomes are typically blameworthy (it may be fitting to blame them) for those outcomes.\(^4\) To hold individuals responsible for good outcomes is to praise them for those outcomes and to hold individuals responsible for bad outcomes is to blame them for those outcomes. The conditions under which individuals are responsible are often entwined with the conditions under which it is fitting to hold them responsible. Nevertheless, it is possible to hold individuals responsible for outcomes without overt expressions of praise or blame, as attitudes that are held need not be faithfully expressed.\(^5\) The conditions under which it is fitting to hold individuals responsible can therefore be distinguished from the conditions that determine whether and by what means expressions of attitudes are fitting.

The distinction between the different conditions that pertain to attitudes and their expression is often alluded to in the philosophical literature. Within the context of discussions about moral responsibility, it is usually referred to in terms of public and private blame. Sarah Gokhale is among the philosophers to develop the distinction and to emphasize one of the ideas referred to in what follows, namely that there

---

\(^4\) The fitting-attitudes view just described represents one of the most common interpretations of praiseworthiness and blameworthiness. While admittedly controversial, it is taken for granted in what follows. For a more general discussion of fitting-attitudes views of value and their flaws, see Włodek Rabinowicz & Toni Rønnow-Rasmussen, “The strike of the demon: on fitting pro-attitudes and value”, *Ethics*, vol. 114, no. 3 (2004), pp. 391–423.

\(^5\) The expression of attitudes is here understood very narrowly, as constitutive of a certain kind of speech-act that must involve a communicative intent on part of the expressor. This may seem narrow to those who think that the expression of attitudes can be carried out in a multitude of ways, not all of which are inherently linguistic. For example, the intuition might be that attraction can be expressed through flirtatious body language, amusement can be expressed through intense laughter, and joy can be expressed through a spontaneous jig.
tend to be higher stakes associated with public blame than there are with its private counterpart.\(^6\) It should also be noted that while the present discussion is explicitly couched in terms of *fittingness*, Gokhale and other philosophers writing about public and private blame often invoke other normative notions. It should therefore be noted that the arguments offered here work just as well when formulated in terms of the fittingness of attitudes and, say, the *rightness*, *propriety*, or *correctness* of their expression.

The important point to notice is that there might well be situations where an individual is responsible for a good outcome (making the individual a fitting target of praise for those outcomes) even though it is unfitting to express praise toward the individual. Similarly, there may be situations where an individual is responsible for bad outcomes (making the individual fitting targets of blame for those outcomes) even though it would be unfitting to express blame toward the individual. Such cases may be rare, but they are not difficult to imagine. Suppose Kate thinks too highly of herself and is at risk of becoming a narcissist were she to become the knowing recipient of any more praise. It can be imagined that if Kate becomes a narcissist, then this will significantly decrease her chances of producing good outcomes in the future. Kate might therefore be the fitting target of praise for the good outcomes that she produces, even though overt expressions of praise toward her are ultimately unfitting.\(^7\)

The case just mentioned illustrates that overt praise of a praiseworthy individual can benefit that individual in ways that she does not deserve and that leads to bad outcomes. While such considerations may affect the conditions that determine whether and by what means expressions of attitudes are fitting, they need not affect the conditions that determine whether an individual like Kate is a fitting target of praise. This can also be illustrated by less exotic examples, whereby the overt praise of a praiseworthy individual benefits the friends and family of that individual in undeserved ways. When Kate is overtly praised by a large group of people for the good outcomes she produces, her friends and family may start to enjoy social benefits which they have not earned or of which they have not made themselves deserving. If the benefits are significant, this may also have relevance for the issue whether and by what means expressions of praise will be fitting.

Consider next Laura who is suffering from severe depression risks harming herself were she to become a knowing recipient of any more blame. It can be imagined that the bad outcomes produced by Laura are minor ones and of a sort to affect

---

\(^6\) See Sarah Gokhale, *Who Needs Blame?: Answerability Without Expressed Blame* (CUNY Academic Works, 2019). For a discussion concerning private and public blame, see also Michael McKenna, *Conversation & Responsibility* (New York: Oxford University Press, 2012), “Quality of will, private blame and conversation: Reply to Driver, Shoemaker, and Vargas”, *Criminal Law and Philosophy*, vol. 10, no. 2 (2016), pp. 243–263; Julia Driver, “Private blame”, *Criminal Law and Philosophy*, vol. 10, no. 2 (2016), pp. 215–220.

