

Punishing Groups: When External Justice Takes Priority over Internal Justice

Johannes Himmelreich and Holly Lawford-Smith

Abstract. Punishing groups raises a difficult question, namely, how their punishment can be justified at all. Some have argued that punishing groups is morally problematic because of the effects that the punishment entails for their members. In this paper we argue against this view. We distinguish the question of internal justice – how punishment-effects are distributed – from the question of external justice – whether the punishment is justified. We argue that issues of internal justice do not in general undermine the permissibility of punishment. We also defend the permissibility of what some call “random punishment.” We argue that, for some kinds of collectives, there is no general obligation to internally distribute the punishment-effects equally or in proportion to individual contribution.

1 Introduction

Groups are regularly punished. Consider three examples. In 2017, Volkswagen was ordered to pay \$4.3bn in fines for manipulating their vehicles to pass emissions tests. The company has set aside €23bn in total to cover the expected costs of the scandal (McGee and Lynch 2017). Although the long-term effects of this form of punishment cannot yet be discerned, shareholders in, and at least some employees of, the company are likely to be affected negatively. Second, in 2002 the accounting firm Arthur Anderson, then one of the largest accounting firms worldwide, was convicted in relation to the Enron accounting scandal. In result, the company had to cease operations. Of the 85,000 employees the company had back then, only 200 are left today working to dissolve the firm—the rest lost their jobs. Third, the US, despite a global increase in travel volume in 2017, has seen a decline in the number of international visitors. Anecdotal evidence suggests that at least some travelers actively avoid visiting the US to quasi-boycott the administration of President Trump (Trejos 2018).¹ Of course, these non-travelers aim at punishing the current US government but the ones feeling the effect of this treatment are individuals living in the US—partners, friends, relatives, and colleagues, of the non-travelers, and those who would have benefited economically from the additional tourist dollars.

¹ This third example raises several complications. For example, citizens are not members of the administration, that is, the group against which the boycott seems to be directed.

These cases illustrate a challenge for the permissibility of punishing groups. Although the punishment is directed at a group, the punishment, in effect, is borne by individuals (Smith 2009; Pasternak 2011b). In this paper, we address this challenge and defend the permissibility of and develop a justification for punishing groups. We are particularly interested in groups with a hierarchical structure, like corporations and states. Is it permissible to punish groups given how punishment-effects² distribute within a corporation, particularly to their low-level employees? How should we think about the distribution of punishment-effects, e.g. for the state's war crimes or for historic injustices, to members of the state – however these are best understood (see discussion in Lawford-Smith forthcoming)? Many of these issues have been addressed elsewhere (Erskine 2010; Wringer 2016, chaps. 7–8). Our focus is on a specific challenge to the permissibility of collective punishment, namely that its distribution is internally disproportionate. We are not attempting to contribute to discussion about the nature of punishment or the possibility of group agency.

We defend two points. One: we distinguish the question of *internal justice*, namely how punishment-effects are distributed, from the question of *external justice* namely whether the group's punishment is justified. We argue that a disproportionate distribution of punishment-effects (internal injustice) does not in general undermine the permissibility of punishment (external justice). Instead, we argue that satisfying the requirements of external justice by punishing a collective wrongdoer does not in general rest or depend on the specific distribution of punishment-effects of that would be required to satisfy internal justice. Two: we defend the permissibility of what could be called “random distribution” of punishment-effects. Although internal justice favors punishment-effects being distributed equally in egalitarian groups (see discussion in Lawford-Smith forthcoming), and in proportion to the member's role in the hierarchy in hierarchical groups (see discussion in Lawford-Smith forthcoming), external justice takes precedence over internal justice, and so there is no general obligation on groups to distribute punishment-effects equally or proportionally to members' roles. In many circumstances a disproportionate, or “random”, distribution is permissible.

The paper is structured as follows. First, we say a bit more about the distinction between punishment and punishment-effects, illuminating the two concepts with some examples from the legal literature. Second, we reconstruct the challenge for the permissibility of punishing groups as it has been put forward by Avia Pasternak (2011b) and Thomas Smith (2009). We identify the

² A note on our terminology. We're using ‘punishment’ for the agent that is condemned, and ‘punishment-effects’ for the impacts of that punishment on members who are not themselves condemned. This helps to make clear that those individuals are not necessarily themselves being punished. It is possible, in rare cases, for a group to be punished without that involving punishment-effects for members, but in most cases, punishment goes hand-in-hand with punishment effects. Moreover, punishment-effects are arguably constitutive for punishment and not merely a side-effect thereof.

limits of these arguments in order to carve out a space for permissive kinds of collective punishment. Third, we distinguish between three different types of groups, each of which is well-established in the literature, to show that on each of these accounts disproportionate distributions of punishment are permissible.

2 Punishment and Punishment-effects, Moral and Legal Responsibility

There is a difference between being punished and being adversely affected by that punishment in some way. When the company Arthur Anderson was punished for its involvement in the Enron accounting scandal, those adversely affected by that punishment were, among others, their employees, but also dependents of these employees. Yet, the employees and their dependents were not punished in the sense that the punishment was directed at them. Instead, the employees and their dependents were affected by the “overspill” of the punitive sanction against their employer (Erskine 2010).

