Group Blameworthiness and Group Rights

Stephanie Collins

This is the final draft of a paper whose definitive version will be published in Inquiry.

Abstract

The following pair of claims is standardly endorsed by philosophers working on group agency: (1) groups are capable of irreducible moral agency and, therefore, can be blameworthy; (2) groups are not capable of irreducible moral patiency, and, therefore, lack moral rights. This paper argues that the best case for (1) brings (2) into question. Section 1 paints the standard picture, on which groups' blameworthiness derives from their functionalist or interpretivist moral agency, while their lack of moral rights derives from their lack of sentience. In Section 2, I add support to a recent argument that this standard picture needs alteration: groups' blameworthiness requires something akin to sentience, which groups acquire from members. Section 3 discusses rights: if groups acquire sentience from members, as Section 2 argues, then can groups have moral rights? I argue that they can, but that groups have only a limited range of moral rights, whose existence depends on (without being ontologically or justificatorily reducible to) the attitudes and actions of humans.

Key Words

Group agency, group blameworthiness, group rights.

Introduction

The following are widely accepted by those working on group agency: (1) groups are capable of irreducible moral agency; (2) groups are not capable of irreducible moral patiency.[1] 'Moral agency' implies the capacity to act, bear moral reasons, reason about morality, acquire duties or obligations, and be blameworthy or praiseworthy. 'Moral patiency' implies the capacity to be benefited, be a source of moral reasons, be an object of concern for those who reason about morality, bear rights or entitlements, and be respect-worthy or consideration-worthy.

Why the consensus? Because many authors accept either functionalism or interpretivism about moral agency. Functionalism and interpretivism allow groups to be moral agents in a non-reductionist way—or so many argue. But neither functionalism nor interpretivism produce moral *patiency* (so the thought goes), because they don't give any essential role to sentience or phenomenal consciousness. In short: groups are not sentient; therefore, they are not irreducible moral patients; but sentience is not required for moral agency; therefore, groups can be irreducible moral agents. I will argue for (1) but against (2).

On the agency side, I focus on blameworthiness. On the patiency side, I focus on (what I will call) fundamental rights. I take this narrow focus because the above-listed aspects of

[1] On (1), see French 1984; Copp 2006; List and Pettit 2011; Huebner 2014; Epstein 2015; Tollefsen 2015; Björnsson and Hess 2017; Hindriks 2018; Collins 2019; Collins 2023. On (2), see Manning 1984; Ozar 1985; Raz 1986: 51; Kymlicka 1995; List and Pettit 2011: 179-182; Pasternak 2017; Hedahl 2017. A recent exception to (2) is Silver 2019, whose argument differs sharply from mine.

moral agency might not rise and fall together; likewise the above-listed aspects of moral patiency.[2] To avoid conflating different aspects, I focus on just these two.

I assume an entity is 'blameworthy' just in case it is apt to blame that entity. Blame, I'll assume, is a bundle of negative reactive attitudes such as anger, resentment, and indignation (Strawson 1962). It might be *apt* to blame an entity even if, all things considered, one *ought not* to blame the entity—for example, if blaming the entity would produce bad consequences or if no one has standing to blame. Blameworthiness, in this sense, is a purely backward-looking assessment of the entity. I argue—as is now standard—that groups agents can be blameworthy. However, my argument is not straightforwardly functionalist or interpretivist. This will have important implications for groups' fundamental rights.

I assume 'fundamental rights' are rights that are not justified via reference to other entities' rights or interests. Whether a right is 'fundamental' is a question of the right's *justification*, not its *ontological status* (as real or ontologically irreducible). For example, consider a corporation's right to buy property. This right is not fundamental, if it is *justified* purely via the rights or interests of humans. The corporation's right might nonetheless be *ontologically* irreducible: the right might not *be* a collection of humans' rights, for example because there are numerous different ways of 'realising' the group's right in attributes of individuals. And if humans' rights or interests were *always* the *entire* justification for groups' ontologically irreducible rights, then groups would not have any *justificatorily fundamental* rights. In what follows, I take it for granted that groups have some *ontologically irreducible* rights, such as corporations' rights to enter contracts. The question is whether these are

[2] For example, Collins (2019: ch. 3; forthcoming) argues non-organized groups can bear *abilities*, but not *duties*, though both are arguably aspects of agency.

justificatorily fundamental, that is, justified by the moral value of the group itself. I answer affirmatively. [3]

Section 1 paints the standard picture, on which groups' blameworthiness derives from their functionalist or interpretivist moral agency, while their lack of fundamental rights derives from their lack of sentience. In Section 2, I add support to a recent argument that this standard picture needs alteration: groups' blameworthiness requires something akin to sentience, which groups acquire from members. Section 3 discusses rights: if groups acquire sentience from members, as Section 2 argues, then can groups have fundamental moral rights? I argue they can, but that groups have only a limited range of fundamental moral rights, whose existence depends on (without being ontologically or justificatorily reducible to) the attitudes and actions of humans.

[3] Ontological irreducibility and justificatory fundamentality are not always clearly distinguished in the literature on groups' rights. But the standard view is that groups' rights are ontologically irreducible—yet justificatorily *non*-fundamental. For example, Lovett and Riedener (2021: 223) argue that "As individuals, we have distinctive interests in participation in successful group agency. To properly respect this interest, we must treat organization as if they had … moral status. … We must keep our promises to them and show them gratitude to respect the claims of their members." On this view, groups *themselves* are owed promises or gratitude (ontologically irreducible group rights), but this is *justified* via individuals' interests (justificatorily non-fundamental group rights). Similarly, those cited on (2) in fn. 1.