\(^7\) Admittedly, while the idea of private blame seems familiar to our everyday thoughts and practices, the idea of private praise sounds somewhat strange. This need not be a problem. Whenever private praise is mentioned in what follows, this could be taken to refer to whatever attitudes are typically expressed through the public praise of agents. Presumably, private praise is constituted by a certain category of admiration or respect, perhaps in connection with judgments about the relevant outcomes having been realized by the target of said attitudes.
her more than anyone else. This seems like a case where Laura is the fitting target of blame for the bad outcomes that she produces even though overt expressions of blame toward her are ultimately unfitting. Indeed, it seems plausible to go a step further and suggest that vociferous praise for the good outcomes produced by Laura is fitting even though this does not quite match the level of praise that she is earned by the value of the outcomes themselves. The reason is that this would represent a positive reinforcement that may facilitate Laura’s recovery from severe depression and thereby decrease the risk that she continues to produce bad outcomes.8

The case just described shows that the overt blame of a blameworthy individual can harm that individual. Again, while this consideration may affect the conditions that determine whether and by what means expressions of negative attitudes are fitting, they do not affect the conditions that determine whether Laura is a fitting target of blame. Just as before, this can also be illustrated by less exotic examples, whereby the overt blame of a blameworthy individual instead harms the friends and family of that individual. When Laura is overtly blamed by a large group of people for the bad outcomes that she produces, her friends and family may be forced to endure social harms. If the harms in question are significant, then this may have relevance for the issue whether and by what means expressions of blame are fitting.

The fitting targets of blame are often the fitting targets of punishment, just as the fitting targets of praise are often the fitting targets of reward. Imprisoning a parent for a crime may have adverse effects on her closest family. In addition to being deprived of support and companionship, their partner may face the emotional and economic hardships that are associated with becoming a single parent. As a result, their children could suffer abandonment and shame, eventually contributing to depression and anger later in life. While considerations regarding these kinds of destabilizing effects on the family can motivate a mitigation of the punishment by courts, this would not necessarily be because imprisonment as such is thought to be unfitting, or that the courts would regard them as punitive for the family.9

Giving a prestigious award to a political activist in an authoritarian nation for their civil right struggles could have adverse effects for the political activist and her

8 It is difficult to find general principles able to capture the conditions that determine the fittingness of attitudes or the fittingness of expressing them. However, rules of thumb can be found that identify salient factors that often contribute to making expression of attitudes unfitting. For example, the substantive example just invites a rule of thumb pertaining to the moral cost of expressing fitting blame. It states that if the effects of expressing blame toward an individual are worse overall than the outcome for which the individual is the fitting target of blame (and if the badness of the effects in question are not outweighed by the intrinsic goodness of the expression as such), then it is usually unfitting to express blame toward that individual.

9 Indeed, even if steps are taken to avoid the morally innocent from being adversely affected by individual criminal sanctions, strictly speaking, such effects do not necessarily undermine the inherent fittingness of the criminal sanctions themselves. Cf., Lewis, op. cit., p. 10; Smith, op. cit., p. 56. I have recently learned of a similar argument being made in Himmelreich & Lawford-Smith, op. cit. There, a distinction is made between internal justice and external justice, where the latter has to do with how the effects of group-punishments are distributed among group members, while the former has to do with whether the groups as such are the fitting targets of punishments. Himmelreich and Lawford-Smith argue that facts about the distribution of punishment-effects among group members do not generally undermine the permissibility of group-punishments as such. The following discussion focuses on attitudes and their expression but occasionally mentions the practice of issuing sanctions and rewards as well. For a detailed discussion about this practice and its potential unfairness, see ibid.
friends. The increased attention resulting from her receiving the reward may exacer-
brate the harsh treatment imposed by the state, who would be incentivized to turn
the political activist and their friends into cautionary tales.\footnote{10} While consid-
erations of this kind may play a part in a decision to withhold any awards until the situation
improves, this would not necessarily imply that the awards themselves are unfitting.
Even if steps are taken to avoid the morally innocent from being adversely affected,
strictly speaking, adverse effects do not undermine the fittingness of the rewards
themselves. Facts about the practical effects of giving a reward to a political activist
is not of the kind to determine whether the reward as such is fitting.

The analogies involving sanctions and rewards are not perfect since they do not
present a close corollary to the distinction between expressing attitudes and merely
holding them. Nevertheless, they do help illustrate the additional point that consid-
erations regarding the practical effects of a response, which may reasonably affect a
decision about whether and by what means it should be carried out, need not affect
the fittingness of the response as such. The examples involving attitudes and their
expression are a special case, for the fittingness of holding attitudes such as praise or
blame need not be affected by global considerations concerning the practical value
of their expression. Inasmuch as such considerations of unfairness have weight, it is
at least not typically to be found in a decision whether to praise or blame, but in a
decision whether and by what means to express these attitudes.\footnote{11} There will be more
on this, but before that we will revisit the unfairness objection.

3 Cashing Out the Unfairness Objection

It seems that the unfairness objection to the practice of collective responsibility
can be cashed out in a multitude of ways depending on how collectives are thought
and spoken about. Suppose Annie quits her office job because of the hostile work
environment created by her male colleagues. When she expresses blame toward the
men at her office for making her quit, she might have in mind each and every man
at the office. This means that her collective talk is used as a shorthand allowing

\footnote{10} Nazi Germany adopted the medieval practice of \textit{Sippenhaft}, whereby the families of known politi-
cal opponents were punished with the confiscation of property and imprisonment. The practice fell out
of favour toward the end of the war, when the regime realized that the practice galvanized and shone an
even bigger light on political opposition. See Robert Loeffel, \textit{Family Punishment in Nazi Germany: Sip-
penhaft, Terror and Myth} (Palgrave Macmillan, 2012). Similar considerations are in play today and so
realistically speaking, giving an award to a political activist, thereby drawing attention to them, might
help protect her from being sanctioned by authorities.

