

# Collective moral agency and self-induced moral incapacity

Niels de Haan 

Department of Philosophy, University of Vienna, Vienna, Austria

## ABSTRACT

Collective moral agents can cause their own moral incapacity. If an agent is morally incapacitated, then the agent is exempted from responsibility. Due to self-induced moral incapacity, corporate responsibility gaps resurface. To solve this problem, I first set out and defend a minimalist account of moral competence for group agents. After setting out how a collective agent can cause its own moral incapacity, I argue that self-induced temporary exempting conditions do not free an agent from diachronic responsibility once the agent regains its moral faculties. For collective agents, any exempting condition is potentially temporary due to the 'malleability' of their constitution. Therefore, in cases of self-induced moral incapacity and subsequent wrongdoing, unlike individuals, every collective agent can be (made) morally responsible for its actions even though it did not qualify as a moral agent at the time of wrongdoing. Hence, this is no reason for skepticism concerning corporate responsibility.

## ARTICLE HISTORY

Received 15 July 2020

Accepted 9 May 2022

## KEYWORDS

Collective moral agency;  
corporate responsibility;  
diachronic responsibility;  
exemptions; moral  
competence; responsibility  
gaps

## 1. Introduction

Numerous philosophers argue that certain structured groups, such as states, corporations, or universities, qualify as collective agents (Collins 2019b; Copp 2006; Donaldson 1982; Erskine 2003; French 1984b; Hess 2014; Hindriks 2018; Lawford-Smith 2015; List and Pettit 2011; Pauer-Studer 2014). Collective agency is to be explicated in terms of a collective decision-making procedure that enables the group to identify and pursue representational and motivational attitudes while satisfying desiderata of rationality in a robust manner. The procedure is part of the organizational structure, which further includes rules, policies and conventions in virtue of which the group coordinates their decision-making and action-taking. This view, which I call *procedural collectivism*, is widely discussed throughout the literature on group agency (especially List and Pettit 2011), so I will not go into detail here. Instead, I want to focus on collective moral agency (CMA). Procedural collectivists claim that in a functional sense akin to individuals, most collective agents can understand and process moral reasons and act accordingly. These collective agents not only stand under moral requirements, but we can hold them morally responsible for their actions. For example, we hold British Petroleum morally responsible for the Deepwater

**CONTACT** Niels de Haan  niels.de.haan@univie.ac.at  Department of Philosophy, University of Vienna, Vienna, Austria

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Horizon oil spill in the Gulf of Mexico in 2010. This is called corporate responsibility, which is non-distributive. This does not mean that the members cannot be blamed for their actions, but corporate responsibility is an additional and non-redundant level of responsibility.

For example, Christian List and Philip Pettit claim that a group agent is fit to be held responsible for doing something if it satisfies the following three conditions: (a) the group agent faces a normatively significant choice, involving the possibility of doing something good or bad, right or wrong; (b) the group agent has the understanding and access to evidence required for making normative judgments about the options; and (c) the group agent has the control required for choosing between the options (List and Pettit 2011, 158).<sup>1</sup>

List and Pettit argue that when a collective agent is faced with a normatively significant choice, the members can put forward normative considerations to the decision-making process as long as they have access to evidence on relevant matters (2011, 158). According to List and Pettit, '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). Once a proposition is presented by a member for consideration, the group will take whatever steps are prescribed in its organizational structure for endorsing it (e.g. via a vote of the committee-of-the-whole, an authorized subgroup or an appointed official).

However, some collective agents may lack or come to lack the necessary abilities for moral agency due to the way they are organized. Collective agents are more 'malleable' in their constitution than individuals. Because the capacities of a collective agent are dependent on its members and how it is organized, certain faculties may be suddenly lost when there are changes in its structure, rules or procedures. This may occasionally involve costs, but it need not involve a loss of essential features for pursuing prudential interests. This means that it could be in the prudential interest of collective agents or other parties to 'get rid of' the collective agent's moral competence by changing their decision-making procedures or organizational structure. If the collective agent was designed as an amoral agent from the start by its designers or made this way by another (collective) moral agent over time, then we can hold these agents morally responsible for the actions of the collective agent. But there is a third possibility. The collective agent can become amoral by its own hand. I call this *self-induced moral incapacity*.<sup>2</sup> The collective agent has caused itself to lose (some of) its essential moral faculties and becomes an amoral agent. This raises a distinctive problem. If a collective agent no longer qualifies as a moral agent, then, arguably, we cannot hold the agent morally responsible for what it did.<sup>3</sup> And here we cannot identify any other agent as responsible for the collective agent's possible wrongdoing. This creates a tension with the very purpose for which the notion of CMA is invoked.

Corporate responsibility gaps concern cases where a collective agent has done something seemingly morally wrong, but no (individual) agent can be held responsible for the actions of the corporation (see among others Braham and van Hees 2011; Collins 2019a; Copp 2006 & 2007; French 1984b; Pettit 2007). Pettit thinks that in cases of corporate wrongdoing, if only responsibility can be ascribed to individuals as enactors or constitutors, then there will be 'gaps in the books' that we can keep on individuals (2009, 170). For example, it is possible that each agent has a legitimate excuse for the wrongful corporate

action, whereas these excuses do not hold for the group agent (Copp 2006). Pettit calls this elsewhere a ‘deficit in the accounting books’ (2007, 194). To fill these deficits, we need corporate responsibility. To the best of my knowledge, this is the main argument of procedural collectivists in favor of corporate responsibility (see also French 1984b, 141).

The problem a self-induced loss of moral competence appears to create is that once we have filled corporate responsibility gaps with corporate responsibility, responsibility gaps may arise in the very same spots, because collective agents can incapacitate themselves as moral agents. Corporate responsibility is invoked to avoid a shortfall in what we should expect the practice of holding agents morally responsible to deliver (Pettit 2009, 170), but self-induced moral incapacity makes it possible that such shortfalls resurface. I call this the *problem of self-induced moral incapacity*. This may fuel a certain skepticism about CMA and corporate responsibility. If collective agents can simply ‘opt out’ of moral agency, then what’s the point? There is something wrong with a moral theory that lets the agents it places under moral demands ‘get off the hook’ for wrongdoing by ‘shutting off’ their moral agency. The skeptic might take this to show that CMA and corporate responsibility is incoherent.