To mark this distinction between punishment and punishment-effects,³ we follow Joel Feinberg (1965) and distinguish between punishment and costs resulting from punishment. Punishment differs from what we call “punishment-effects” in that the former but not the latter has an expressive or communicative component.⁴ Specifically, punishment carries the sign of condemnation of those against whom it is directed. Feinberg (1965, 400) writes that punishment expresses “attitudes of resentment and indignation, and ... judgments of disapproval and reprobation”. Punishment-effects, by contrast, are those disadvantages that result from someone being punished and that are constitutive of their punishment. More specifically, punishment-effects are those disadvantages that are borne by individuals in virtue of their association with a group to which the punishment is meted out. The form of this association can vary. The association by which punishment-effects are borne by an individual can be that of membership, identification, employment, or investment. We focus only on those individuals who bear punishment-effects and are related to a group in a way resembling membership in that group (so for example, in the case of Arthur Anderson, we would be focusing on the employees who lost their jobs, but not their dependents, at one further remove, who were negatively affected by the employees’ loss of earnings). This focus is important because, as we will argue, the nature of this relationship matters

³ Instead of “punishment-effects” Pasternak (2011b), who also draws this distinction, talks of “liability.”

⁴ This is true on Feinberg’s theory, and on expressive theories of punishment in general, which we are assuming here. It’s not necessarily true on alternative theories of punishment, such as retributivist, rehabilitative, reconciliatory, and deterrence theories. See discussion in Duff and Hoskins (2018).

for questions of internal justice concerning how the punishment-effects are distributed within the collective.

Since we investigate the permissibility of punishment when it is directed against groups, we first need to think about what reasons there are to punish someone in general. One important precondition for punishment is that of moral responsibility. We focus on moral responsibility because moral responsibility is the condition on which the challenge against collective punishment rests.

What does it take for someone to be responsible for their actions in a way that would justify punishment? Because the precise conditions for responsibility are not the central focus of this paper, we simply gloss the conditions for moral responsibility as ‘causation plus intention’. There is widespread consensus on the view that for an individual to be morally responsible for her action, she must have both caused that action and intended the performance of that action.⁵ The individual must have intended the action under the relevant description or must be aware of its foreseeable consequences and yet not have taken steps to avoid the action.

These two requirements—of causation and intentionality—although with subtle differences, also show up in criminal law. Criminal law assigns legal responsibility on the basis of *actus reus* (the conduct, or action) plus *mens rea* (the intention, or at the very least, negligence or recklessness). Furthermore, criminal law has developed several guidelines as to how and why groups can be punished. In criminal law, we find at least three different treatments of legal responsibility. One treatment of criminal law can be found in individual culpability, two further treatments can be found in the UK’s sentencing guidelines on Joint Enterprise and Corporate Manslaughter. The treatment of the first, of individual culpability, has been outlined above. Individual culpability is the presence of *actus reus* plus *mens rea* within the relevant individual.

The legal treatment of Joint Enterprise applies the conditions for individual culpability to the collective case. The Guidelines on Charging Decisions (Crown Prosecution Service 2012), although suspended pending a legal review, clearly stress that there can be no responsibility on the grounds of membership alone. For example, a person who is normally a member of the group is not culpable if she remains home sick on the evening they perform the relevant act. Likewise, a person is not culpable if she suffers an attack of conscience and flees before the relevant act is performed. However, the treatment of Joint Enterprise goes beyond individual culpability in that it recognizes different modes of involvement in an offence by designating ‘principals’ and ‘accomplices’. A principal must be directly causally involved in the relevant conduct, and an accomplice must encourage or assist in the performance of the conduct (this extends to e.g. posting encouraging remarks on social media websites). In short, the culpability under Joint Enterprise –

⁵ A third condition – next to causation and intentionality – is the epistemic condition concerning whether a harm etc. has or could have been foreseen by the agent. For our purposes here, we set the epistemic condition aside.

although it makes room for the distinction between different modes of involvement – is individualistic (see also discussion in Lawford-Smith 2018).

Corporate Manslaughter, on the other hand, is given a collectivist treatment. It does not require what is known as the *identification principle*, that a senior individual who is taken to embody the company, need to be identified as guilty of the offence in order for the company to be guilty of the offence. Instead an offence under the Corporate Manslaughter and Corporate Homicide Act is defined as an offence committed by organizations through the way they are managed or organized (*Corporate Manslaughter and Corporate Homicide Act 2007*, sec. 1). Furthermore, it requires that “the way in which its activities are managed or organised by its senior management is a substantial element in the breach” (ibid). Noteworthy about this model of Corporate Manslaughter is that although the *actus reus* may come from anywhere within the corporation (the limiting case being from just one low-level employee), the *mens rea* must come from within the corporation’s senior management.⁶ The individualist model is, in the law on Corporate Manslaughter, applied to the corporation itself, requiring the action and the intention to be present within the corporation – as opposed to within the individual person. The corporation may be punished with fines, by being required to compensate victims through a compensation order, by being made to remedy the conditions that allowed the breach through a remediation order, and through a publicity order that makes public of their offense (Sentencing Guidelines Council 2015, 23–28).

In summary, we are employing an expressive rationale for punishment; we are distinguishing punishment from punishment-effects and focusing in particular on the effects on closely associated individuals, in particular *members* of the punished group; and we are assuming that both causation and intention are preconditions of moral, and legal, responsibility, and that moral responsibility is a precondition of morally justified punishment, and legal responsibility is a condition of legally justified punishment.

3 The Challenge for Punishing Groups

In this section, we reconstruct the argument for why punishing groups is morally problematic. The heart of the argument itself is quite simple and consists of two main assumptions. First, the argument starts with the assumption that when a group is punished, punishment-effects are borne by its members. In other words, the assumption says that *punishment is distributive* with respect to its effects (Smith 2009). The second assumption is that if punishment-effects are borne by individuals against whom the punishment is not directed, then this is morally problematic. The punishment spills over to places where it was not meant to go (Erskine 2010). These two

⁶ Although the requirement is explicitly less demanding than the earlier existing legal test of the ‘controlling mind.’

assumptions form the main part of the argument that collective punishment, when it occurs and when it affects individuals against whom it is not directed, is morally problematic.