\footnote{11} It is worth considering whether the conditions under which attitudes are fitting can overlap with the
conditions that determine whether and by what means it is fitting to express them. The issue may be
approached by considering cases where the unfair effects of even holding attitudes are so severe as to
put constraints not only on whether and by what means it is fitting to express them, but also on whether
it is fitting to hold them in the first place. In effect, such cases would put into question the necessary con-
nection between the fittingness of attitudes and the notion of being either praiseworthy or blameworthy.
Suppose an evil demon threatens the world with destruction in the event that anyone blames him for his
threatening ways. Perhaps it makes sense to suggest that even though the evil demon is blameworthy in
this case, it is not, in fact, fitting to blame him. Understood this way, the issue has a strong relation to the
more general \textit{wrong kind of reasons problem}. For an influential discussion of the problem, see Rabinow-
icz & Rønnow-Rasmussen, op. cit.
Annie to express blame toward many individuals without needing to mention every one of them. This aggregate way of treating individuals appears commonplace and so it is worth considering how it might be relevant to assessments of the unfairness objection. In fact, the most obvious worry in this context is very similar to the worries that were mentioned in the previous section, about the collateral damage that might result from assignments of moral responsibility specifically to individuals.

The lesson is that when talk of responsible collectives is used as a shorthand for talk of responsible individuals, care needs to be taken to mitigate the fact that some individuals are blamed, praised, sanctioned, or rewarded (or otherwise benefited or harmed) for outcomes for which they are not responsible. It is conceivable that if Annie were to clarify what individuals she has in mind, she might admit that she does not blame each and every man at her office. While the number of responsible men may still seem large enough to warrant the shorthand, a lack of contextual clues can make it difficult to infer whom Annie has in mind. As a result, her male colleagues might be unsure whether they are among the wrongdoers, just as the rest of us are unclear about where to direct our blame. If the practice of collective responsibility always worked like this, then it is easy to see how it might lead to unfair effects.

Indeed, part of the reason why people are often uncomfortable with the practice of collective responsibility could be that it is often talked about specifically as the unjust practice of blaming or punishing innocents for the behaviour of others. A classic example is the authoritarian practice of kin punishment, whereby individuals are punished for the deeds of a family member. Another classic example is the teacher’s practice of punishing children for the behaviour of their classmates. Note, though, that while cases such as these may often be described as cases of collective responsibility (or perhaps more commonly, collective punishment), they could just as well be described as misuses of the practice of individual responsibility. After all, they essentially involve holding one or many persons individually responsible for outcomes for which they are not ultimately responsible.

This understanding of collective responsibility is reflected in the philosophical literature as well. When H. D. Lewis expresses his unfairness worry about the practice of collective responsibility, it seems clear that he has in mind the practice of blaming individuals for the outcomes produced by other individuals. He insists that “we cannot answer for another or share each other’s guilt (or merit), for that would imply that we could become directly worse (or better) persons morally by what others elect to do—and that seems plainly preposterous”. What Lewis objects to, in other words, is the idea that an individual could produce an outcome and share responsibility for it with another individual who did not produce that outcome. While this is right as far as it goes, it seems fair to suggest that it neglects senses of collective responsibility that do not require shared guilt (or merit).

12 The examples illustrate misuses when the aim is to assign moral responsibility for past outcomes, but the practice can have other uses, like spreading fear, severing loyalties, and influencing future behaviours. See footnote 17.

13 Lewis, op. cit., p. 12.

14 Marion Smiley, “Collective responsibility”, Edward N. Zalta, ed., The Stanford Encyclopedia of Philosophy (Summer 2017 Edition), section 4.
The idea is that rather than using her collective talk as a sort of shorthand for expressing blame toward one or many individuals, Annie might be targeting her blame toward a genuinely collective entity, like a \textit{collection of agents} or even a \textit{collective agent}.\footnote{Ibid.} The distinction between the two is rooted in the intuition that there are differences between groups comprised of independently acting agents and integrated collectives that form intentions in unison, perhaps in accordance with established decision procedures.\footnote{See Peter French, \textit{Collective and Corporate Responsibility} (New York: Columbia University Press, 1984). Examples of mere collections of agents that nevertheless have the potential of acting jointly might include such things as school classmates and a concert audience, while examples of collective agents might include such things as corporations and military units. See e.g. Pettit, op. cit., for some additional examples.} Annie might be forced to quit her job because each and every man at her office happens to be rude and lacking in empathy. Such a case, where a regretful outcome is realized by independently acting men, is different from a case where the men are parts of an integrated community, characterized by goals and values that lead to harsher conditions for other office workers.