List and Pettit do briefly discuss the possibility of amoral group agents. They note that the procedures of the group agent may restrict its agenda to propositions of a purely descriptive kind. They think this need not be disturbing for two reasons. First, few group agents are likely to impose procedural restrictions against forming moral judgments about the options they face. And second, it would be a serious design fault, from the perspective of society as a whole, to allow any group agents to avoid making such judgments. Society should regulate group agents to ensure condition (b) is met. Groups seeking to be incorporated would thus be legally required to have procedures in place whereby they give due consideration to evaluative matters and form collectively endorsed judgments on them. (List and Pettit 2011, 159)

While I agree with List and Pettit’s suggestion, this answer is not sufficient to address the worries of the skeptic. First, how likely or unlikely it is in practice does not do much to remove the worry that a group agent *can* restrict members from making moral contributions, especially given that this may be in the prudential interest of the group agent. Second, legal requirements indeed should be imposed on groups incorporating as a group agent. The possibility that group agents may become amoral not only by design but over time suggests that there may also be a need for imposing legal requirements on existing group agents to enforce the continuity of their moral agency. But considering that morality precedes legality and that there will likely be societies where such regulations are not yet or not anymore in place, such contingent facts do not address the theoretical worry that CMAs can become amoral agents by their own hand. Thus, we need an answer to the problem of self-induced moral incapacity.

In the remainder, I aim to provide an answer to this problem and show that skepticism about CMA is ultimately unwarranted. First, I set out a minimalist account of moral competence for group agents in Section 2. In Section 3, I consider a few objections against my view. With a clear picture of how moral competence functions within a collective agent, I explain how a collective agent can cause its own moral incapacity in various ways in Section 4. Finally, I address the problem of self-induced moral incapacity and resurfacing responsibility gaps in Section 5. I distinguish between temporary and permanent exempting conditions and between synchronic and diachronic responsibility. Self-induced

temporary exempting conditions do not free an agent from diachronic responsibility once the agent regains its moral competence. Similar to individual cases, we can trace corporate responsibility for wrongdoing during the state of self-induced moral incapacity back to an earlier culpable decision. To answer the skeptic, I argue that we can always rely on such tracing arguments, because, unlike persons, a collective agent can always regain its moral competence *ex post facto*. Hence, with respect to self-induced moral incapacity, we have less reason to be skeptical about the responsibility of collective agents than of individuals.

## 2. Moral competence: a minimalist account

In this section, I will set out a minimalist account of moral competence for group agents. This account explicates what I call the general sense of collective moral agency. List and Pettit's conditions for the fitness of a collective agent to be held responsible concern a particular choice or instance of its behavior. This relates to a particular understanding of what moral agency means. We can understand moral agency as an instance of agency that is subject to moral appraisal, which is the sense List and Pettit have in mind. But we can also understand moral agency in a more substantive and robust sense as being an agent that, generally speaking, stands under moral obligations, that is capable of acting rightly and wrongly, and that can be morally responsible for its actions and attitudes. The conditions List and Pettit invoke are related to specific abilities, because these abilities are related to the particular instance of behavior that is under evaluation. For example, the agent may or may not have had the ability to gather relevant evidence in this particular circumstance. If an agent has a specific ability to perform an action, this (arguably) entails she has the general ability to do so.<sup>4</sup> But the absence of a specific ability does not necessarily imply the absence of a general ability. The distinction differentiates between 'what an agent is able to do in a large range of circumstances, and what the agent is able to do now, in some particular circumstances' (Whittle 2010). I might not be able to serve in tennis right now, because I am miles away from the tennis court, but that does not mean that I do not have the general ability to serve (Maier 2018). A general sense of collective moral agency relies on general abilities, and it is this sense of moral agency that I am after.

The reason for this is that, as Pettit (2017, 29) also acknowledges in later work, if certain general abilities are absent, then the agent will be exempt from responsibility. Let a specific ability be absent and you will be excused by the factor that impedes it, assuming that you are not responsible on an independent basis for letting that factor get in the way.<sup>5</sup> I take it that Pettit agrees with the following reading: Excuses operate locally, they give us reason to withdraw our reactive attitudes we would ordinarily take in response to a particular action, but excuses do not provide any reason to view the agent in a different light altogether. Exemptions, however, invite us to suspend our reactive attitudes towards an agent altogether, at least for a certain period of time, and to take what Peter F. Strawson (1962) famously called the objective view. We see agents not as ones to be esteemed or resented, but as ones to be controlled, managed, manipulated or trained. Exemptions block responsibility for a particular act by showing that a normally impermissible act has been done by someone who is not, in general, a morally responsible agent (Wallace 1994, 156).<sup>6</sup> The problem of self-induced moral incapacity and resurfacing

corporate responsibility gaps has much to do with how exemptions work. If the agent lacks one of the general abilities necessary for moral competence, then it appears the agent must be exempted from responsibility. Hence, it is important to get clear what these general abilities exactly are.

Before I turn to moral competence, it is important to say more about abilities. I will not adopt any specific account of ability, but I'll rely on two plausible assumptions.

First, I'll assume that ability is a success notion (see e.g. Greco 2009; Jaster 2020; Sosa 2015). Abilities are a matter of success across a sufficient proportion of modal space (Jaster 2020). For example, I have the ability to hit the bullseye on a dart board only if I successfully hit the bullseye in a sufficient proportion of cases where I intend to hit the bullseye.<sup>7</sup> There needs to be an appropriate modal tie between the intention and the successful performance of the action.<sup>8</sup> For me to have this ability, it is not sufficient if I hit the bullseye once as a fluke, but I need not hit the bullseye every single time I intend to hit it. Modal reliability is part of the success-factor of a general ability: the agent has to be able to  $\phi$  across a whole variety of situations, meaning the agent must be successful in a sufficient proportion of the wide range of cases where the agent intends to perform the action.

Second, I'll assume that group abilities supervene on the individual abilities of members: A group-level ability supervenes on a set of individual abilities if and only if for each change in the group-level ability there is some change in the members' individual abilities.<sup>9</sup> For example, a corporation's ability to undertake a hostile takeover supervenes on a complex set of individual abilities of its members. This group ability (whether specific or general) can be *multi-realizable* (see also Collins 2019b, 78). The supervenience base can have different configurations of member-level abilities, each of which is sufficient to realize the relevant group-level ability. However, not all abilities of members should be countenanced towards what the group agent is able to do. The possible supervenience base for group abilities is restricted to the abilities of members that broadly speaking fall within the purview of their membership of the group agent.