The first assumption, that punishment-effects are distributive, can be read in any one of four different ways. First, distributiveness can be understood weakly (distributive to *some* members) or strongly (distributive to *all* members). Second, the alethic modality of the assumption can be understood weakly (as in “generally” or “usually”) or strongly (as in “necessarily” or “essentially”).

Generally, that some *F* is distributive over a plurality means that *F* applies to the plurality just in case *F* applies to some or each of the members of that plurality.⁷ This distributiveness can be understood in two ways (Smith 2009). Understood strongly, if *F* is distributively possessed by a plurality then *F* is possessed by *each*. Understood weakly, if *F* is distributively possessed by a plurality then *F* is possessed by *some*. Specifically, punishment would be weakly distributive in its effects if punishing a group entailed that at least *some* members of that group bear punishment-effects as a result. Punishment would be strongly distributive if punishing a group entailed that *all* members of the group bear punishment-effects.

Many authors hold that punishment is necessarily weakly distributive. For example, Thomas Smith (2009, 55), although talking about the distribution of “harms” and not of punishment effects, contends that punishment is “essentially harmful: I cannot penalize or punish if I do not thereby harm. Of course, some costs are easily absorbed, and some deprivations are scarcely noticed, but these can still be *harms*, and an attempt to punish that does not harm does not succeed.”⁸ Similarly, Jackson (1988, 264) contends that punishment is necessarily weakly distributive. Likewise this claim is made by Pasternak (2011b, 212, 2011c, 189), Feinberg (1968, 687), Stilz (2011, 194), and Coffee (1981, 401).

Is this assumption sound? We think it is not. Group punishment is not necessarily or even essentially distributive. Pasternak (2014, 225) provides a counterexample. Suppose a group has inherited assets from past generations. In the case that Pasternak discusses, a church has an emergency fund that was set up by members who have meanwhile passed away. Suppose that the group, as a matter of punishment-effects, owes compensation to a third party and the group decides to use the emergency fund to pay for this compensatory claim. In this case, there clearly seem to be punishment-effects in the form of monetary costs but these costs are not actually distributed to group members. This puts into question that collective punishment is necessarily distributive.

A more plausible version of this first assumption would be that collective punishment is *contingently* distributive. Punishing some groups entails that punishment-effects are borne by at least some of their members. Taking the assumption to be true only contingently limits the force

⁷ Distributiveness can be understood as a feature of predicates or as a feature of properties.

⁸ Smith goes on to reject the structurally analogous claim about blame, arguing that there can be non-distributive blameworthiness, but maintains the claim about punishment.

of the challenge against collective punishment. The challenge is not a challenge for punishing groups in general. The challenge arises at most in cases where some punishment-effects are in fact distributive. It seems now more a practical than a conceptual challenge.

Let us now consider the second main assumption of the challenge. This second assumption, most generally, says that if punishment-effects are borne by individuals, then this is morally problematic. By “morally problematic” we mean that the fact that punishment is sometimes weakly distributive in its effects constitutes a *pro tanto* reason against the permissibility of punishing a collective. Even proponents of collective punishment acknowledge that collective punishment is morally problematic in some way (French 1984, 190; Erskine 2010). Pasternak (2014, 230), for example, writes that “some of the side effects of corporate task responsibility are inherently unfair, and that we ought to do what we can to mitigate them.” (It should be noted, however, that punishment-effects are not merely “side effects” as this quote suggests. Rather, punishment-effects are arguably *constitutive* of the act of punishing).

What makes collective punishment morally problematic? There are different answers to this question that are compatible with one another. The most general answer, on which we will concentrate here, is that the infliction of punishment-effects on individuals is not justified. More specifically, the act of punishing the group might lack justification in that individuals’ having to bear punishment-costs is unfair (Pasternak 2014). This objection of unfairness has a broadly contractualist flavor insofar as it rests on the idea that the distribution of punishment-effects is not justifiable in terms of principles that could not be reasonably rejected. On a more deontological approach, a lack of justification could be due to the fact that some individuals bear punishment-costs although they are not morally responsible for the thing for which the group is punished. Although the group might meet the necessary conditions for responsibility and punishment, the group members might not. For example, not all members that have to bear punishment-effects caused or intended the offence for which they seem to be held responsible.

There is no need to spell out these arguments in detail because we do not aim to object against any of them specifically. Instead, we take a constructive approach. We sketch ways to provide the justifications that these arguments claim are lacking, to thereby argue that the arguments fail to establish that collective punishment, when punishment-effects are borne by individuals, is impermissible.

The challenge for collective punishment consists in the absence of a justification for the imposition of punishment-effects. We think this challenge can be met. Other authors have already explored possible avenues. Pasternak (2014) specifically, explores two kinds of justifications for the imposition of punishment-effects in collective punishment, which she calls the “membership-based duties approach” and the “side-effects approach”. With our paper we offer further alternatives to supplement those identified by Pasternak (2014). We spell out in additional detail the membership-based approach.