Attempts have been made to account for collective agency, but no such account enjoys wide acceptance within the literature. While the existence of mere collections of agents is far less controversial, their lack of joint agency is often taken to suggest that they are the unfitting targets of responsibility assignments.\footnote{Gunnar Björnsson expresses reservations about this idea. He at one point maintains that moral responsibility can be “essentially joint”, meaning that agents that do not constitute a collective agent and that lack individual responsibility for an outcome may still be jointly responsible for that outcome. See Gunnar Björnsson, “Joint responsibility without individual control: Applying the explanation hypothesis”, Nicole A. Vincent, Ibo Van de Poel, Jeroen Van den Hoven, eds., \textit{Moral Responsibility} (Library of Ethics and Applied Philosophy vol. 27, 2011), “Collective responsibility and collective obligations without collective moral agents”. Saba Bazargan-Forward & Deborah Tollefsen, eds., \textit{Handbook of Collective Responsibility} (Routledge, 2020). Cf., also Smith, op. cit.} Note that if mere collections of agents cannot be morally responsible, then the main objection to the practice of holding them responsible is not that doing so leads to unfair effects for individuals, but that the practice as such is misguided. On the other hand, if mere collections of agents \textit{can} be the fitting targets of praise and blame, then the practice of holding such collections responsible is subject to many of the same unfairness worries as the practice of assigning moral responsibility to collective agents. The following discussion therefore takes on a general perspective that is meant to apply regardless of whether the practice of collective responsibility involves mere collections of agents or collective agents.

The point emphasized in what follows is that both the practice of collective responsibility and the practice of individual responsibility may result in collateral damage and that both are susceptible to misuses that mirror the collective punishment cases described earlier. It should be noted that some of the philosophers that present an unfairness objection to the practice of collective responsibility would agree. Sverdlik is among them, for he is careful to point out that it would be unfair, “whether we are considering a result produced by more than one person’s action...
or by a single person, to blame a person for a result that he or she did not intend to produce” (emphasis added). 18 This remark displays an awareness and appreciation of the fact that both the practice of collective responsibility and the practice of individual responsibility involve the potential of unfair effects. This fact remains relevant regardless of whether the practice involves mere collections of agents or collective agents.

The question, then, is whether holding collective entities responsible is more likely to result in unfair effects than the practice of holding individuals morally responsible. Making progress on the issue requires looking closer at what it means to hold collectives morally responsible in the first place, as well as the commitments that might be involved in blaming, praising, sanctioning, and rewarding them. Some reasons have already been given for believing that the practice of collective responsibility need not be more problematic than the practice of individual responsibility, but it is worth pausing on some of those reasons for a bit longer. This will help rule out potential sources of unfairness and further highlight the parallels between the two practices. 19 The distinction between collective agents and mere collections of agents will be reconsidered later, when the question is asked whether there might be more specific formulations of the unfairness objection that remain to be explored.

4 Holding and Expressing Attitudes in Collective Cases

Philosophers worrying about the practice of collective responsibility are often focused on the issuing of sanctions. For example, when Björn Petersson and Govert den Hartogh express their respective unfairness worries, both bring up cases where individual group members are not just blamed but punished for outcomes that are produced by the their respective collectives. 20 Petersson warns us that “collective

18 Sverdlik, op. cit., p. 68. The objection might be made that there are cases where we can fittingly blame individual agents for outcomes that they produce unintentionally. Some of the more obvious examples are cases of negligence and recklessness. Whether they ever intend to crash their cars into things, it seems intuitive to suggest that drunk drivers are responsible for the damage that they cause others by drinking and driving. Of course, the argument might be made that, strictly speaking, drunk drivers are not responsible for their car crashes, but for the many actions that precede them, i.e. drinking too much, not getting a designated driver, and so on. See footnote 31.

19 It is worth saying something more to separate out the unfairness objection from some of the wider problems that may arise for the practice of collective responsibility. The idea that holding collective entities responsible necessitates the blame or praise of individual group members could be propped up by the metaphysical argument that there are no irreducibly collective entities in the first place and so whenever collective agents are apparently held responsible, individuals are the ones actually being blamed or praised. While there may be something to this argument, it seems that it would fail to prop up the unfairness objection as an independent argument against the practice of collective responsibility. Opponents to the practice could just as well ignore considerations of unfairness and focus exclusively on the allegedly dubious existential commitments that the practice involves.

20 See Petersson, op cit; den Hartogh, op. cit. Also note that while it is quite common to assume that punishing agents is just a way of blaming or at least holding them responsible, Thomas Scanlon has argued that sanctions are typically meant to inflict harm upon the agent being sanctioned and that this is not the point of blame. For this reason, he wants to discourage the habit of describing blame as a type of sanction and vice versa. See Scanlon, What we owe to Each Other (Cambridge, MA: Harvard University Press, 1998), ch. 6.
sanctions strike individual members” and that we therefore need to demarcate collective agents so that the only individuals that end up struck are those “that justifiably can be blamed for what the group has done”. Den Hartogh similarly states that when groups are sanctioned “some individuals will in the end have to take on the burdens of that punishment” and that this is unfair “if they have done nothing for which they personally deserve to be blamed, or at least to be blamed to that extent”.