With these assumptions in mind, let's turn to moral competence. Simply put, moral competence is the capacity to be appropriately responsive to moral reasons. I take moral competence to be sufficient for qualifying as a moral agent. To be morally competent does not mean that one is infallible. Moral agents can and regularly will make mistakes. Instead, being (sufficiently) morally competent means that the agent is capable of understanding and responding to moral reasons past a certain threshold. Moral competence is the combination of moral understanding and control (Wallace 1994). An agent has moral understanding if and only if (and to the degree to which) it has the ability to acquire moral knowledge (Sliwa 2017; cf. Hills 2016).<sup>10</sup> Moral knowledge includes knowing that an action is right or wrong, knowing *why* an action is right, just, or fair, and so on. There are different ways in which an agent can acquire moral knowledge (depending on the kind of agent), this need not only be via moral reasoning, but it can be via imagination, perception, intuition, affective responses or even testimony (Sliwa 2017, 548).<sup>11</sup> The agent needs to have a mechanism that, if working correctly, yields moral knowledge when the agent is presented with relevant moral evidence. If the agent also has the ability to control its goal-seeking states and actions in light of its moral understanding, then the agent is morally competent.

I can now formulate the minimalist account of moral competence for group agents:

**Moral Competence:** Group agent GA has the general ability of moral competence if and only if

- (1) GA has moral understanding, that is, (a) the general ability to grasp moral reasons that possibly govern GA's actions; and (b) the general ability to relate such reasons to GA's available evidence; and
- (2) GA has the general ability to control its goal-seeking states and actions accordingly in light of (1).

Let me expand briefly on each group-level ability. Let's say that to be able to grasp a moral reason, one must be able to form an appropriate truth-tracking belief concerning the reason that constitutes knowledge.<sup>12</sup> The group's performance as an epistemic agent depends on its aggregation procedure (List 2005), and this can have an important moral dimension. For a group agent to be able to grasp a moral reason, then, it must be able to collectively endorse an appropriate truth-tracking belief concerning the reason that constitutes group knowledge. The ability of the group agent to do so supervenes on members' individual abilities to grasp moral reasons, to provide moral contributions to the (central or sub-group) decision-making process and to vote on relevant matters. One might think that members' grasping of moral reasons happens at the same level of decision-making, but this need not necessarily be the case. Low-ranked employees outside of the decision-making may formally or informally have the opportunity to provide moral contributions to the agenda (cf. Collins 2019b, 160).<sup>13</sup> For example, a low-level employee might recognize that the collective agent is potentially doing something morally wrong, and as long as he is able to put this moral consideration forward somehow, to pass it up the chain of command, the moral reason can be taken up in the decision-making.

But note that the ability to grasp moral reasons in general is not sufficient for a group agent to have moral understanding. To have the ability to acquire relevant moral knowledge, the group agent must have an organizational structure which, if it works correctly, yields moral knowledge when it is presented with *relevant evidence*. This means that grasping the moral reasons that actually (and not just possibly) pertain to the group agent's action-taking can only be done when the group agent has group-level access to the available evidence. If the collective agent cannot update its representational and goal-seeking states in light of novel evidence, it cannot reason morally about its possible actions, and this would severely impede its moral agency. But available evidence need not be merely a bunch of spreadsheets and data. Here too the collective agent's group-level ability supervenes on all members their abilities to evaluate and register the evidence that is available to them, and not just those at the locus of decision-making (cf. Collins 2019b). This is because the access to available evidence is spread out all over the ranks within the collective agent. For example, in a large organization a high-ranked manager will not notice some faults in a production process. As long as the members can recognize what relevant evidence is available and ensure this is registered within the organization at the appropriate level, the collective agent can not only grasp moral reasons generally, but it can apply such moral reasons based on its available evidence.<sup>14</sup> This is essential in order for the collective agent to be able to act in a morally informed manner.

Control is best explained from a top-down perspective. At the heart of the collective agent lies the decision-making machinery. When sufficiently large, plausibly not all

members of the group agent will take part in this, but only a subset of members higher up in the hierarchy. In some organizations the decision-making may be dispersed over various authorized sub-groups. The point is that collective decisions (typically) result in actions by members.<sup>15</sup> In order for the collective agent to have the relevant control for moral competence, the collective agent must be able to let its goal-seeking states and actions be sensitive to its moral understanding. The ability of the group agent to do so supervenes on members' abilities to vote on relevant matters in a manner that is coherent with the group agent's moral knowledge. This means that the moral understanding and control of the group agent must be inter-linked. If this link somehow breaks down, the agent cannot control its goal-seeking states and actions accordingly even when the group agent knows that it may (potentially) be doing something morally wrong.

This account is minimalist in the sense that the conditions for moral competence are not very demanding. Before discussing how a group agent can cause its own moral incapacity, I will defend my account against some objections that my account is too minimalistic, and that the conditions are too weak to suffice for CMA.

### 3. Is the minimalist account of moral competence too minimalistic?

Some may think that more is needed in order for a group agent to qualify as a moral agent. In light of space, my defense here will not be exhaustive. Instead, I'll target two foreseeable objections that concern moral understanding and moral emotions, primarily because the rebuttal to these objections is instructive for understanding how moral competence (potentially) functions within collective agents and how self-induced moral incapacity can and cannot come about.

Some might object that a group agent has moral competence only if the agent has an *effective moral perspective*, as Hindriks (2018) argues.<sup>16</sup> In order for a prudentially rational collective agent to be capable of moral understanding it must have its own moral perspective. A moral perspective consists of an appropriate set of moral policies that serve to systematically bring moral considerations to bear on the decisions of the agent. These collective moral policies are an instance of what Bratman (2004) calls 'policies of shared valuing', which are general intentions to attribute weights to particular values when deliberating. These moral policies ensure that moral considerations are brought to bear on the decision-making process. For example, the group agent may set environmental standards on product development, safety regulations in case of disasters, or fair-trade policies concerning the import of resources. When the group agent has adopted an appropriate set of such moral policies, it has a moral perspective. This moral perspective must be *effective*, that is, collectively accepted by its members (Hindriks 2018, 12). This means that the moral policies must be integrated with the decision-making procedures and the members must have formed relevant corresponding collective intentions to apply certain procedures and policies.<sup>17</sup> If such moral policies are in place and they are effective, then a substantial proportion of the members will be disposed to bring moral considerations to bear on the decision process, and in this way 'a corporate policy can give rise to a social practice that informs and shapes corporate decisions' (Hindriks 2018, 10). Thus, the objection is that in order for a group agent to have moral understanding, it is not sufficient for a group agent to have (a) the group-level ability to grasp moral



reasons and (b) the group-level ability to relate such reasons to its available evidence, but the group agent must further have (c) an effective moral perspective.