At this point, it will be helpful to distinguish more precisely, what exactly needs to be justified. Three elements are relevant to the justification of punishment-effects. First, the nature of the effects might be relevant. The punishment-effects might consist in actual monetary cost but also in any other form of disadvantage. Second, the pattern of the distribution of punishment-effects might be relevant. Pasternak (2011b) distinguishes broadly between three patterns of distributing punishment-effects. Punishment-effects can (a) be distributed in proportion to the individual's causal contribution to the overall outcome for which the group is punished, (b) be distributed equally among all members, (c) be distributed randomly. We add a fourth option to this list: (d) punishment-effects can be distributed at the group's discretion. In both (c) and (d), the punishment effects are distributed neither proportionally nor equally, and in (d) they are also not distributed randomly in accordance with a fair procedure. Rather, the group gets to decide how to distribute them, with the limit case being deliberate scapegoating. Third, also of relevance might be the way by which a decision was made as to how the punishment-effects are to be distributed. In this paper, we focus on the distributional pattern of the effects and—to the extent that there is a choice between different distributional patterns—on cases in which such a decision between different distributional patterns is made by a strict subset of group members on behalf of all members of the group.

In sum, we have now a clearer view of the scope and the structure of the challenge. The scope of the challenge is narrow because punishment is not necessarily distributive. The structure of the challenge is a demand for justification. Punishment-effects need to be justified in those cases in which punishment is in fact distributive. It is important to note the relatively low bar of this challenge. Because punishment is only contingently distributive, all that is required is a contingent justification for punishment-effects being borne individuals. Moreover, justifications can add up, they need not be exclusive. Hence, the bar for justification is not only low, it is also long: there might be several contingent justifications for why punishment-effects may be borne by individuals that are compatible with one another. We take such a pluralistic approach to the justification of collective punishment in this paper.

4 Justifications from Accounts of Collective Action

In this section, we present three different accounts of collective actions. We show that on all three kinds of account, the distribution of punishment-effects from collectives to members is permissible.⁹ Because readers are likely to accept at least one of these kinds of account, we take this to be a kind of proof by exhaustion. We will look at accounts that rest chiefly on participatory intentions (Kutz 2000), joint actions (Bratman 2014; Shapiro 2014), and group agents (Gilbert

⁹ We bracket the question of what it takes for an individual member of a punishable collective to be excused from a share of distributed punishment-effects here, but see discussion in (Lawford-Smith forthcoming, chap. 6 Sec V).

2006; List and Pettit 2011). The accounts are ordered from weak to strong in the sense that the latter entail the former. Joint actions entail some kind of joint or participatory intention, and group agents are created by certain joint actions (Pettit and Schweikard 2006). A quick clarification, before we start: some accounts of collective action are singular—one (collective) agent performs an action—while others are plural—a group of (individual) agents perform an action together. It is not clear whether punishment-effects are technically distributive, in the sense just discussed, on the latter. They do not distribute ‘from’ a collective agent ‘to’ members. Rather, they bear on members together *as* the punishment. We bracket this technicality in what follows and assume that the punishment is directed at the group (whether singular or plural) and has effects on members.

4.1 Participatory Intentions Accounts

The account put forward by Kutz (2000) rests on the notion of a “participatory intention”, which Kutz takes to be the “common element” to all forms of collective action. Before we get to participatory intentions as such, there are three prior conditions that are necessary for joint actions on the participatory intentions account. First, an agent’s intentions need to exhibit *strategic responsiveness*, that is, her intentions need to be sensitive to “beliefs or predictions about what others intend to do”. Second, members need to have a *shared goal*. Kutz writes that “there is at least one token activity or outcome, involving the actions of the other, whose performance or realization would satisfy the intentions of each”. Third, members need to exhibit *mutual openness*, that is “each must also be favorably disposed towards the other’s possible knowledge of this strategic sensitivity”. These necessary conditions, or conditions of a similar nature, can also be found in alternative accounts of joint actions (cf. Pettit and Schweikard 2006; Bratman 2014).

Fourth, and finally, the participatory intentions account requires, of course, that each member of a group engaged in a joint action has a *participatory intention*. That is, each must see herself as a member of a group and as a participant to the actions of the group. She must see her action as a contribution towards the group’s end. Kutz stipulates further that members must not be deluded about matters of fact figuring into the participatory intention. For example, a member must not see her action as a contribution when in fact no one else shares the end towards which she contributes.

The participatory intentions account extends to hierarchical groups. An individual can participate in an end over which others have ultimate authority, e.g. those higher up in the hierarchy of the relevant group. Accordingly, Kutz distinguishes between *executive intentions* whose “content is an activity or outcome conceived of as a whole” and *subsidiary intentions* which aim at achievement of a part of the total outcome or an activity and that are “generated and rationalized by an executive intention”. For example, consider an orchestra performing the joint action of giving a concert. While the cello player has only the subsidiary intention to perform her cello part, the conductor has the executive intention that the orchestra play the piece of music. We might

think of executives' and board of directors' intentions over corporations' ends as executive intentions similar to the conductor's, whereas we think of employees' intentions over corporations' actions as subsidiary intentions similar to the cello player's.

Specifically, the participatory intentions account extends to corporations or states. The account extends to a great number of individuals as members of such groups. Many if not all employees of a corporation will have subsidiary participatory intentions in performing the actions required of them in their employee role, which further the corporation's ends without having any discretion over these ends. Many if not most government employees will have subsidiary participatory intentions in performing the actions required of them in their formal role, e.g. contributing to the running of the particular division of whatever department or agency they are a part of. By contrast, most if not all employees and elected officials in the upper echelons of government will have executive participatory intentions over what the state does. If an elected official or employee were to disagree with her state's ends, she may cease to perform the normative, expressive, or practical actions of associated with her formal role, thereby ceasing to hold the participatory intention.