1. Groups' Blameworthiness and Groups' Rights: The Standard Picture

On the functionalist picture of agency, attitudes such as belief, desire, hope, regret, and so on are characterised by the function those attitudes play within some system. This system is the entity that bears the attitude. The functions might include causing, grounding, or blocking other attitudes and behavioural dispositions. An entity with functionalist agency has internal mechanisms that ensure (or at least, are oriented towards ensuring) that the entity can arrive at attitudes that interact in a coherent way, such that the entity tends not to believe contradictions, tends to desire the means to its ends, and so on.

Many groups have mechanisms that are oriented towards precisely such coherence. The mechanisms take a variety of forms—such as committees, compromises, commands, consultation, and so on—that allow the production of group-level attitudes that (more or less) cohere with one another. As has been emphasised at length in the literature, the long-run operation of such mechanisms can mean that the group ends up holding attitudes that no members hold. The group's attitudes are determined by the internal mechanisms, which can be so complex as to allow sharp departures from the private attitudes of members. Such 'functionalist' group agents can be *moral* agents if the internal mechanisms render the group able to reach more-or-less correct attitudes towards moral propositions.[4]

On the interpretivist picture of agency, the question is whether it's possible (or rational, or reasonable) for an external observer to adopt the 'intentional stance' towards an entity. From within the intentional stance, one views the entity's behaviours as explained by its beliefs, desires, hopes, fears, regrets, and so on. The interpretivist view renders an entity an agent if

[4] French 1984; Copp 2006; Huebner 2014; Epstein 2015; Björnsson and Hess 2017; Hindriks 2018; Collins 2023.

adopting the intentional stance enables good explanations and predictions of the observed behaviour of the entity.

Again, many groups are agents on the interpretivist picture. For example, when we adopt the intentional stance to an organisation such as a bank, we are able to explain and predict its behaviour: we can explain and predict a bank by attributing to it certain motives (e.g., profit) and beliefs (e.g., about what customers wants). Furthermore, on the interpretivist picture, we can attribute *moral* agency to a group if the group's behaviour is well-rationalised by attributing to the group the capacity for correct moral attitudes.[5] Of course, one might think none of a bank's behaviour is well-rationalised by attributing to it the capacity for correct moral attitudes. But even the most cut-throat for-profits do play lip-service to correct moral attitudes, for example when they emphasise 'ethics' their mission statements and advertising. This behaviour is well-explained by attributing to them the *capacity* for correct moral attitudes, even if they do not often *exercise* that capacity.[6] And the capacity is all that's needed to be a moral agent, with the propensity to be blameworthy.

When viewing groups as moral agents, we need not 'pick' between functionalism and interpretivism. For example, in their influential book *Group Agency*, Christian List and Philip Pettit appeal to interpretivism at some points (e.g., List and Pettit 2011: 11, 13, 23) and functionalism at other points (e.g., List and Pettit 2011: 171). They start with a general picture of agency that is interpretivist: a picture that focuses on an entity's *outward performances*. Yet the bulk of their discussion concerns the *internal mechanisms* of group agency, which are more the province of functionalism.

[5] The interpretivist account is applied to group agents by Tollefsen (2015).

[6] On collective moral capacities, see Hindriks 2018.

In fact, when we are assessing whether groups are irreducible agents, it is useful to assess them *both* by functionalist lights *and* by interpretivist lights. After all, if groups can satisfy both the functionalist and the interpretivist 'tests,' then our arguments for their irreducible agency are stronger than if group passes just one of those tests.[7] On the functionalist test, groups' irreducible agency derives from the fact that the group's internal mechanisms produce attitudes that can depart systematically and consistently from the attitudes of members. On the interpretivist test, their irreducible agency derives from the fact that we must view the group as a whole agential system before we can explain and predict its behaviour. The tests are overlapping and complementary. Groups that pass both tests have a stronger case in favour of their irreducible agency.

Arguments about groups' moral *patiency* take a different starting point. Central here is the moral value of sentience or experience.[8] As Kendy Hess says when rejecting groups' fundamental rights, humans are vulnerable to

> certain kinds of *experiences*, both painful and pleasurable, and it is precisely this vulnerability that accounts for the rights and protections that are typically awarded … Without the possibility of hunger, humiliation, and hatred, it wouldn't really matter

[7] This is not to say that an entity *must* pass both tests to be an agent; it's simply to say that our case for an entity's agency is *stronger* if it passes both tests (see also Collins 2023, ch. 1). One could also introduce other tests, such as representationalist tests, which Kendy Hess argues group agents can pass (2018: 70). On functionalism versus interpretivism about group agency, see Strohmaier 2020.

[8] On the moral importance of sentience in non-group contexts, see Singer 1993; McMahan 2002.

> whether a person's property rights were acknowledged, her voice heard, her decisions and bodily integrity respected. (2013: 333, emphasis added)

As Hess puts it, "There is no phenomenal point of view to accompany their [groups'] rational point of view." (2013: 334) Therefore, groups lack fundamental irreducible rights.

Likewise, List and Pettit briefly discuss whether group agents have moral worth equal to humans (2011: 179-182). They reject this, writing in an endnote that "[w]e take it as a working assumption that group persons do not have whatever functional characteristic it is that makes individual human beings distinctively valuable, such as sentience of the right kind or other distinctively human qualities" (2011: 227).[9] Thus, List and Pettit endorse "the view that something is good only if it is good for individual human, or more generally sentient, beings." (2011: 182).[10]

The idea here is not that it is impossible to characterise sentience in functionalist terms. Perhaps sentience is just a matter of a high level of functional complexity. 'Functional complexity' might mean 'information integration,' as under the 'integrated information theory' of consciousness—or it might mean some other kind of functional complexity. Regardless, the

[9] Similarly, those cited on (2) in fn. 1.