It is instructive to see why one might think this. Plausibly, not every rational agent is a moral agent. Hindriks notes that List and Pettit go some way in making an adequate distinction between rational and moral agency, because if the collective agent's members are systematically constrained from contributing moral propositions to the decision-making process, it essentially functions as a prudentially rational psychopath. However, Hindriks claims that 'it is far from obvious that the absence of such a constraint suffices for a collective agent to be a moral agent' (2018, 7). Members may contribute moral considerations only on an infrequent and *ad hoc* basis, especially when that collective agent is geared towards goals such as profit-maximization. The organizational context 'crowds out' moral considerations. Following May (1992), Hindriks claims that other values may trump or replace moral values within the corporate culture such that members feel bound to uphold these values rather than moral ones. Contributing moral considerations would be overly costly for members. Because of this, Hindriks thinks that grounding moral competence solely in the abilities of members is too weak to suffice for CMA (2018, 8). While any collective agent can put some such policies on the books, not all collective agents can employ them effectively (2018, 9). Due to external pressures, such as the level of competition, the corporation would not apply any of the moral policies it has, meaning its moral perspective is ineffective. In Hindriks' view, moral agents are responsive to moral reasons in a more systematic and robust way. Therefore, according to Hindriks, a collective agent has moral competence only if it has an effective moral perspective.

In my view, requiring a group agent to have an effective moral perspective in order to have moral understanding is far too demanding. I'll first show that a group agent need not have a moral perspective that is *effective* in order to have moral understanding. After this, I'll show that a group agent need not have a moral perspective at all in order to have moral understanding.

First, it is not clear why the moral policies must be *effective*. If there are moral policies in place, but members do not support or follow such moral policies, they are essentially ignoring weighty moral reasons they once collectively endorsed as important and relevant for their action-taking. It is not evident that external pressures make a difference to this. Of course, it is possible that external pressures, such as a competitive environment, may possibly *excuse* certain behavior of CMAs, because the control condition for moral responsibility has not been met. This is something we may encounter with structural injustices, for example. Think of Iris Young's (2011) case of the global apparel industry. In order for lower-level and mid-level enterprises to keep afloat, they must keep labor and production costs to a minimum, resulting in very bad labor circumstances for employees at the end of the global production chain. Possibly, some of these businesses are not to be blamed for their behavior due to the external pressure.<sup>18</sup> But an excuse is not an exemption. These businesses are not exempted from moral responsibility altogether, because the lack of control in this specific instance does not necessarily imply that any of the general abilities essential to moral competence are absent.

But why should we think that an ineffective moral perspective implies that the group agent no longer has the capacity of moral understanding? Why should it be an exempting factor? Suppose that external pressures indeed lead to a corporate culture where moral values are 'crowded out' and replaced with values related to the group agent's prudential



interest. Suppose the group agent does not follow its moral policies and that members contribute moral considerations only on a very infrequent and *ad hoc* basis. Hindriks claims that unless members put forward moral propositions as input into corporate deliberation in a systematic and robust manner, the group does not have moral understanding. But here Hindriks confuses the group ability of moral understanding with a successful exercise of this ability. Hindriks mistakenly takes a *well-functioning* moral agent as the baseline for determining the capacities relevant for moral agency in general. In effect, if the infrequency and ad-hocness of moral contributions by members indeed removes the group agent's ability of moral understanding, as Hindriks claims, then the group-level ability depends on what its members are actually *disposed* to do. But this is clearly false. Consider a law firm that never settles. This is part of the firm's image and success story. All of its members are disposed to never opt for settlement when part of a legal team on a case. However, this clearly does not mean that the firm lacks the ability to settle cases. This is because the group ability supervenes on what members are able to do rather than on what they are (actually) disposed to do. Members are able to perform actions conducive to settling cases, and this is why the group agent is able to settle cases.

Similarly, if the group agent ignores its moral policies and its members contribute moral contributions only on a very infrequent basis, this will most certainly affect the group agent's behavior, but not whether the group agent has the general ability of moral understanding. We should indeed not ignore the possible effects of the organizational context on members. The corporate culture can certainly be important for the moral functioning of the collective agent. But we should not confuse the unlikelihood or irregularity of members' moral contributions with the inability of members to provide such contributions. Of course, a culture of corporate greed can have a strong effect on members' behavior. In extreme cases, such a culture may impose high costs or risks for members who provide moral contributions. However, this does not remove a group agent's moral competence. I'll return to this in Section 4. A corporate culture of greed simply reflects a collective failure to appropriately value morally relevant matters, and, in a sense, it is similar to an individual who continuously ignores weighty moral reasons because of particular desires or emotions. When members only contribute moral considerations on an *ad hoc* basis, often let other values override moral values, and ignore whatever moral policies are in place, the collective agent is not letting moral considerations bear sufficient weight on its collective decisions. If this results in a bad state of affairs, this is exactly when collective agents are to be held morally responsible. This is precisely why certain collective agents end up cutting costs recklessly, pollute the environment or do a number of other things that are morally wrong. When an individual applies moral reasons on an *ad hoc* basis or lets other values trump moral values, this does not necessarily mean that they lack moral agency, and neither does it for collective agents.

Second, setting aside the effectiveness of the moral perspective, a group agent need not have a moral perspective at all in order to be capable of moral understanding. Remember that on my minimal account of moral competence, the group agent's ability of moral understanding supervenes on the members abilities to grasp moral reasons, to provide moral contributions to the decision-making process, to vote on relevant matters and to evaluate and register relevant evidence for the group agent's

action-taking. This does imply that there must not be policies that impede such abilities. Following Bratman (2017), it is helpful to distinguish moral (and other normative) policies from shared policies of procedure. The latter specify (among other things) how a group is to deliberate, for example what decision-making procedure the group follows. Some shared policies of procedure must be in place in order for the group to have an organizational structure. But no moral policy, that is, no shared commitment to give weight to some moral value or consideration during deliberation, needs to be in place for the group to have the ability of moral understanding. Otherwise, this would imply that whenever an issue would come up for which there is not yet a moral policy, the collective agent would not be able to grasp and process such a reason. But this is false. To see this, let's consider how a group agent could come to collectively accept a moral policy.