The previous section distinguished individualistic readings of collective talk from the more robustly collective interpretations discussed here, but not enough was said about what the practice of collective responsibility involves. Part of the practice is assumed to involve the adoption of blame and praise, but what remains to be explored is the connection between the fittingness of such attitudes and the fittingness of issuing sanctions and rewards. Petersson and den Hartogh seem to assume that the practice of issuing sanctions is central to the practice of collective responsibility and this grounds some of their unfairness worries. Roughly, the worry is that individual group members cannot be left alone while the collective entities that they constitute are fined, imprisoned, or made to pick up roadside garbage. Many punishments that are directed toward collective agents will strike individual group members as well.

The argument might be made that while issuing sanctions and rewards is central to practices of legal responsibility, it has little to do with practices of moral responsibility, but this would go too far. Firstly, it seems very plausible to suggest that the factors that make things the fitting targets of attitudes often contribute to making them the fitting targets of sanctions and rewards as well. Secondly, some of the sanctions and rewards that are involved in practices of moral responsibility may not be of the legal but of the social variety (consider cases where agents become the fitting target of exclusion from certain social arrangements). What sanctions and rewards are called for by instances of collective responsibility and whether these differ from those that are called for by instances of individual responsibility is an open question that deserves more attention, but it has to be set aside for now.

Notice that sanctions and rewards seem to be characterized by a certain kind of goal-directedness. This is reflected in the intuitive distinction between being

21 Op. cit., p. 850.
22 Den Hartogh, op. cit., p. 119. Mark Reiff expresses similar thoughts but has normative concerns that go beyond the kind of unfairness worries discussed here. For example, he argues that thinking in collective terms has disastrous effects on how people understand morality and the obligations they have toward themselves and the world. See Mark Reiff, “Terrorism, retribution, and collective responsibility”, Social Theory and Practice, vol. 28, no. 3 (2008), pp. 442–455. For an argument stating that thinking about collective responsibility can lead to moral progress, see e.g. Tracy Isaacs (2011), Moral Responsibility in Collective Contexts (New York: Oxford university press, 2011).
23 Cf., Himmelreich & Lawford-Smith, op. cit.
24 While factors that make attitudes fitting can contribute to the fittingness of actions as well, perhaps this is not always the case. There might be blameworthy agents that it is not fitting to punish, just as there are praiseworthy agents that it is not fitting to reward. There is of course a problem of accounting for when the fittingness of attitudes coincides with the fittingness of actions, but this appears to be a very general problem in ethical theory.
punished or rewarded, on the one hand, and being affected by the punishments or rewards bestowed on someone else, on the other.\textsuperscript{25} When the courts punish a parent for a crime, their aim is to inflict some harm upon the parent that is constitutive of the punishment. While the punishment is also likely to have negative effects on the family of the accused, it is not the aim of the justice system to bestow any harm upon them and so they are not punished. This suggests that what is done to collectives need not be done to their individual group members. T. H. Smith makes the very same point and he thinks it applies just as much to attitudes as it does to sanctions and rewards:

Now, an act or attitude can be targeted at a plurality without it, any part or portion of it, or anything of the same kind as it, being targeted at any one of them: an actor who fears his audience targets his fear at them, but may not target it, or any part or portion of it, or any act or attitude of the same kind as it, at any one of his audience. Likewise, an award can be given to a plurality without it, or any part or portion of it, or anything of the same kind as it, being given to any one of them. Sometimes, perhaps, a single award goes to each of a plurality, as when a tie occurs. But when the achievement recognized has been done in collaboration (one thinks of Oscars and Nobel prizes), an award will go to a plurality, none of whom receive it, any part or portion of it, or anything of the same kind as it.\textsuperscript{26}

It seems right to suggest that collective entities can be held responsible without their individual group members necessarily becoming praised or blamed. As Smith indicates, the claim can be supported by making comparisons with the attitudes that are called for by values other than blameworthiness and praiseworthiness. A lovable person can be fittingly loved without all her constituent parts or characteristic features necessarily becoming loved, just as an admirable artwork can be fittingly admired without all its brushstrokes necessarily becoming admired, just as a preferable world can be fittingly preferred to another possible world without every state in the former necessarily becoming preferred to every state of the latter. It may be that when some attitudes are directed toward an object, some of the properties or constituent parts of the object will be featured in the intentional contents of the attitudes. Nevertheless, as Noah Lemos rightly points out, this does not mean that

\textsuperscript{25} Cf., Nagel’s version of Aquinas’ doctrine of double-effect: “Not everything that happens to others as a result of what one does is something that one has done to them” (Thomas Nagel, Mortal Questions (Croydon: Cambridge University Press, 2012), p. 60). See also Himmelreich & Holly Lawford-Smith, op. cit., p. 136.