[10] List and Pettit see little need to argue for this. They state that "[i]ndividual persons create and organize group agents and it is in the co-reasoning of individuals, not that of corporate bodies, that the case should be made for or against equal rights." (2011: 181) If the second "and" means "so," this is a non-sequitur. After all, parents create and organize their children, yet the co-reasoning of parents alone is not where the case should be made for or against children's rights. (If the second "and" is not read "so," then the quote does not contain an argument.)

idea goes, *whatever it is* that undergirds sentience in humans and other animals does not inhere at the group level in group agents, where 'whatever it is' might be functional complexity or might be something else. I follow the orthodoxy here: functionalist accounts of sentience do not apply at the group level, for any realistic group made of human-like members.[11]

Of course, sentience is not the only game in moral patiency town. On a Kantian conception of moral rights, it is persons' distinctive capacity for autonomy that gives rise to moral rights. Combining a Kantian view with the functionalist-interpretivist picture painted above, groups do seem capable of autonomy and, therefore, seem to be bearers of Kantian moral rights. However, philosophers have tended to take this as a reason to reject Kantianism, rather than a reason to endorse groups' fundamental rights (Hess 2014; Wringe 2014). I won't explore the Kantian space here.

Another prominent account of moral rights sees moral rights as grounded in the right-bearer's emotional capacities (e.g. Jaworska 2007; Sepinwall 2017). If emotional capacities are understood as clusters of attitudes (judgements, motivations, etc) and behaviours, which are in turn understood along functionalist-interpretivist lines, then group agents do indeed have emotional capacities in an irreducible way.[12] However, the most prominent accounts of

[11] Dissenters exist. For example, Schwitzgebel (2015) argues that groups like 'the United States' can satisfy the integrated information theory of consciousness. List (2018) responds that such groups lack the requisite integration of information. Kramer (2021) sides with Schwitzgebel. Along different lines, Silver (2019: 263) argues that "we are in a particularly bad place epistemically to judge whether corporations are conscious given the kinds of creatures we are." I side with the orthodoxy here to pick battles with that orthodoxy elsewhere, discussed shortly.

[12] Groups' functionalist-interpretivist emotions are endorsed by Björnsson and Hess 2017, though they do not endorse the emotion-based view of moral patiency.

emotions do not understand emotions along purely functionalist-interpretivist lines. Emotions are normally understood as involving a certain phenomenology or 'what it is like'—that is, as involving sentience or consciousness (Scarantino and de Sousa 2018). If emotionality grounds moral patiency, then the standard view of emotions brings us back to the conclusion that groups lack fundamental rights because they lack phenomenology or sentience. [13]

In Section 3, I will assume that only sentient beings have fundamental rights. I will argue that this produces more support for groups' fundamental rights than is typically acknowledged. To get there, I must first explain the role of sentience in groups' blameworthiness.

2. Blameworthiness: An Alternative Account

As explained in the Introduction, moral agency is multi-faceted. For some facets, the functionalist-interpretivist view is probably adequate.[14] However, contra the standard view, blameworthiness is not one such facet. To see this, consider that present-day complex machines satisfy functionalist-interpretivist accounts of agency.[15] Yet they are not blameworthy. This

[13] One could explore other rights-bases as well, such as 'being an important cultural artefact' or 'being an ecosystem' (thanks to the Editors for noting this). I mention Kantianism and emotions because they are the most prominent alternatives.

[14] For example, elsewhere I have argued that groups' decision-making procedures (understood in functionalist terms) enable groups to *act* and *bear duties* (Collins 2019). List and Pettit (2011, ch. 7) argue their functionalist-interpretivism enables groups to *reason about morality*.

[15] From now on, I will discuss "machines" rather than "present-day machines"; by this, I do not mean to include hypothetical future machines that have significantly higher levels of sophistication than their present-day counterparts.

suggests functionalist-interpretivist agency is insufficient for blameworthiness—even if sufficient for other facets of agency, including other facets of moral agency.

List and Pettit are a useful foil, because they apply their functionalist-interpretivist account of agency to robots (2011: 19-24), while using that account to argue that group agents can be blameworthy (2011: ch. 7).[16] In their view, an entity can be blameworthy if it meets three conditions:

> Normative significance. The agent faces a normatively significant choice, involving the possibility of doing something good or bad, right or wrong.
>
> Judgmental capacity. The agent has the understanding and access to evidence required for making normative judgments about the options.
>
> Relevant control. The agent has the control required for choosing between the options. (2011: 155)

Not all functionalist-interpretivist agents satisfy these conditions. Presumably, List and Pettit would say many machines fail 'judgmental capacity,' despite being agents in the functionalist-interpretivist sense. Certainly, this seems true of the simple robot that they use to introduce their account of agency. That simple robot rolls around a tabletop, setting cylinders upright. Thus, it might seem the leading functionalist-interpretivist account of group agency—List and Pettit's—correctly gives the verdict that group agents can be blameworthy, while machines cannot, because of the latter's lack of judgmental capacity.

[16] List and Pettit use 'blameworthiness' or 'responsibility' as I do: they distinguish blameworthiness from causal responsibility and accountability, saying blaming "typically means indulging in resentment" (2011: 154).

Yet not all machines are so simple. Machines can be taught—and can teach themselves—methods for arriving at a wide variety of types of judgments. There's no reason to think that normative judgments are ruled out for machines, where normative judgments are simply judgments about what 'should' be done or what 'should' be the case. Machines have access to a wide range of evidence, as required by the judgmental capacity condition. Of course, we might debate whether they truly understand or cognize their evidence or the normative judgments they make as a result. But exactly the same is true of group agents—as illustrated by the 'China brain' thought experiment (Block 1978; Dennett 1991, ch. 14).[17] If List and Pettit (and other functionalist-interpretivists) want to hold that group agents have the requisite 'understanding' to meet the judgmental capacity condition, then it seems they had better allow that machines can also have the requisite understanding. After all, the resources that functionalist-interpretivists use to attribute that understanding to group agents (internal mechanisms, external interpretations) are available to machines as to groups.[18]

Yet complex machines—who meet List and Pettit's three conditions—are not candidates for blameworthiness, for a familiar reason: they lack sentience. As Robert Sparrow

[17] In this thought experiment, each citizen of China simulates the action of one neuron in a brain, using telecommunications to mimic the connections between neurons.