Suppose a collective agent has a number of shared procedural policies but no background of moral policies. The group agent is a moral *tabula rasa* if you will. It has not instantiated any moral policy yet. A member brings up a moral consideration that she thinks is relevant to consider, for example, we should ensure that our products are environmentally friendly. This consideration itself presents the group agent with grounds for accepting a relevant moral policy. It forces them to think about their relation to the environment as a group agent. Of course, if she brings up something irrelevant for the group agent's current practices and actions, for example, we should save polar bears on Antarctica (or insert any random irrelevant moral consideration), this can and should be ignored. In my view, what matters is that the group agent can acquire moral knowledge (e.g. that it is morally wrong to produce products that harm the environment) by collectively accepting such propositions. The group agent can acquire such moral knowledge even in the absence of any moral policies as long as its members have the relevant abilities. The group agent can translate this moral knowledge into a particular moral policy. If a collective agent consistently fails to adopt such policies, this does not necessarily mean that the agent is not able to acquire moral knowledge.

What Hindriks gets absolutely right, in my view, is to tell a plausible story about how a collective agent can move from the recognition of moral reasons to ensuring that its action-taking is sensitive to such reasons in a robust manner. Effective moral policies robustly ensure that corporate actions are taken in a morally informed manner. Hindriks is certainly right that collective acceptance of moral policies could be key to a *well-functioning* CMA. But that is something rather different from satisfying the necessary conditions for moral agency. We must not equate a group agent's actual responsiveness to moral reasons with the group agent's ability to do so. Of course, a group agent that is morally competent may not always respond appropriately to moral reasons. But that is precisely why in many cases group agents are morally responsible for their immoral behavior.

Some may object that the conditions of my account are nonetheless too weak for CMA, because all moral agents must have the capacity to have moral emotions. Hindriks claims that a collective agent has this capacity exactly if it possesses a suitably broad range of collectively accepted moral policies that are appropriately supported by collective member emotions (Hindriks 2018, 19). For example, when collective member feelings of guilt support and are informed by a suitable corporate policy (e.g. the group agent is to recompense any victims of exploitation), they can be taken to constitute corporate guilt. Others have likewise argued for the possibility of group-level emotions (Björnsson and Hess 2017; Gilbert 2002).

This raises an interesting methodological question: To what extent should we base our conditions for CMA on theories of individual moral agency? We must ask ourselves what sort of abilities are truly necessary for all kinds of moral agency, and which conditions are remnants of an individualistic methodological bias. In my view, postulating group-level emotions as a necessary condition for CMA is unnecessarily demanding. Moral emotions are ‘only’ derivatively necessary for moral agency. To be sure, moral emotions enable humans to recognize and understand moral reasons and the interests of others, and in light of this, we are able to make correct moral judgments and act accordingly. But *group-level* emotions are not necessary in order for *members* to be able to recognize and understand moral reasons and the interests of others. The group agent’s ability of moral understanding supervenes on the individual abilities of members. In order for a group agent to be able to acquire relevant moral knowledge, it need not have group-level emotions, because its members can take up and reason from the group’s perspective (e.g. see Tuomela 2013) and provide moral contributions in relation to this perspective. As long as members are not prohibited from feeling emotions in relation to the collective agent’s doings, and there is a relevant supervenience base of individual abilities, then the group agent has the capacity of moral understanding. Thus, in my view, whether or not collective agents can have group-level moral emotions is irrelevant to the question as to whether they qualify as CMAs.<sup>19</sup>

Of course, if members have relevant member emotions that are supported by appropriate moral policies, this is likely to aid to the moral functioning of the CMA (and this is true even if this does not constitute a group-level emotion as Hindriks claims). And, as Stephanie Collins (2018) argues, although a collective agent cannot itself feel any emotions, it can have duties over the organizational-level functional aspects of emotions and duties to influence and encourage morally valuable emotions that its members feel within and because of their membership in the collective agent. But, again, this is key to a *well-functioning* CMA. We mustn’t confuse factors that aid to the moral functioning of the CMA with the necessary and sufficient conditions for CMA.

While a group agent need not have an effective moral perspective or group-level emotions in order to have moral competence, this does not mean that there cannot be amoral group agents (cf. Hindriks 2018, 9). If any of the general abilities of moral competence (as set out in Section 2) are lacking, the group agent is not a moral agent. With a better understanding of how moral competence (potentially) functions within group agents, I’ll discuss how exactly a group agent can cause its own moral incapacity.

#### 4. Self-induced moral incapacity

If a group agent lacks any of the three general abilities underlying the minimal account of moral competence, then it fails to qualify as a moral agent.<sup>20</sup> To remind the reader, these are the general abilities to grasp moral reasons that possibly govern GA’s actions; to relate such reasons to GA’s available evidence; and to control its goal-seeking states and actions accordingly in light of its moral understanding. This means there are broadly speaking three ways in which a collective agent can become amoral by its own hand. This need not necessarily be done intentionally. It may happen unintentionally over time in a more gradual process without the collective agent necessarily aiming for this. I will consider each ability in turn.

First, consider the general ability to grasp moral reasons that possibly govern the collective agent's actions. To be able to grasp a moral reason, the group agent must be able to form an appropriate truth-tracking belief about the reason. This group-level ability supervenes on the member abilities to grasp moral reasons, to provide moral contributions to the (central or sub-group) decision-making agenda and to vote on relevant matters. As discussed, for a group agent to lack this ability (and thus moral understanding), it is not sufficient that members are simply unlikely to exercise these abilities or do so in an *ad hoc* manner. Enough members must lack these abilities. As List and Pettit already pointed out, the group agent can remove relevant member abilities by redefining the formal procedural policies of its decision-making. In effect, the group agent can restrict the agenda(s) (of its relevant sub-groups) to specific topics, meaning members do not have the opportunity to place relevant moral topics on the agenda(s), or simply ban moral considerations from the decision-making altogether. This means that no member has the ability to provide moral contributions, because the procedural policy prevents any success, meaning the group agent lacks the ability to acquire moral knowledge (as long as such a rule is in place). An explicit ban of moral considerations from the agenda seems to involve an intentional act of the group agent. But when certain new procedural policies restrict the agenda(s) (of various sub-groups) to specific topics, thereby causing the group agent to lose its general ability to grasp moral reasons, this could be unintentional.