\textsuperscript{26} Smith, op. cit., p. 57. Note that Smith rejects the idea that there are collective agents—or, more accurately, he denies that we need to appeal to such agents to make sense of cases where collectives appear blameworthy. He argues that several things (a plurality, but not a collective entity) can be the target of attitudes without any individual agent thereby becoming the target of those attitudes and actions. Nevertheless, some of his arguments can be used to support the idea that if there are such things as collective agents, then these can be the targets of attitudes and actions without any individual agent thereby becoming the target of those attitudes and actions.
each such property or part is as much of a target of the attitudes as the object that bears them.\footnote{27}

The properties or constituent parts of an object could be featured in the intentional contents of attitudes indirectly, not as the intended targets of the attitudes in question, but as the things that characterize and help identify their intended targets.\footnote{28} The objection might be made that attitudes cannot be directed toward an object without the attitudes thereby being directed toward some of its parts, but it is unclear what reasons there are for thinking that this is true. It is not shown to be true by the mere fact that fitting attitudes are always directed toward objects for some of their constituent parts or characteristic features.\footnote{29} A consideration can make certain attitudes fitting without thereby becoming a fitting target of the same kinds of attitudes.

What this shows is that there are wider axiological discussions to be had about the nature of attitudes and whether favouring and disfavouring objects requires the favouring and disfavouring of their respective parts or properties.\footnote{30} From this wider perspective, it seems intuitive to suggest that favouring and disfavouring objects does not necessarily require the favouring and disfavouring of their respective parts or properties. The point is that if those general intuitions have any validity, then they apply here too, and lend support to the view that holding collective entities

\footnote{27} Noah Lemos, *Intrinsic Value: Concept and Warrant* (Cambridge: Cambridge University Press, 1994), p. 10.

\footnote{28} Cf., Toni Rønnow-Rasmussen, *Personal value* (Oxford: Oxford University Press, 2011), pp. 68–73.

\footnote{29} While attitudes are in principle capable of significant discernment, human psychology is admittedly messy and so attitudes directed toward collective agents could still have the unfortunate tendency to “trickle down” to their individual group members. However, it seems that such facts about the caprice of human psychology are just as likely to disrupt the practice of individual responsibility as it does its collective counterpart.

\footnote{30} Several philosophers have argued convincingly that there are responsibility gaps in the sense that there are collective agents that are morally responsible for outcomes even though their individual group members are not. For a discussion, see e.g. Pettit, op. cit.; David Copp, “The collective moral autonomy thesis”, *Journal of Social Philosophy*, vol. 38, no. 3 (2007), pp. 369–388; Matthew Braham & Martin van Hees (2012), "Responsibility voids", *The Philosophical Quarterly*, vol. 61, no. 242 (2012), pp. 6–15; Hein Duijf, “Responsibility voids and cooperation”, *Philosophy of the Social Sciences*, vol. 48, no. 4 (2018), pp. 434–460. Arguments in favour of responsibility gaps often focus on intuitive examples as well as more technical considerations borrowed from decision theory and legal philosophy. For example, much of the literature considers the possibility of discursive dilemmas where individual group members vote on decisions and produce results that are not immediately reflected in the preferences of any individual voter. Cf., Lewis A. Kornhauser, “Modelling collegial courts II: Legal doctrine. Journal of Law”, *Economics & Organization*, vol. 8, no. 3 (1992), pp. 441–470; Bruce Chapman, “More easily done than said: Rules, reasons and rational social choice”, *Oxford Journal of Legal Studies* vol. 18, no. 2 (1998), pp. 293–329. Note that if there are indeed responsibility gaps, then there should be cases where it is fitting to adopt a pattern of attitudes such that a collective agent is blamed or praised even though no individual group member is blamed or praised. This is important, because if it turns out to be impossible to adopt such a pattern (for holding collectives morally responsible results in the blame or praise of individual group members as well), then such a pattern is not fitting. Arguably, a pattern of attitudes that cannot even be adopted cannot be fitting and so the very notion that there are potential responsibility gaps already speaks in favour of the view that holding collectives morally responsible does not necessitate the blame or praise of individual group members. On the other hand, if there are no responsibility gaps to begin with, then there is of course no need for any unfairness objection, seeing as collectives are then the unfitting targets of blame and praise to begin with. While this response to the unfairness objection is a tempting one, it is left aside here as it appears to stack the deck against proponents of the objection. The paper proceeds not from questions about whether there are collective agents or whether collective agents can be responsible when none of their individual group members are, but from the question whether the practice of holding collectives responsible leads to unfair effects for individual group members.
responsible does not necessarily require holding individuals responsible. Proponents
of the unfairness objection might acknowledge this point and maintain that though
the practice of collective responsibility does not generally require the blame and
praise of individual group members, there are instances where it does.