Despite its relaxed conditions for involvement in joint agency, the participatory intentions account is not entirely individualist about responsibility. Individuals can be responsible together for what they do together. Groups whose members meet the conditions of the account have joint agency.¹⁰ These groups are the kinds of things that can be the objects of claims and obligations by other individuals and other collectives. Similarly, a group of people with overlapping participatory intentions, focusing on its normative role and setting aside metaphysical questions, can be an object to be punished for their joint action.

The participatory intentions account can help to justify that members of such a group may come to bear punishment-effects. Since each member accepts a shared goal and sees herself as participating in the pursuit of this goal via her participatory intention, punishment-effects may be justifiable though each individual's acceptance of this shared goal and her participatory intention. If this shared goal involves the offence for which the group was punished, then each group member, in virtue of sharing this goal, can be made to bear the punishment-effects that result from the punishment of the group. This relates to the *mens rea* requirement of individual punishment. When a group member furthermore was open to responding strategically and saw herself as participating in that goal, this gets close to also meeting the *actus reus* requirement. For example, if a road construction team illegally destroys part of a protected forest in order to access their construction site more easily, then they together can be punished for their destruction of the forest. They might be made to pay a fine. This, however, does not mean that the punishment can be

¹⁰ Kutz himself he does not want to call this "collective agency" because of his reductionist commitments about collective-level explanations. He claims that "groups are nothing more or less than agents who intend to participate in collective action" (Kutz 2000, 31).

directed against individual members. The members are punished together for what they have done together.

Even if the shared goal does not involve the offence for which the group was punished, group members can still be made to bear punishment-effects in virtue of sharing the group's goal in one specific kind of case. The actions of distributing and bearing punishment-effects can be part of the shared goal, even if the offence for which a group is punished is not part of this goal. The actions of distributing and bearing punishment-effects can be part of the conditions that group members accept with their participatory intentions. In this way, participatory intentions can provide a justification for directing punishment-effects to group members, even if they did not have a participatory intention with respect to the offence. Participatory intentions cannot justify directing punishment-effects at group members when either the shared goal is unrelated to the offence, or when the shared goal does not include the distribution of punishment-effects. In short, on the participatory intentions account, collective punishment (that is weakly distributive) is sometimes justified, and sometimes not.

This way of thinking about punishment in the case of groups on the participatory intentions account vaguely resembles the rationale behind the guidelines for joint enterprise in the UK. Of course, there is an open question about thresholds: does any weak participatory intention without any causal contribution to the shared goal provide sufficient justification for an individual to bear punishment-effects? This is the question that has been debated in the UK legal context. For our argument we can leave open the question concerning a threshold of participation.

Assuming now that a decision can be made with respect to how the punishment-effects are to be distributed, which patterns of distribution are permissible for a group, as seen from the perspective of the participatory intentions account? We think that a distribution of punishment effects in proportion to each individual's causal contribution to the group outcome is most immediately justifiable. Although a corporation may be punished even if it delegates punishment-effects fully to low-level employees (given further possible justifications and in compatibility with a side-constraint that we outline below), a distribution in proportion to causal contribution seems to be easily conceptually accommodated by the participatory intentions account and seems to align with the guidelines of joint enterprise sentencing, which it resembles. Along the lines of the distinction between executive and subsidiary intentions, we might distinguish two tiers of responsibility. This opens up the possibility to distribute punishment-effects not proportional only to causal contribution, but proportional also to the type of contribution members made. Finally, insofar as the actions of distributing and bearing punishment-effects are part of the shared goal of a group, and given that this shared goal is accepted by each, participatory intention accounts have the resources to justify distributions of punishment-effects that are disproportionate to individual contributions.

4.2 Joint Action Accounts

A more specific account of joint action for large-scale hierarchical groups is put forward by Shapiro (Shapiro 2011, 2014). Shapiro amends the celebrated account of joint action due to Bratman (1992, 2014), by relaxing Bratman's general assumption of "modest sociality" and specifically the condition that relations between group members must be symmetric. Hierarchical groups violate this symmetry assumption because in hierarchical groups members essentially stand in an asymmetric structural relation to one another: some are higher up than others. Accordingly, Shapiro allows an asymmetry in members' responsiveness to one another instead of requiring mutual responsiveness as in Bratman's original account. Compared to the general participatory intentions account, this account of joint actions of hierarchical groups, at least in the abstract, is a refinement.

The central innovation of this account of joint action is that it rests on the notion of a "plan". Plans play an important role in practical reasoning (Bratman 1987). Importantly, plans have structure. Not only do plans structure agency over time, plans can also coordinate collective activities. Bratman has extended plans from individual to collective settings by introducing the idea of meshing subplans. Shapiro now extends this idea further to hierarchical groups with "vertically interlocking subplans". This account of joint action, while still relying centrally on the notion of a plan, allows for a horizontal as well as a vertical division of labor.

The role of authorities and superiors is even more pronounced in this account of joint action than it is in the participatory intentions account. First, this can be seen in the how authorities not only set the ends for a group unilaterally, as they may also in the participatory intentions account with their executive intentions, but the authorities may moreover delegate tasks and decide how the different subplans that individuals follow work together. In the words of Shapiro, authorities "are 'mesh-creating' mechanisms". The delegation and assignment of individual roles are captured in the plan that is made by those higher up but shared by all. Second, in contrast to participatory intentions, which suggest a certain kind of identification with a group, the way in which a group member thinks of her own membership and role in a group is more transactional on Shapiro's account. Participants to a joint action can even be alienated from the group activity (Shapiro 2014).