Second, a corporation may cause itself to lose its general ability to relate moral reasons to its available evidence. If the collective agent lacks this ability, it becomes impossible for the collective agent to realize that it is in fact facing a morally significant choice. The collective agent becomes unable to make correct moral judgments because it cannot take essential evidence into consideration. Due to certain changes in the organizational structure and hierarchy, the link between the access to evidence and the moral decision-making may be severed due to such restructuring or reorganization. For example, the essential information about relevant moral issues that is gathered at the 'ground-level' will not reach the authorized sub-group or appointed official(s) that are tasked with evaluating the options, because the members at the ground-level are not able to register the relevant evidence and/or put forward relevant moral propositions based on the evidence. Such restructuring could in theory be done purposely in order to escape moral agency. But in practice more often than not this may happen gradually over time. This is, of course, a question of degree, but at some point, the collective agent can no longer be said to have the ability to apply moral reasons based on its available evidence when members do not have the ability to register relevant evidence. The available evidence will not reach the locus of decision-making. This could lower the success-rate of the group agent's moral understanding to such an extent that, in a very high number of scenarios, the group agent is not able to acquire the relevant moral knowledge related to its action-taking.

Third, consider the general ability to control one's goal-seeking states and actions in light of its moral understanding. A collective agent can decide to sever this link by restructuring and re-organizing such that the part of the organization where the locus of moral understanding is placed does not have any authority over the collective action-taking. This may in some cases be conducive to the collective agent's aims. For example, think of a national intelligence agency. Suppose their aim is to gather as much intelligence

as possible about everyone who resides in the country. Any policies regarding privacy will impede that aim. It may set-up a sub-unit that is tasked with moral understanding but ensure via its structure that this sub-unit is powerless and cannot influence the actions of the collective agent. Quite similar to Hindriks' worry about ineffective policies, such a sub-unit may be able to acquire moral knowledge, but this knowledge cannot be translated into decisions about actions, because the sub-unit does not possess the right kind of authority or centrality within the collective agent. Because all moral contributions are processed by this sub-unit, such considerations do not find their way to the locus of decision-making related to the group agent's action-taking. Thus, the locus of decision-making for action-taking is essentially in the dark about any relevant evidence and moral considerations. In such a case, the agent is incapable of translating its moral knowledge into action, because it does not have control over its goal-seeking states and actions *in light of its moral understanding*.

What is interesting is that all three ways of self-induced moral incapacity involve a problem in the group agent's *organizational structure*. Some may think that certain extreme corporate cultures could similarly cause a group agent's moral incapacity. The group agent could either intentionally create such a culture or this may unintentionally develop over time. I find it instructive to see why a corporate culture on its own cannot cause a group agent's moral incapacity.

*Prima facie*, it does seem that an extreme corporate culture can impact the moral functioning of a group agent in a substantial way. Imagine an investment firm where every single member is a cutthroat businessperson with extreme greed. The cost of providing moral contributions related to the firm's action-taking in such an environment is essentially demotion or getting fired. Wouldn't such an extreme corporate culture effectively instantiate an informal decision-making rule that bans moral considerations from entering the decision-making process? How is this any different from a formal decision-making rule?

It is important to clearly distinguish between facts that 'merely' influence whether one *should* do something and facts that affect whether someone *can* do something (cf. Lawford-Smith 2015, 464). The fact that making a moral contribution will lead to demotion (or worse) is a weighty *pro tanto* prudential reason not to do so, but that doesn't mean the agent lacks the ability to do so. Still, one might think that even if the costs do not affect the members' abilities, we must be more precise in our ability-analysis. What matters is that moral considerations materialize in a vote. One could argue that members lack the ability to place moral considerations *on the decision-making agenda*, because the relevant success-rate is extremely low due to the corporate culture. The unwillingness of other members prevents the member from being capable of doing so. It seems that the internal group dynamic of the group agent impedes the relevant member abilities, and the group agent therefore lacks moral understanding.

However, it is important to see that in this case what blocks the member abilities is the behavior of other members, whereas in the other cases of self-induced moral incapacity, what constricts the member abilities is the organizational structure. This is an important difference. We must not leave out other relevant abilities of other members when determining the group ability. Other members are able to be receptive to moral considerations, to support members when they provide such considerations and to refrain from penalizing members when making such contributions. Again, they may have weighty prudential

reasons not to do so, but this doesn't imply they lack these abilities. If the successful exercise of certain member abilities is impeded by the behavior of other members, but these other members are able to refrain from such behavior, then, strictly speaking, the group agent still has the ability. Remember that we learned from the law firm case that group abilities do not supervene on the actual behavior of members. Similarly, if the actual behavior of members lowers the success-rate of certain member abilities, but these members are able to refrain from such behavior, then the group still has the relevant group ability, because the supervenience base also includes the abilities of members to refrain from this behavior. Clearly, members are less likely to exercise these abilities due to the costs imposed by such a corporate culture, but a corporate culture (alone) cannot cause a group agent to come to lack moral competence.

While a corporate culture cannot give rise to self-induced moral incapacity, I have shown that there are nonetheless three ways in which a group agent can self-induce their own moral incapacity due to faulty organizational structures. Next, I can finally argue why this does not give us reason to worry about resurfacing corporate responsibility gaps.

## 5. No reason for skepticism

Exemptions relate to the absence of a general ability necessary for moral agency. But following R. Jay Wallace, it makes sense to further distinguish between two types of exempting conditions: temporary and persistent. Temporary exempting conditions are conditions that make it inappropriate to consider an agent as responsible during a particular restricted segment of the agent's lifetime, e.g. hypnotism, extreme stress, physical deprivation, short-term effects of drugs, and so on. Persistent exempting conditions are conditions that make it the case that the agent's normal state is such that one (or more) of the general abilities necessary for morally responsible agency are absent, for example, insanity or mental illness, extreme youth, (possibly) psychopathy, and the effects of systematic behavior control or conditioning (Wallace 1994, 155).

I am particularly interested in temporary exempting conditions. The reason for this is that temporary exempting conditions do not completely free an agent from responsibility when they are self-induced. For example, consider hypnosis, I think most would agree that a person who is unwillingly hypnotized is exempted from moral responsibility. Consider what Wallace says:

What is distinctive about hypnotism is that the desire on which the agent acts becomes effective in a way that disables the agent's powers of reflective self-control. If posthypnotic suggestion leads the agent to violate our moral obligations, we will suppose that the agent lacks the power to control her behavior in light of those obligations at the time when she acts. But even if the agent is led to act in accordance with our moral obligations, it will not be because she has grasped the reasons that support those obligations and has chosen to comply with them. (Wallace 1994, 175)

During the timeframe of hypnosis, we must treat the person from the objective stance, as an agent to be managed or controlled. It makes no sense to reason with the hypnotized agent because the agent lacks reflective self-control. Next, consider the following case.