In other words, the idea might be that there are specific situations, perhaps
involving certain types of collectives and outcomes, where the practice of collective
responsibility does involve the unfair blame and praise of individual group mem-
bers. For example, perhaps there are salient differences between mere collections
of agents and collective agents such that the practice of holding one responsible for
an outcome tends to produce more unfair effects than the other. This means that the
unfairness objection is weaker than it might first appear, involving no general claim
about the connection between the practice of collective responsibility and the unfair
blame or praise of individual group members. Of course, there is then a challenge on
part of its proponents to explain what the problematic instances look like (i.e. what
it is about the relevant types of collectives and outcomes that leads to a disruption
in the goal-directedness of actions and the intentional structure of attitudes), just as
they need to show that they involve cases where the practice of collective responsi-

bility truly does end up being unfair.

For now, it seems safe to suggest that the unfairness objection is on somewhat
shaky ground if it is based on the general claim that holding collectives responsible
for outcomes necessarily requires holding potentially innocent individuals responsi-
ble as well. The points made above about the goal-directedness of actions, as
well as the discernment and intentional structure of attitudes, indicate that this is
not a guarantee. Nevertheless, while holding collectives responsible for outcomes
does not necessarily require holding individuals responsible for those outcomes,
individuals could nonetheless be affected in unfair ways by the practice of collective
responsibility. Velasquez shares some of the worries of Petersson and den Hartogh
and thinks that they apply both to the adoption of attitudes and the issuing of sanc-
tions and rewards. He has this to say about the effects of holding collectives morally
responsible:

But in fact it is not possible to impose blame or punishment upon an organiza-
tional structure without having that blame or punishment fall on the shoulders
of the corporation’s members. It is the members who will feel all the effects
and bear all the injuries if the corporation’s structures are “fined,” if its “public

31 The objection might be made that if there is too large of a gap between individual responsibility and
its collective counterpart, individuals may ask what it is to them that the collective agents they are parts
of are blameworthy. Leave it to the collective agents themselves, they might say, to react appropriately
to being blamed. However, adherents of the practice of collective responsibility agree that when collec-
tive agents are responsible for certain outcomes, their individual group members tend to be responsible
for other closely related outcomes, like becoming part of wrongfully acting collective agents, not doing
enough to influence the trajectory of the relevant collective intentions, and so on. See e.g. Lewis, op.
cit., p. 13; Karl Jaspers, The Question of German Guilt, Ernst Basch Ashton, trans. (New York: Capri-
corn, 1961); Held, op. cit., p. 479; Hannah Arendt, “Collective responsibility”, in: James W. Bernauer,
ed., Amor Mundi: Explorations in the Faith and Thought of Hannah Arendt (Boston: Martinus-Nijhoff,
1987); Larry May, Sharing Responsibility (Chicago: University of Chicago Press, 1992); Copp, op. cit.,
p. 374; Pettit, op. cit., p. 192–194; den Hartogh, op. cit., p. 123.
image” is “tarnished,” or if these structures are altered or perhaps even dissolved. These members are therefore being unjustly forced to bear the punishment for another entity’s moral responsibility (in addition to any punishment they may have to bear for their own moral responsibility for the act).32

This is perhaps one of the clearest formulations of the unfairness objection in the philosophical literature and it does not say that holding collective agents morally responsible necessarily requires holding potentially innocent individuals morally responsible. Instead, what it says is that potentially innocent individuals will inevitably be affected by the practice of collective moral responsibility in other ways that are unfair. Even if there is something to this, it is once again unclear whether this makes the practice of collective responsibility different from its individual counterpart. We saw in an earlier section that there is a potential for unfair effects in cases of individual responsibility as well. The examples that were used to illustrate this point can be adapted to cases where collectives are held morally responsible for an outcome.

There may be conceivable situations where collectives are responsible for good outcomes (making them fitting targets of praise for those outcomes) even though it is unfitting to express praise toward them. Suppose a collective produces a good outcome that lies beyond the intentional control of its individual group members. Imagine too that if the collective as such becomes a target of overt praise, then some of its individual group members will suffer bad outcomes. These may well represent the kinds of situations where the collective as such is worthy of praise, even though overt expressions of blame toward the collective ends up being unfitting. Perhaps the example could be imagined in more detail, as a collective counterpart to the previously discussed example involving Kate. The collective group members think, intend, and act in such a way as to produce a collective action that leads to a good outcome. The collective may to that extent be praiseworthy, but because each member of the collective is as narcissistic as Kate is, the knowledge that they are parts of a praised collective might decrease their chances of thinking, intending, and acting in the right way again.