[18] Two leading accounts of moral understanding seem applicable to both machines and group agents. On Alison Hills' (2016: 102-3) account, moral understanding involves giving and following moral explanations and making moral inferences. On Paulina Sliwa's (2017) account, moral understanding involves a high degree of moral knowledge. Assuming we can give functionalist-interpretivist analyses of giving explanations, following explanations, making inferences, and knowledge, it seems 'understanding' should not be ruled out for machines or groups.

writes in an article arguing against the blameworthiness of 'killer robots' that commit war crimes:

> To hold that someone is morally responsible is to hold that they are the appropriate locus of blame or praise and consequently for punishment or reward. A crucial condition of the appropriateness of punishment or reward is the conceptual possibility of these treatments. Thus in order to be able to hold a machine morally responsible for its actions it must be possible for us to imagine punishing or rewarding it. … the most plausible accounts of the nature and justification of punishment require that those who are punished, or contemplate punishment, should suffer as a result. While we can imagine doing [damaging] things … to a machine, it is hard to imagine it *suffering* as a result. (2007: 71, emphasis added)

Here, Sparrow focuses on *punishment* as apt only for creatures who suffer. But similar things can—and should—be said about *blame* itself. According to recent accounts,[19] one crucial function of blame is to induce felt remorse in the target. If this account is correct, then an apt target of blame is something that has the capacity to feel remorse—which requires sentience, and more specifically a negative feeling or affect. Moreover, David Shoemaker (2019) has recently argued that entities can be forgiven only if they are capable of feeling sincere remorse—and that entities are 'blameable' only if they are forgivable. In what follows, I will follow these authors in assuming the capacity for the *feeling* of remorse is important for blameworthiness.[20]

Thus, there are at least three separate arguments against machines' blameworthiness: direct arguments that blameworthiness entails the capacity to suffer; arguments from the

[19] Macnamara 2011, esp. 90-91; Shoemaker 2015, 110, 171; Fricker 2016

[20] See also Collins 2023, ch. 5.

remorse-inducing function of blame (combined with the claim that machines do not feel remorse); and arguments from the sentience-based mechanisms that underly the transition from blame to forgiveness (where, again, these mechanisms are lack in machines, due to their lack of sentience). These three arguments together suggest that there is a further fourth condition—a condition that encompasses something like suffering or feeling remorse—that needs to be added to List and Pettit's three conditions.

I have recently proposed a condition on blameworthiness that plays this role: to be blameworthy, an entity must have the capacity for 'moral self-awareness,' characterised thus:

> An entity has moral self-awareness when it has a phenomenal belief-like attitude ('awareness') to the proposition 'I will do wrong,' 'I have done wrong,' or 'I am doing wrong.' There are three components to moral self-awareness: a wrongness component (the 'moral' part), a first-personal self-identifying component (the 'self' part), and a phenomenal belief-like component that is best captured by the notions of awareness, grasping, feeling, or acquaintance (the 'awareness' part). (Collins 2022: 6-7)

Moral self-awareness can be understood as the sentient, phenomenal, or felt component of self-directed reactive emotions such as remorse, guilt, or regret. If the capacity for moral self-awareness were added as a fourth condition on List and Pettit's list, this would get the correct result that machines are not candidates for blameworthiness. The condition furthermore dovetails with the three arguments against machines' blameworthiness.

Do group agents have the capacity for moral self-awareness? I have argued 'yes.'[21] Groups are made of humans. These humans can 'house' or 'hold' or 'realise' the group's moral

[21] Collins 2022. There are other accounts of collective self-awareness (e.g. Schmid 2014; Pettit 2018). Schmid focuses on *shared* (not *collective*) self-awareness; Pettit does not purport to

self-awareness. Nonetheless, the moral self-awareness is *the group's*—it should not be eliminated or attributed to the individual alone—because that moral self-awareness is constituted by irreducibly group components. More specifically:

> A collective is morally self-aware if, and because, a member is aware that 'this first-personal agent did (or is doing, or will do) wrong' (1) within the remit of their role, (2) because of their role, (3) while performing their role, and (4) while adopting the point-of-view of the collective. (2022, 15)

The use of 'this first-personal agent' refers to the collective itself, since the member is referring to themselves first-personally while adopting the collective's point-of-view. Importantly, the 'role' and 'point-of-view' are features that inhere at the group level, not the individual level. The 'role' exists only when embedded in the structure of the group agent, where this structure specifies how the role relates to other roles (for example, through relations of co-deliberation, reporting, or delegation). And the 'point-of-view' is that of the collective itself. Following the functionalist-interpretivist accounts outlined in Section 1, this point-of-view is a bundle of attitudes (beliefs, preferences, etc) that are attributable to the group itself (not to the members), whether because of the group's internal mechanisms or outward performances. These attitudes—and the point-of-view they compose—are attributable to the group (not the members) for two reasons. First, the attitudes are produced by internal mechanisms that do more than straightforwardly aggregate members' attitudes when arriving at the group's attitudes (this is the functionalist emphasis). Second, we can rationalise observed behaviour best when the

---

defend collective *phenomenology* (rather, he defends collective indexicality). On the difference between Schmid's shared self-awareness and my collective self-awareness, see Collins 2022: 11-13.

attitudes that compose the point-of-view are attributed to the group as a whole (this is the interpretivist emphasis).