**Self-Hypnotized Bank Robber:** A moral agent has found a way to hypnotize himself. The agent has taken precautions to ensure that once hypnotized, there is a posthypnotic suggestion that would lead him to rob a bank. During the bank robbery, the agent reacts to a heroic bank teller, and he shoots and kills the bank teller.

It is counterintuitive to say that the agent is exempted from moral responsibility altogether, precisely because the episode is self-induced. During the hypnotic state we must treat the agent from the objective standpoint. Suppose the police has surrounded the bank, it would be futile to reason with the agent during a hostage negotiation. Of course, the police are likely not to realize the agent is hypnotized, but if they would, this should certainly change their stance towards the agent. Next, suppose that the agent is taken into custody, and the agent is no longer under hypnosis. As soon as his moral competence resurfaces, our stance is altered and we occupy, as Peter F. Strawson calls it, the participant standpoint. We would blame the agent for his misdeeds during the hypnotic state, and rightfully so, because this state was self-induced. There is not really a (worrisome) responsibility gap, because the agent is responsible for its actions as soon as the agent comes out of the temporary state of moral incapacity.

What can we learn from this? Following French (1984a & 2017), it is important to distinguish between *synchronic* and *diachronic responsibility*.<sup>21</sup> Synchronic responsibility concerns responsibility for a morally wrongful action (or attitude) at,  $t_1$ , the time of the action (or attitude). Diachronic responsibility concerns responsibility for an action (or attitude) at a later timeframe,  $t_1 + n$ , than the action (or attitude). Note that Self-Hypnotized Bank Robber is very different from cases where agents become morally incompetent at  $t_1 + n$ , but who were morally competent at  $t_1$ . This happens all the time. Murderers die, bank robbers may develop dementia and forget where they hid the spoils, and so on. In Self-Hypnotized Bank Robber, there is diachronic responsibility but no synchronic responsibility. This lack of synchronic responsibility is not problematic, because at a later stage there is diachronic responsibility.

Now, why is there diachronic responsibility at  $t_1 + n$ ? First, because the episode is self-induced by a moral agent.<sup>22</sup> The agent intended to dodge his moral agency in order to do wrong. As Aristotle (2009, 46) already noted, moral agents are not just responsible for their actions, but also for maintaining their moral capabilities. Second, because the agent is no longer morally incapacitated. The agent has regained all the features so that it can understand the moral address. If the agent would not regain its moral competence, we must continue to treat it from the objective stance. Although the agent is not synchronically responsible, because he is exempted at  $t_1$ , the agent can be diachronically responsible, because his self-induced moral incapacity is not permanent, and he is no longer exempted at  $t_1 + n$ .

Some may think that the agent is synchronically responsible as well, because prior to his self-induced hypnosis, at  $t_0$ , he already had the intention to rob the bank. Hence, we must adjust the timeframe of the action such that the formation of the intention is the start of the action. This is not a good idea, because not every intention results in an action, and acting upon an intention constitutes (additional) wrongdoing. Moreover, even if this would be true for the bank robbery itself, it would not account for the killing of the bank teller. The intention to shoot the bank teller originates during his hypnotic state, therefore the agent cannot be synchronically responsible for the killing. This suggests that synchronic responsibility does not limit diachronic responsibility for the



same event. Not merely in terms of degree, involving increased or diminished degrees of diachronic responsibility (cf. French 2017, 60), but in terms of scope, as in what the agent is responsible for.

Next, suppose that at  $t_1$  a corporation intentionally self-induces its own moral incapacity via one of the ways discussed in Section 4. The corporation subsequently does all sorts of wrong at  $t_2$ , say it causes an environmental disaster. Can we say something similar?

Thomas Donaldson translates Aristotle's claim into an interesting insight about CMA: Analogously, CMA implies responsibility for maintaining corporate moral faculties, such as certain corporate policies, rules, and procedures (1982, 30). A CMA must have the capacity to control its goal-seeking states, actions *and* its organizational structure of policies, rules, and procedures in light of its moral understanding. The control condition of moral competence extends to the corporation's own moral faculties. The group agent has a group-level moral duty to safeguard and maintain its moral decision-making machinery. The corporation clearly violates this group-level duty at  $t_1$ .

Next, the exempting condition is necessarily temporary. In order for a collective agent to lack moral agency, it must lack one of the general abilities required for moral competence. Consider what Pettit says when considering a group agent that has restricted its agenda to propositions of a non-evaluative kind: 'To be sure, members will always be able to change the constitution and to usher in evaluative judgment, but the constitutional restriction may make that difficult and ensure that the exercise of the evaluative ability is only a remote prospect' (2007, 187). Crucially, although the collective agent may not have moral competence at a certain time, if members are capable of changing the collective agent's constitution, that is, to fix the faulty organizational structure, then any self-induced exempting condition is not really permanent for a collective agent. This may be difficult to change, of course, but it would be a mistake to liken this state to say psychopathy, insanity, or any other more persistent condition. This self-induced loss of moral competence can be reversed. Although the collective agent is not synchronically responsible at  $t_1$ , the collective agent can potentially be diachronically responsible at  $t_2$ . The corporation is (potentially) diachronically responsible for its actions *ex post facto*, because the agent is (potentially) culpable for violating its duty to safeguard and maintain its moral decision-making machinery.

Of course, the collective agent can only be diachronically responsible when it regains its moral competence at  $t_2$ . An interesting question is whether the collective agent can reverse its own self-induced loss of moral competence. It may appear that in some cases, namely when the agenda is restricted to non-normative propositions, the collective agent is strictly speaking not capable of recognizing the reasons for changing its constitution. But even if the collective agent is incapable of recognizing the reason to overturn the self-induced loss of moral competence, other agents may enforce this, for example high-ranked members or external (institutional) agents. The need for outside interference by itself is not problematic. Suppose the self-hypnotized bank robber cannot snap himself out of the hypnotic state. This does not mean that he is permanently exempted. Other agents can ensure that he regains his moral competence, and he will again be fully responsible for his misdeeds. If someone would become permanently unstable, we would have to exempt that person from moral responsibility even if the condition is self-induced. However, following Pettit, this is impossible for collective agents. The very same malleability of collective agents that creates the problem also solves the problem.