With the notion of plans, and the power of authorities to delegate tasks and specify subplans, this account provides further justification for the distribution of punishment-effects to group members. As with the participatory intentions account and its distinction between executive and subsidiary intentions, the notion of plans being determined by authorities helps to justify a distribution of punishment effects in proportion to individuals' causal contributions. That authorities can create subplans seems to relate closely to the distinction between principals and accessories known from joint enterprise law.

In addition, going beyond the participatory intentions account, the planning account of hierarchical groups can render permissible forms of what we called "random" distributions of

punishment-effects. This is because a group may plan for how punishment-effects are distributed. Just as stocks or bonds in a corporation may come with different privileges and liabilities, so membership and participation in a group may have different liability-exposures specified as part of the sub-plans attached to these roles. The plans can hence serve as a vehicle to specify in advance members' liability, and accordingly, any distribution of punishment-effects for the eventuality that the group may be punished for its conduct. Whereas the plans specify the substance of the justification, what is the form of this justification? The distribution of the punishment-effects is essentially justified in virtue of members voluntarily accepting their plans. This does not, in contrast to the participatory intentions account, involve that the members make the group's end their own end. But it requires that the members voluntarily play their role. Of course, this notion of voluntariness would need to be spelled out further. Nevertheless, that acceptance of a role in a group, even if it is only for monetary incentives and not out of genuine identification, can provide grounds to distribute punishment-effects in line with what group members accepted.

4.3 Group Agency Accounts

Accounts of group agency, as we understand the term, differ from accounts of joint actions in that the former give an account of actions that are the group's own whereas the latter give an account of actions performed by several individuals together (cf. Pettit and Schweikard 2006). We are interested in examining how accounts of group agency, such as that of French (1979), List and Pettit (2011), and to some extent that of Gilbert (1992, 2006), may provide further justification for different distributions of punishment-effects on group members.

The feature of group agents (in contrast to shared agency or joint actions) that we want to focus on here is that in group agents there is an agent 'over and above' the individual agents. This agency has also a normative dimension. Group agents may be responsible as a group (Pettit 2007; List and Pettit 2011, chap. 7). Group agents may enter into contracts and thereby incur rights and obligations as a group. A corporation, as one example of a group agent, has in the words of Ciepley (2013) "contractual individuality" that "separates the property and liabilities of the corporation from its members".

This moral or normative standing of a group agent is important along two fronts: externally and internally. First externally, a group agent is a distinct entity that stands in normatively relevant relation to other individuals or groups. Likewise, moral predicates might apply to the group but not to its members. As Gilbert (2006, 113) observes, for example, it might well be true that members of a group agent can truthfully assert "we are to blame" thereby including themselves as one of the "we", "without implying her personal blameworthiness". Second internally, a group agent has a normative standing vis-à-vis its members. This standing is investigated in the literature in political theory, for example, on the question of on what grounds citizens ought to pay debts incurred by their state (Parrish 2009; Pasternak 2011a; Stilz 2011; Pasternak 2011c). One answer

that might be given is that the state has the right normative standing vis-à-vis its members. This is not the place to determine what this right standing vis-à-vis the members consists in. Likely, what the right standing is that allows for members to bear punishment-effects for offences committed by their group agent, depends crucially on the kind of association that forms the group agent. Concretely, the normative standing of a group agent vis-à-vis its members differs depending on whether the group agent is a state, a corporation, or a religious association.

The important point that we want to carry home is the following: Group agents, in virtue of their existence as agents – in a normative or moral sense – over and above their members, allows the group to stand in normatively meaningful relations not only to outside agents but also internally vis-à-vis its members. It is these relations between members and the group agent as a further entity that can provide various justifications for the distribution of punishment-effects to members.

This insight differentiates what Pasternak (2014) called “associative obligations” under her membership-based approach. Whereas accounts of shared agency—such as those based on participatory intentions or those based on plans—characterize the standing that group members may have towards one another, accounts of group agency also make room for a separate normatively relevant relation to obtain between the group agent as such and its members.

In terms of their relation to the examples from criminal law, the accounts of group agency relate, of course, to the law on corporate manslaughter in that it can be a group agent itself that is responsible or punishable. The way in which the effects of the punishment are distributed within members of associations is an internal matter of the group agent. The distributions that can be chosen within a group agent might be morally better or worse, some might in specific cases not be justified at all. But group agents as such can be set up in such a way that they have an internal decision structure—be it consensus or majoritarian or dictatorship or something else entirely—by which they can create justifications for their members to accept distributions of punishment-effects. The justification of this distribution of punishment-effects is derivative. The distribution of the effects will be justified in virtue of the decision mechanism being justified by which the distribution was determined. Clearly, however, many group agents have internal decision mechanisms that are sound in the sense that the outputs that these mechanisms generate come with an obligation on the members to be followed. In this sense, in principle, any pattern of distribution—proportional, equal, random, or discretionary—can be justified in the context of a group agent.

5 Two Separate Domains

What we have said so far suggests that, depending on the kind of group under investigation—a looser group based on participatory intentions, a structured group with hierarchical planning, or a group agent—different justifications for distributions of punishment-effects are available. Whether

these justifications are in fact available will depend on each specific case and group in question. What we have shown is only that there are various plausible justifications that are captured by the accounts of collective agency. These justifications render various distributions of punishment-effects permissible. Each account can justify distributing punishment-effects unequally or disproportional to individual contributions. It should be noted that we only make a claim about permissibility and not about the relative strengths of justifications.