Although groups' moral self-awareness is thus irreducibly group-level, it nonetheless depends on the intrinsic attributes of the group members. Group members' capacity to *feel* that the group has done wrong is an intrinsic feature of them as humans. It is only because members have the capacity to 'house' the group's moral self-awareness that the group has the capacity for moral self-awareness. A group's capacity for moral self-awareness, then, is an amalgam of irreducible group-level components (the role and the point-of-view, characterised via functionalist-interpretivism) and intrinsically member-level components (the member's capacity for awareness or phenomenology).

One might worry: if groups' capacity for moral self-awareness thus depends on members, should we attribute the capacity for moral self-awareness to *the group* irreducibly? And if this condition on blameworthiness is not attributed to the group irreducibly, then can the group be irreducibly blameworthy? The answer, in short, is the dependence does not entail reduction: the capacity for moral self-awareness, and the blameworthiness of which it is a condition, belongs to the group. Indeed, List and Pettit agree that groups' blameworthiness depends on attributes of group members, though irreducibly. Specifically, when explaining groups' abilities to meet the judgmental capacity condition, they write:

> Since the members of any group are able to form judgments on normative propositions in their individual lives, there is no principled reason why they should not be able to propose such propositions for group consideration and resolution – that is, for inclusion in the group's agenda. (2011: 159)

Their suggestion seems to be that groups' blameworthiness depends on members' moral agency, because groups cannot arrive at moral understanding, use moral evidence, or make

normative judgments without inputs from members. But this does not imply reductionism, any less than a cake can be reduced to a single ingredient. Likewise for the phenomenal component of groups' moral self-awareness: phenomenology is housed in members, but phenomenology is just one ingredient in groups' moral self-awareness, so we should not be reductionists about the latter. (The other ingredients—i.e., the role structure and rational point-of-view—are features of the group itself.)

The account of groups' irreducible blameworthiness sketched in this section differs from Section 1's functionalist-interpretivism in two crucial ways. First, my account asserts that the capacity for sentience (specifically, moral self-awareness) is a condition on blameworthiness. Second, it asserts that groups have this sentience, when members have the right sentient states while performing their role and adopting the group's point-of-view. Yet the account is similar to functionalism and interpretivism, insofar as the 'role' and 'point-of-view' are characterizable along functionalist-interpretivist lines. Again: we need an *amalgam* of irreducibly group elements (the role and the rational point-of-view), together with intrinsically human-housed elements (the sentience), in order to produce an overall mental state (the group's moral self-awareness) that is a fourth condition on blameworthiness. In the next section, I consider the implications for groups' irreducible fundamental moral rights.

3. From Blameworthiness to Rights

I now suggest that groups can experience all sorts of mental states that have both a phenomenal component and a functional-interpretivist component, along similar lines to moral self-awareness. In turn, these mental states can be the basis of sentience-based fundamental rights for groups.

Consider two of the states Hess mentions in her denial of groups' fundamental rights: humiliation and hatred. Following the formula for moral self-awareness, group agents can experience these states, including their phenomenal components, just as they experience moral self-awareness. The formula for group humiliation would run:

> A collective is humiliated if, and because, a member is aware that 'this first-personal agent did (or is doing, or will do) something humiliating' (1) within the remit of their role, (2) because of their role, (3) while performing their role, and (4) while adopting the point-of-view of the collective.

Likewise, the formula for group hatred would run:

> A collective experiences hatred if, and because, a member is aware that 'this first-personal agent hates something' (1) within the remit of their role, (2) because of their role, (3) while performing their role, and (4) while adopting the point-of-view of the collective.

Crucially, the 'awareness' in both cases involves sentience or phenomenology: the member's 'awareness' implies that the member *feels* humiliation or hatred. Such feelings are the attribute that many take to rule out groups' fundamental rights.

More generally, members' sentient or phenomenal mental states are attributable to the group whenever those states are experienced (i) within, because of, and while performing their role and (ii) while the member adopts the group's point-of-view. Some such mental states will have a phenomenal component. Mental states that can satisfy (i) and (ii), and that have a phenomenal component, include moral self-awareness, alongside humiliation, hatred, and numerous other states.

Importantly, this does not commit us to endless attributions of group-level sentience or phenomenology. For example, on this view, members' felt hunger is usually not attributable to

the group. This is because hunger is not the kind of phenomenal state that can be had 'within' one's role, that is, as required by one's role. Most group roles are completely silent on how often members should eat, so hunger—even when experienced 'while on the job'—is not a state that's experienced *within* one's role.[22] Likewise, the particular phenomenology of seeing red, tasting lemons, or feeling cold are not the kinds of phenomenal states that are usually incorporated into group roles.

Yet some phenomenal states are required by group roles, where 'required' crucially includes 'required by the informal norms, expectations, and culture of the group agent.' Paradigmatically, these phenomenal states include the phenomenal components of *emotions*—such as humiliation, hatred, jealously, and joy. These are complex states that have both a phenomenal and a non-phenomenal component. The non-phenomenal component includes beliefs, preferences, and other attitudes that can be characterised along functionalist-interpretivist lines. For an emotion to attributed to the group, the non-phenomenal components must derive from the group's rational point-of-view. When the non-phenomenal component

[22] An exception might be a protest group engaging in a hunger strike. Here, the hunger might be group-level, but only if the hunger somehow 'includes' (as a component part) a propositional attitude that is acquired from the group's bundle of mental states. Plausibly, hunger does not include *any* propositional attitude—it is 'merely' a feeling—so it cannot be attributed to a group, even a hunger-striking one. In short: condition (4) (requiring that the member adopts the group's point-of-view) must do some substantive work in determining the content of the mental state; the group's point-of-view must produce an attitude that is part of the mental state. I refer to this attitude below as the 'non-phenomenal component' of the group-attributed mental state.

does thus derive—and when the phenomenal component is experienced within, because of, and while performing one's role—then the whole emotion gets attributed to the group.