There may also be situations where collectives are responsible for bad outcomes (making them fitting targets of blame for those outcomes) even though it would be unfitting to express blame toward them. Suppose that a collective produces a bad outcome that lies beyond the intentional control of its individual group members and that those group members would suffer bad results if the collective as such became the target of overt blame. The bad outcomes suffered by the individual group members could even be worse than the bad outcome that they originally produced as a collective. This represents the kind of situation where the collective as such is worthy of blame, even though expressions of blame toward it ends up being unfitting. This example could be described as a collective counterpart to the one involving Laura. While each member acts in a blameless way, the collective action that results produces slightly worse outcomes for all of them. The collective may to that extent be blameworthy, but because each member of the collective is as emotionally fragile

32 Velasquez, op. cit., p. 124.
as Laura is, they risk harming themselves were they to become knowing parts of a blamed collective.

Local considerations affecting the conditions under which it is fitting to hold collective agents responsible ought not to be confused with wider considerations determining whether and by what means it is fitting to express praise or blame toward collective agents. A similar point can be made regarding actions directed at collectives, although such cases tend to become more sophisticated and difficult to get a firm grip on. For example, in the arena of international politics, governments often impose economic sanctions on nations that result in significant negative effects for their citizens. The effects of these sanctions may be so severe that they start to affect the extent to which the citizens can satisfy their basic needs. There is a sense in which the sanction may still be fitting in respect of the collective, but its unfair effects on individual citizens may be such as to recommend that it not be carried out.

The argument has just been made that holding collective agents responsible for an outcome does not necessarily require holding individuals responsible for that outcome. Furthermore, the argument has also been made that the unfair effects of the practice of collective responsibility do not necessarily undermine the fittingness of blaming or praising collectives. The examples of previous sections do not just help to illustrate the generality of these specific points, but also help to show that the considerations of unfairness that crop up for cases of collective responsibility are ubiquitous in individual cases. The overt blame of blameworthy individuals can produce harms for those individuals and their loved ones, just as the overt praise of praiseworthy individuals can lead to undeserved benefits for those individuals and their friends. The practice of individual responsibility and that of collective responsibility therefore appear to be on equal ground as far as considerations of unfairness go.

If the practice of collective responsibility is unfair, then, this is because of factors that make the practice of individual responsibility unfair as well. Indeed, most of the points that have been made in this paper can be generalized beyond the domain of moral responsibility altogether, as considerations of unfairness are likely to crop up for many cases where values other than blameworthiness and praiseworthiness call for the adoption of attitudes. The previously mentioned examples involving lovability and admirability suffice to show that this is true. That Laura is worthy of love does not guarantee that her friends are worthy of the benefits resulting from her becoming loved. That an artwork is worthy of admiration does not guarantee that its creator is worthy of the benefits resulting from it becoming admired. Given the ubiquity of these kinds of considerations, it seems unlikely that they should present any deep problem specifically for a practice of collective responsibility.

For all that has just been said, there is admittedly something intuitive about the suggestion that the practice of collective responsibility is somehow more sensitive to considerations of unfairness, even though the same considerations are evidently relevant to all cases where the presence of value calls for the adoption of attitudes. Perhaps this is because practices of moral responsibility—be they individual or collective—have a social dimension that is meant to be more pronounced than in many other such practices. For example, responding to the moral responsibility of a person by either blaming or praising them is in some sense associated with greater moral stakes than responding to the beauty of a person by adopting an attitude of
aesthetic appreciation. Another remnant of the unfairness objection is therefore a challenge to account for what these stakes might be and precisely how they could affect the fittingness of holding and expressing attitudes such as blame and praise.

5 Concluding Remarks

The purpose of the paper was not to justify indifference in respect of the outcomes that might be produced by the practice of collective responsibility. We have seen that care needs to be taken in its application to make sure that its outcomes do not affect individuals too unfairly. Rather, the paper was meant to show that these considerations of unfairness are ubiquitous and relevant in both collective and individual cases. The result is that it is difficult to see how the unfairness objection could present a deep and serious problem specifically for the practice of collective responsibility. Blaming or praising collectives as such does not necessarily require the blame or praise of individual group members, nor does the punishment of collectives necessarily require the punishment of individual group members. Moreover, while the practice of collective responsibility can affect individual group members in ways that might be considered unfair, this makes it no different from the practice of blaming or praising individual agents. Whenever an individual is targeted by attitudes such as blame and praise, or punished for the outcomes that they have produced, there is the risk that this will affect other individuals as well. Proponents of the unfairness objection therefore need to state more clearly what the effects are that supposedly undermine the practice of collective responsibility and how they differ from the effects that might result from the adoption of fitting attitudes in other contexts.

Acknowledgements The paper was drastically improved as a result of the insightful and constructive feedback given to me by two anonymous reviewers. I owe them a great debt of gratitude. Very early drafts of the paper were also helped by the suggestions of Olle Blomberg, Björn Petersson, Matthew Talbert, Mattias Gunnemyr, Jakob Green-Werkmäster, Thomas Schmidt, Benjamin Kiesewetter, Jared Parmer, and participants at the ethics colloquium at Humboldt University of Berlin.

Funding Open Access funding provided by Lund University. Vetenskapsrådet (201806612)

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.