There is a broader picture that emerges from these investigations. This picture is that there are two domains of justice when it comes to the permissibility of punishing groups. There is the external domain concerned with the question of whether wrongful collective action renders a group punishable. Then there is the internal domain concerned with the question of whether and how the distribution of punishment-effects to group members is justified. The two domains are not independent in the sense that considerations about internal justice can never impact considerations about external justice. However, as we discussed above, there is no argument in principle about internal justice that would place necessary restrictions on external justice, in contrast to what some authors have suggested. Importantly, the fact that a group is punishable for an offence might persist even if there is insufficient justification on the issue of internal justice to distribute the punishment-effects in any particular way. And whether or not a distribution of punishment-effects is justified depends largely on inherent properties of the group and its members. This is the sense in which the two domains are independent.

However, this claim about the independence of the two domains is subject to an important constraint that we need to mention before concluding. The impetus of this constraint is that certain protections are needed to safeguard individuals against subordination to the group (a version of this claim can be found in Susan Okin's critique of Will Kymlicka's justification of multiculturalism (Okin 1999; Kymlicka 1995)). We suggest that the protection of individuals' negative human rights be understood as a constraint upon any internal distribution of punishment.

As an example, consider the following scenario. Suppose a corporation has been involved in the contamination of a water supply. The after-hours cleaner for a large corporation, involved neither in the corporation's decision to perform that particular action, nor the performance of the action itself, is scapegoated to take the public fall for the wrongful action performed by the corporation. He will be required to publicly admit fault, and to resign from his position. The corporation, to demonstrate remorse for its employment of the cleaner, will make a substantial donation to a charitable organization doing medical work in the region to help victims of the harm for which it has been punished.

We can distinguish two ways in which the cleaner can be made to resign from his position. In the first way, the cleaner is extorted, threatened, or otherwise exploited by someone in the corporation to take the fall on its behalf. This undermines, rather than involves, the cleaner's autonomy, and violates his legal rights (not to be blackmailed) and his moral rights (not to be used

as a mere means for ends which are not his own). In the second way, the cleaner is approached and made an attractive offer, for example, in exchange for his public admission of fault and resignation, he will be given a substantial remuneration package sufficient to support him and his family for several years (more than long enough for him to secure further employment). The second offer, as long as it is not an “offer he can’t refuse”, is not problematic in the same way as the first, because it does not involve a violation of the cleaner’s negative rights. So long as such constraints are observed, any internal distribution of punishment by the collective among its members is permissible, given the different avenues of justification outlined above.

6 Conclusion

In this paper, we have addressed a central worry that is brought forth against the idea of collective punishment. We have reconstructed the worry as an argument, we have found that the argument is less strong than it initially appeared, and we have pointed out different ways in which the weaker challenge, that the argument in fact poses, can be met.

Concretely, the worry is that collective punishment is always morally problematic, and can hardly be justified, because of the effect that collective punishment has on members of a group. A group may be punished but individuals will bear the costs and the effects of this collective punishment. However, there seems to be no justification for why individuals should be made to bear the punishment-effects. This imposition of significant effects without justification is morally problematic, or so goes the worry against collective punishment.

We have argued that this worry is not as powerful as it may seem. This is because punishment is not necessarily distributive. Collective punishment is, hence, not morally problematic in principle. Instead, collective punishment is only morally problematic in cases where no justification is available for individuals bearing punishment-effects. We have added to the existing literature by sketching different ways in which such justifications can be provided out of different accounts of collective action. We have broadly distinguished between three families of accounts of collective actions: those based on participatory intentions, those based on shared plans in hierarchical groups, and those based on the idea of group agency in both hierarchical and egalitarian groups. We have argued that each account—participatory intentions, plans, and the group agency—provides justifications for the distribution of punishment-effects.¹¹ Whereas the first two families of accounts of collective actions resemble the legal models of joint punishment in the criminal law literature (which are individualist), the latter model of group agency resembles the legal model of corporate manslaughter (which is collectivist).

¹¹ Cf. Lawford-Smith (*forthcoming*), who accepts only the third of these—collective agency—as a justification for all of collective responsibility, culpability, blame, and punishment.

There is the question of external justice—whether a collective is punishable and there is the question of internal justice—how the resultant distribution of punishment-effects can be justified. Subject to a constraint concerning individual autonomy and human rights, which we have illustrated above, these two domains are largely independent in a specific way. The external justifications for punishment are not undermined by internal worries about the distribution of punishment-effects. Furthermore, different compatible internal justifications are available. The justification of a distribution of punishment-effects among group members might be encapsulated in the collective end of a participatory intention, in the content of a shared plan, or in the moral standing that a group agent has vis-à-vis its members. The challenge to collective punishment is not only weaker than it initially appeared, it can also be met in a variety of ways.