For example, if my university drops in international rankings, it might hold the non-phenomenal propositional attitudes that are components of humiliation. I might adopt these group-derived non-phenomenal propositional attitudes. At the same time, I might also *feel* humiliation as a university staff member—that is, within, because of, and while performing my university role. This might occur if such humiliation is required of the informal norms at my university. Perhaps, as a private person, I couldn't care less about university rankings. But I can nonetheless feel humiliated within, because of, and while performing my role, and while I adopt my university's point-of-view (that is, while taking up its non-phenomenal attitudes that are components of humilitation). In this case, my university feels humiliation, whose phenomenal component is 'housed' within me. My felt humiliation is the university's felt humiliation in those cases where the rational pressure to care about rankings comes from the university's rational point-of-view—including from its informal culture, ethos, or norms. Organizations' points-of-view often embed themselves in a workplace's culture, ethos, or norms, so this is not an outlier case.

Of course, such group-level phenomenal states depend on members: the states are realised by, or housed in, members. But the complex mental state—the state that has *both* phenomenal *and* non-phenomenal components—should be attributed to the group in a non-reductionist way, because the non-phenomenal components belong irreducibly to the group. The complex state of 'humiliation' cannot be disaggregated into its component parts: it is one amalgam state, which is a fusion of its phenomenal and non-phenomenal aspects. The amalgam should be attributed at the group level, with the member who houses the phenomenology being understood as a material part of the group.

Do these group-level emotions give rise to irreducibly group-level fundamental rights? For example, does my university have a right not to be publicly shamed, tormented, or scorned for dropping in the rankings, where this right is not justified purely by reference to the rights of non-group entities? One might think not: it might seem that my university has a right against shaming only if shaming would negatively affect humans like me.

But notice that the humiliation isn't really bad for humans like me. It's bad for humans like me only insofar as we operate *as* the university, not in our private human capacity. For it is only when I operate as the university that I feel the humiliation at all. That is: the 'me' for whom the humiliation is bad *is* a part of the university.[23] So any claims that the humiliation gives rise to—claims against shaming, say—are, strictly speaking, claims of my university. It is me-qua-university-part—not me-qua-private-agent—who has the negative experience that partly constitutes humiliation.

This would remain true even if I did privately care about university rankings. If I privately cared about university rankings, then there would be two types of humiliation that I could house: my own personal humiliation (experienced while conditions (i) and (ii) are not met) and the university's humiliation (experienced while conditions (i) and (ii) are met). My private humiliation might give rise to *my* claim against being shamed. But the university's humiliation (whose phenomenology is housed within me) would also give rise to a claim. This claim is held by me-qua-university-part—or, we can say more straightforwardly, simply by the university itself. This is just as other claims I have qua-university are really claims of the university: when I have a right to set deadlines for students, this is my exercise of the university's right to set deadlines for students. I am the realiser of the university's right.

[23] I defend the metaphysics behind this picture in Collins (2023, chs 2-3).

This picture of the relation between me and my university can be taken further, via the metaphysics of Carol Rovane (1998; 2014). In Rovane's view, some parts of human lives are, strictly speaking, not parts of the human-agent but are rather parts of the group-agent to which those humans belong. On this view, the agent we'd normally identify as 'me' is smaller than human-sized: that agent is the size of a human, *minus* that part of the human that is a part of the university. This university-devoted 'part' of my human life can be understood as a collection of time-slices rather than a collection of body-slices. That is: on this metaphysics, the time-slices of my human body that act for the university do not, strictly speaking, act for me. So, I am smaller than we typically think, in terms of spatio-temporal extension. Some time-slices of my body that we'd normally think of as 'me' are, in actuality, time-slices of my university. When those time-slices are humiliated, the humiliation is my university's, not mine. So, any rights against that humiliation must belong to my university. Of course, one need not accept Rovane's metaphysics in order to agree with the argument I have made thus far. But Rovane provides one way of further conceptualizing this landscape.

At this point, one might contest the possibility of compartmentalizing a human life this much. Further, one might contest the desirability of doing so: doesn't it display a morally problematic lack of integrity to feel humiliated 'merely' as part of the university and not in one's personal capacity?[24] In response, I do not deny that there are—nor that there should be—*causal links* between a human's feelings-qua-group and feelings-qua-private-agent. One might feel humiliated-qua-private-agent as a causal consequence of feeling humiliated-qua-group. Perhaps this is even psychologically unavoidable or morally desirable; I take no stand on that. My claim is rather than there is no *constitutive link* between these two modes of humiliation.

[24] I thank Olof Leffler for this. The relevant kind of integrity comes from Williams (1973: 97-99).

One mode adopts the group's rational point-of-view; the other mode adopts the private agent's. These are conceptually and constitutively distinct, even if causally related.

Another objection is more ontological in nature: consider an analogy to verbal utterances. When a verbal utterance is attributed to a group, the utterance should *also* be attributed to the human member who makes the utterance.[25] The group makes the utterance *in virtue of* the member making the utterance within, because of, and while performing her role, and while adopting the university's rational point-of-view. A group's verbal utterances are realized by bodily states of humans, such as the human's lips and tongue moving in a certain way. The bodily states of an utterance are fully housed within a human's body, even though those utterances are attributable to the group agent. There is just *one utterance*—one instance of lips and tongue moving in a certain way—and that one utterance is attributable to two agents: the human and the group.

The objection contends that groups' phenomenal states are like groups' verbal utterances. They require a physical housing within a member. But this doesn't mean there are *two* phenomenal states in the world (one that is the member's and one that is the group's), any less than there are *two* utterances in the world when a member speaks for a group. Instead, there is just one phenomenal state—housed in the member—which happens to be *attributable* to both the member and to the group. It is attributable to the member because she entirely houses the phenomenology. It is attributable to the group because the group's role structure and point-of-view play a constitutive role in the state.

[25] Assuming the utterance is made by a human member—I leave aside the complication of machines who can make verbal utterances for groups.

One might use this analogy to insist that *the houser* of the phenomenology is the only one that has fundamental moral rights, because the houser is what grounds the attribution of the phenomenology to the group agent. It would be an egregious kind of double-counting if we posited two subjects of experience; likewise, one might think, it is egregious double-counting to posit two bearers of fundamental rights. This, one might think, would be like saying that the group has its own mouth and tongue, which move when the member makes an utterance on the group's behalf. But there is only one mouth and tongue—and these belong, fundamentally, to the member. Likewise for phenomenology.

In response: I agree that there is just one phenomenal state when my university has a pang of humiliation that is housed within me. I also agree that there is just one 'subject of experience,' as philosophers of mind might use that term to describe the site of phenomenology. I agree that groups' phenomenal states are like groups' verbal utterances in that they require a physical housing within a member. And I agree that groups do not have phenomenal states that are 'purely' their own, any more than they have their own mouths and tongues. However, if a complex mental state such as an *emotion*—which includes not just phenomenal components, but phenomenal *and* non-phenomenal components *in amalgamation*—is the basis of a right, then the group bears the emotion, so the group has the right. The member themselves does not have the emotion (because the emotion's non-phenomenal components reside entirely at the group level), so the member does not have the right, where we take 'member' to refer to a human considered independently of group membership.

Does this establish only the justificatory fundamentality, or merely their ontological irreducibility? One might grant that 'my' humiliation is sometimes really my university's humiliation, in a way that is ontologically irreducible to my humiliation, because that humiliation is experienced while conditions (i) and (ii) are met. So, if there is a right against being shamed into humiliation, then that right must belong to my university, since the

humiliation is my university's. But, one might continue, this does not mean that the *justification* for that right concerns the university at all. Instead, one might insist that my university's humiliation has moral import only because of its effects on me (or others like me).

Again, though: who is this 'me'? On the one hand, if we mean 'me' understood in my private capacity, then that person is either not affected by my university's humiliation (if I couldn't care less about university rankings) or is affected in a way that doesn't exhaust the sentience-related upshots of the humiliation (because my private humiliation is not the entirety of the sentience in the picture). On the other hand, if we mean 'me' understood *as a part of the university*, then 'my' humiliation—the complex emotional state that has functionalist-interpretivist components that are irreducibly group-level—is the university's humiliation. This humiliation is at least part of the justification for my university's right. So, the justification of my university's right does not purely concern the claims or interests of humans.

Crucially, this view does not give groups a carte blanche to claim all the moral, social, and political rights held by humans. My claim is a limited one: *some few* sentient states are properly attributed to groups, so any rights grounded in *those specific sentient states* are both ontologically and justificatorily group-level. For each right that we attribute to groups, we should be justificatory holists about that right only if the correct justification for that right ultimately refers to some group-level sentient state, following the formulae given for humiliation and hatred earlier in this section.

How would this work for groups' rights to speech, conscience, religion, voting, assembly, and so on? One would have to show that groups' holdings of those particular rights are justified on the basis of group-level emotions, using a formula like that provided above for hatred and humiliation. Perhaps a corporation's right to free speech is justified by its

humiliation when silenced. Perhaps a university's right to vote is justified by its taking offence at being less than an equal citizen.

However, it is implausible that political rights can be justified solely via the value of such emotions. Consider the justifications for such rights as held by humans. These justifications might *partly* refer to humans' emotions (such as humiliation and offence). But more commonly, the justifications refer to human dignity, or the equal moral worth of all humans. My argument for groups' emotions (and, therefore, their sentience; and, therefore, their rights) does not generate groups' dignity, or groups' equal moral worth to humans. Thus, while groups' sentience plausibly gets them *some* rights—for example, rights that others do not gratuitously create negative group emotions—I doubt that the political rights listed above are included. Humans' dignity and equal moral worth are beyond the scope of this paper, but I take it that our uniquely human ways of relating to each other, and the relatively open-ended and domain-general nature of our agency (when compared to groups), will play a role in justifying our political rights. I have said nothing about how groups fare on these dimensions. The mere existence of group emotions is insufficient to establish their political rights.

Further issues would arise if one wanted group rights formalised in law. Irreducible fundamental group rights will often conflict with individuals' rights. Perhaps individuals' rights are always more important than groups' rights: perhaps individuals' sentient states are more important or more numerous than the groups' ones, for example because there are distinctively human ways of relating, or because humans' agency is more 'all-purpose' than groups' agency. If so, individuals' rights will often play trumps in the game of codification. Again, this seems plausible to me. Groups may be capable of emotions and sentience, but that does not make them the same kind of creature as us.

Conclusion

In the wake of arguments such as List and Pettit's, an orthodoxy has emerged: groups can be irreducible agents but they cannot be irreducible patients. Much is packed into 'agent' and 'patient,' so I have focused just on two facets: blameworthiness and fundamental rights. I agree with the orthodoxy that groups can be irreducibly blameworthy, but I disagree on the reasons why. Functionalist-interpretivist agency is not enough for blameworthiness; if it were, machines would be blameworthy. Irreducibly blameworthy group agents must have the capacity for moral self-awareness, where the phenomenology of moral self-awareness is housed in human members. However, this opens the door to groups' fundamental rights, since a range of emotions—not just moral self-awareness—are properly attributable to group agents, where those emotions' phenomenology is housed in human members. If those emotions ground rights, then those rights are irreducibly group-level. I have suggested that we should walk through this open door, but that the space beyond is not as large as we might fear.

References

Björnsson, Gunnar and Kendy Hess. 2017. "Corporate Crocodile Tears? On the Reactive Attitudes of Corporate Agents." *Philosophy and Phenomenological Research* 94(2), 273-298.