7 References

- Bratman, Michael. 1987. *Intention, Plans, and Practical Reason*. Harvard University Press.
- . 1992. “Shared Cooperative Activity.” *The Philosophical Review* 101 (2): 327–41. <https://doi.org/10.2307/2185537>.
- . 2014. *Shared Agency: A Planning Theory of Acting Together*. Oxford: Oxford University Press.
- Ciepley, David. 2013. “Beyond Public and Private: Toward a Political Theory of the Corporation.” *American Political Science Review* 107 (01): 139–58. <https://doi.org/10.1017/S0003055412000536>.
- Coffee, John C. 1981. “‘No Soul to Damn: No Body to Kick’: An Unscandalized Inquiry into the Problem of Corporate Punishment.” *Michigan Law Review* 79 (3): 386–459. <https://doi.org/10.2307/1288201>.
- Corporate Manslaughter and Corporate Homicide Act*. 2007. <https://www.legislation.gov.uk/ukpga/2007/19/contents>.
- Crown Prosecution Service. 2012. “CPS Guidelines on Joint Enterprise Charging Decisions.” http://www.cps.gov.uk/legal/assets/uploads/files/Joint_Enterprise.pdf.
- Duff, Antony, and Zachary Hoskins. 2018. “Legal Punishment.” Edited by Edward N. Zalta. *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2018/entries/legal-punishment/>.
- Erskine, Toni. 2010. “Kicking Bodies and Damning Souls: The Danger of Harming ‘Innocent’ Individuals While Punishing ‘Delinquent’ States.” *Ethics & International Affairs* 24 (3): 261–85. <https://doi.org/10.1111/j.1747-7093.2010.00267.x>.
- Feinberg, Joel. 1965. “Expressive Function of Punishment.” *The Monist* 49 (3): 397–423.
- . 1968. “Collective Responsibility.” *The Journal of Philosophy* 65 (21): 674–88. <https://doi.org/10.2307/2024543>.

- French, Peter A. 1979. "The Corporation as a Moral Person." *American Philosophical Quarterly* 16 (3): 207–15.
- . 1984. *Collective and Corporate Responsibility*. Columbia University Press.
- Gilbert, Margaret. 1992. *On Social Facts*. Princeton University Press.
- . 2006. "Who's to Blame? Collective Moral Responsibility and Its Implications for Group Members." *Midwest Studies In Philosophy* 30 (1): 94–114. <https://doi.org/10.1111/j.1475-4975.2006.00130.x>.
- Jackson, Frank. 1988. "Understanding the Logic of Obligation." *Proceedings of the Aristotelian Society, Supplementary Volumes* 62: 255–83.
- Kutz, Christopher. 2000. "Acting Together." *Philosophy and Phenomenological Research* 61 (1): 1–31. <https://doi.org/10.2307/2653401>.
- Kymlicka, Will. 1995. *Multicultural Citizenship: A Liberal Theory of Minority Rights*. Clarendon Press.
- Lawford-Smith, Holly. Forthcoming. *Not In Their Name: Are Citizens Responsible for Their States' Actions?* Oxford: Oxford University Press.
- . 2018. "What's Wrong with Collective Punishment?" *Proceedings of the Aristotelian Society* CXVIII (3). <https://www.aristoteliansociety.org.uk/pdf/lawford-smith.pdf>.
- List, Christian, and Philip Pettit. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.
- McGee, Patrick, and David J Lynch. 2017. "US Court Approves VW's \$4.3bn Emissions Penalty." *Financial Times*, April 21, 2017. <https://www.ft.com/content/51537baa-26a3-11e7-8691-d5f7e0cd0a16>.
- Okin, Susan Moller. 1999. *Is Multiculturalism Bad for Women?* Edited by Joshua Cohen, Matthew Howard, and Martha Craven Nussbaum. Princeton, N.J: Princeton University Press.
- Parrish, John M. 2009. "Collective Responsibility and the State." *International Theory* 1 (01): 119–154. <https://doi.org/10.1017/S1752971909000013>.
- Pasternak, Avia. 2011a. "The Collective Responsibility of Democratic Publics." *Canadian Journal of Philosophy* 41 (1): 99–123. <https://doi.org/10.1353/cjp.2011.0002>.
- . 2011b. "The Distributive Effect of Collective Punishment." In *Accountability for Collective Wrongdoing*, edited by Tracy Isaacs and Richard Vernon, 210--230. Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511976780.009>.
- . 2011c. "Sharing the Costs of Political Injustices." *Politics, Philosophy & Economics* 10 (2): 188–210. <https://doi.org/10.1177/1470594X10368260>.
- . 2014. "The Impact of Corporate Task Responsibilities: A Comparison of Two Models." *Midwest Studies In Philosophy* 38 (1): 222–31. <https://doi.org/10.1111/misp.12025>.
- Pettit, Philip. 2007. "Responsibility Incorporated." *Ethics* 117 (2): 171–201.

- Pettit, Philip, and David Schweikard. 2006. "Joint Actions and Group Agents." *Philosophy of the Social Sciences* 36 (1): 18–39. <https://doi.org/10.1177/0048393105284169>.
- Sentencing Guidelines Council. 2015. "Health and Safety Offences, Corporate Manslaughter and Food Safety and Hygiene Offences – Definitive Guideline." <https://www.sentencingcouncil.org.uk/wp-content/uploads/HS-offences-definitive-guideline-FINAL-web.pdf>.
- Shapiro, Scott. 2011. *Legality*. Cambridge: Harvard University Press.
- . 2014. "Massively Shared Agency." In *Rational and Social Agency: The Philosophy of Michael Bratman*, edited by Manuel Vargas and Gideon Yaffe, 257. Oxford: Oxford University Press.
- Smith, Thomas H. 2009. "Non-Distributive Blameworthiness." *Proceedings of the Aristotelian Society (Hardback)* 109 (1pt1): 31–60. <https://doi.org/10.1111/j.1467-9264.2009.00257.x>.
- Stilz, Anna. 2011. "Collective Responsibility and the State." *Journal of Political Philosophy* 19 (2): 190–208. <https://doi.org/10.1111/j.1467-9760.2010.00360.x>.
- Trejos, Nancy. 2018. "Number of International Visitors to the USA Declines." *USA Today*, January 19, 2018. <https://www.usatoday.com/story/travel/roadwarriorvoices/2018/01/19/number-international-visitors-usa-declines/1047492001/>.
- Wringe, William. 2016. *An Expressive Theory of Punishment*. Springer.