Block, Ned. 1978. "Troubles with Functionalism." *Minnesota Studies in the Philosophy of Science* 9, 261-325.

Collins, Stephanie. 2019. *Group Duties: Their Existence and Their Implications for Individuals*. Oxford University Press.

Collins, Stephanie. 2022. "I, Volkswagen." *Philosophical Quarterly* 72(2): 283-304.

Collins, Stephanie. 2023. *Organizations as Wrongdoers: From Ontology to Morality*. Oxford University Press.

Collins, Stephanie. Forthcoming. "Abilities and Obligations: Lessons from Non-agentive Groups." *Erkenntnis*.

Copp, David. 2006. "On the Agency of Certain Collective Entities: An Argument from 'Normative Autonomy.'" *Midwest Studies in Philosophy* 30: 194–221.

Dennett, Daniel. 1991. *Consciousness Explained*. Back Bay Books.

Epstein, Brian. 2015. *The Ant Trap: Rebuilding the Foundations of the Social Sciences*. Oxford University Press.

French, Peter A. 1984. *Collective and Corporate Responsibility*. New York: Columbia University Press.

Fricker, Miranda. 2016. "What's the Point of Blame? A Paradigm Based Explanation." *Nous* 50(1), 165-183.

Hedahl, Marcus. 2017. "Collective Directionality: A New Possibility for Collectives as Objects of Normative Consideration." *Journal of Value Inquiry* 51: 233-250.

Hess, Kendy. 2013. "If You Tickle Us…: How Corporations Can Be Moral Agents Without Being Moral Persons." *Journal of Value Inquiry* 47(3), 319-335.

Hess, Kendy M. 2014. 'The Free Will of Corporations (and Other Collectives).' *Philosophical Studies* 168, 241-260.

Hess, Kendy M. 2018. "Does The Machine Need a Ghost? Corporate Agents as Nonconscious Kantian Moral Agents." *Journal of the American Philosophical Association* 4(1), 67-86.

Hills, Alison. 2016. "Understanding Why." *Nous* 50: 661-688.

Hindriks, Frank. 2018. "Collective Agency: Moral and Amoral." *Dialectica* 72(1), 3-23.

Huebner, Bryce. 2014. *Macrocognition: A Theory of Distributed Minds and Collective Intentionality*. New York: Oxford University Press.

Jaworska, Agnieszka. 2007. "Caring and Full Moral Standing." *Ethics* 117: 460–497.

Kramer, Max F. 2021. "What it Might Be like to Be a Group Agent." *Neuroethics* 14: 437-447.

Kymlicka, Wil. 1995. *Multicultural Citizenship: A Liberal Theory of Minority Rights*. Oxford: Clarendon Press.

List, Christian. 2018. "What Is It Like To Be A Group Agent?" *Nous* 52(2), 295-319.

List, Christian and Philip Pettit. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.

Lovett, Adam and Stefan Riedener. 2021. "Group Agents and Moral Status: What Can We Owe to Organizations?" *Canadian Journal of Philosophy* 51(3): 221-238.

Macnamara, Coleen. 2011. "Holding Others Responsible." *Philosophical Studies* 152(1), 81-102.

Manning, Rita C. 1984. "Corporate Responsibility and Corporate Personhood." *Journal of Business Ethics* 3: 77–84.

McMahan, Jeff. 2002. *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press.

Ozar, David T. 1985. "Do Corporations Have Moral Rights?" *Journal of Business Ethics* 4(4): 277–281.

Pasternak, Avia. 2017. "From Corporate Moral Agency to Corporate Moral Rights." *Law and Ethics of Human Rights* 11(1), 135-159.

Pettit, Philip. 2018. "Consciousness Incorporated." *Journal of Social Philosophy* 49(1), 12-37.

Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Oxford University Press.

Rovane, Carol. 1998. *Bounds of Agency: An Essay in Revisionary Metaphysics*. Princeton: Princeton University Press.

Rovane, Carol. 2014. "Group Agency and Individualism." *Erkenntnis* 79: 1663-1684.

Scarantino, Andrea and de Sousa, Ronald. 2018. "Emotion." *Stanford Encyclopedia of Philosophy*.

Schmid, Hans Bernhard. 2014. "Plural Self-Awareness." *Phenomenology and the Cognitive Sciences* 13(1), 7-24.

Schwitzgebel, Eric. 2015. "If Materialism is True, the United States is Probably Conscious." *Philosophical Studies* 172, 1697-1721.

Sepinwall, Amy. 2017. "Blame, Emotion, and the Corporation." Pp. 143-166 in E.W. Orts and N.C. Smith (eds), *The Moral Responsibility of Firms*. Oxford University Press.

Shoemaker, David. 2015. *Responsibility from the Margins*. Oxford: Oxford University Press.

Shoemaker, David. 2019. "Blameworthy but Unblamable: A Paradox of Corporate Responsibility." *Georgetown Journal of Law and Public Policy* 17: 897-918.

Silver, Kenneth. 2019. "Can A Corporation Be Worthy Of Moral Consideration?" *Journal of Business Ethics* 159, 253-265.

Singer, Peter. 1993. *Practical Ethics*. Second edition. Cambriddge University Press.

Sliwa, Paulina. 2017. "Moral Understanding as Knowing Right from Wrong." *Ethics* 127: 521-552.

Strohmaier, David. 2020. "Two Theories of Group Agency." *Philosophical Studies* 177(7), 1901-1918.

Tollefsen, Deborah. 2015. *Groups as Agents*. Cambridge: Polity Press.

Williams, Bernard. 1973. "A Critique of Utilitarianism." In JJC Smart and Bernard Williams, *Utilitarianism: For and Against*. Cambridge, Cambridge University Press, 77-150.

Wringe, Bill. 2014. “May I Treat a Collective as a Mere Means?” *American Philosophical Quarterly* 51(3), 273-284.