The example of self-induced hypnosis is intentional, but the same holds for unintentional self-induced states. For example, suppose a corporation unintentionally causes its own moral incapacity because it no longer has (sufficient) access to its available evidence due to restructuring. Following Donaldson, we know that CMAs have group-level obligations to maintain and safeguard their moral faculties, hence the CMA ought to ensure that it continues to have access to its available evidence. If the collective agent fails to do so and allows for a breakdown to occur, then this is a case of negligence. When a CMA undergoes a restructuring, takeover or some other re-organization, the group agent ought to take this into consideration, precisely because it ought to maintain its corporate moral faculties.

Some may object that it is odd to think that there is no difference in responsibility when the CMA intentionally dodges moral agency versus cases where this has happened unintentionally. If there is a difference, which seems intuitively plausible, would this not lead to a possible responsibility gap in the latter case? No, here it is important to keep in mind the distinction between the degree and the scope of responsibility. If the self-caused episode was unintentional, this may affect the degree of responsibility, but not *what* an agent is responsible for.

To sum up, a collective agent can fail to be a moral agent at the time of wrongdoing and yet be (potentially) a morally responsible agent. This is because some exemptions are only temporary. A self-induced temporary exempting condition, whether induced intentionally or unintentionally, exempts an agent at the time of the action. This means that the agent cannot be synchronically or diachronically responsible for its actions during the timeframe of the exemption. However, diachronic responsibility is no longer blocked once the collective agent regains its moral competence. Any agent that regains its moral competence after a self-induced temporary exempting condition can be diachronically responsible for its actions during the timeframe of the exempting condition. The crucial point is that for collective agents any self-induced incapacity is (potentially) temporary because a change in the constitution suffices in order to regain whatever ability that was lost. Hence, a CMA can never by its own hand be completely free from (the potentiality of) diachronic responsibility for its actions.

What does this mean for resurfacing corporate responsibility gaps? The worry was that self-induced moral incapacity reintroduces the same deficits in the accounting books that corporate responsibility is meant to fill. What I have shown is that such deficits are at best temporary, because collective agents can always regain their moral competence due to the 'malleability' of their constitution and be diachronically responsible for their past wrongdoing. Note that individuals may not be able to regain their moral competence. They could permanently damage themselves and subsequently commit wrongful acts. Collective agents cannot do this. Therefore, comparing the temporary deficits of corporate responsibility with the possibly permanent deficits of individual responsibility, if any deficit is to be worrisome, then it is at the individual rather than at the collective level.

## 6. Conclusion

The malleability of the constitution of collective agents gives rise to both interesting questions and problems. For example, if the collective agent was amoral by design, whichever agents constituted the collective agent will incur the responsibility for its actions. One might object that this does give rise to a responsibility gap if the designers are deceased.

This is not particularly worrisome, because individuals can also die after doing wrong. Interestingly, although the possibility of a group agent regaining its moral faculties might be a remote one, it is always a possibility. Given this, even when designers of the amoral collective agent have died, we may hold certain members of the collective agent responsible for failing to transform the collective agent into a moral one. Perhaps we cannot hold them responsible for the initial actions of the collective agent, but they might not be completely free from responsibility either.

Another interesting question is how we should account for cases where other (collective) agents, either internal or external, have caused the collective agent's temporary moral incapacity. In such cases, we can hold the manipulating agent responsible if this causes the collective agent to do something wrong. However, it is not so clear that the collective agent is necessarily completely free from diachronic responsibility when it regains moral competence. The question is whether it is possible for another (collective) agent to incapacitate a CMA without the CMA being responsible for failing to safeguard its moral faculties, because the CMA had a group-level duty to guard its moral faculties from outside interference.

Finally, collective agents may certainly take it to be in their prudential interests to 'shut off' their moral agency. When collective agents do so, given that a collective agent's self-induced moral incapacity is not necessarily permanent, once the collective agents have regained their moral competence, they are diachronically responsible for their wrongdoing. This raises interesting questions about how to ensure and who must bring about this regaining of moral competence and our legal framework concerning incorporating groups. I will have to leave these questions open for now. What I have shown is that self-induced moral incapacity is no reason to be more skeptical about corporate responsibility than individual responsibility. To keep our accountability books in order, we best include corporate responsibility in our ethical theory.

## Notes

1. Regarding (c), List and Pettit argue that just as it is wrong to think the neurons rob an individual agent of control, it is wrong to think that the members rob the group of control. The group agent has control in the sense that it relates as a 'programming cause' to the 'implementing cause' represented by the enacting individual (List and Pettit 2011, 162). For my purposes here, I assume that this is an adequate account of control over its action-taking. With respect to the choice itself, List and Pettit (2011) argue that the group agent develops a 'mind of its own', which gives rise to the kind of autonomy they ascribe to group agents.
2. To avoid confusion, I don't call it self-induced moral incompetence, because incompetence is ambiguous between a poorly functioning moral agent and an amoral agent.
3. Elsewhere I argue that non-agential groups can be collectively responsible (de Haan 2021a & 2021b; see also Chant 2015; Feinberg 1968). The agents are blameworthy for their moral wrongdoing *together* in a purely collective sense. Because not all procedural collectivists accept this, I don't discuss this here. This does raise questions about the collectivists' argument involving corporate responsibility gaps, I hope to address this elsewhere.
4. I think this must be true when ability is related to success (Jaster 2020), and we hold the circumstances appropriately fixed.
5. Pettit makes these claims about general and specific capacities related to the fitness of a *conversable* agent being held responsible (2017, 28), but the conditions and underlying abilities are basically the same as in List and Pettit (2011).

6. Wallace calls this moral accountability. I have in mind the sense of moral agency that justifies being held morally responsible for one's actions, meaning the agent is a legitimate target for reactive attitudes such as resentment and indignation.
7. To be clear, I think an adequate account of ability must include a condition along these lines, but this can be spelled out in different ways (e.g., counterfactually).
8. I don't mean to suggest that all abilities involve actions, see Jaster (2020) on agentive versus non-agentive abilities.
9. Some may object that the supervenience base must be wider than mere member abilities, because the group agent may select an external enactor to achieve its end. I consider such cases as instances of proxy agency. The abilities of external enactors are not part of the supervenience base of the group agent's ability. But the group-ability is in a different way dependent on the enactor's ability, namely in the same way as an individual ability would depend on the proxy's ability. For example, I am able to win an auction, because my proxy is able to bid for me.
10. Sliwa (2017) calls this the Moral Knowledge account.
11. Sliwa does not discuss CMAs, but plausibly a CMA can acquire moral knowledge not only via moral reasoning but also via testimony.
12. Readers are encouraged to plug-in their favored conditions for knowledge. For example, List (2005, 30–31) adopts Nozick's (1981) conditions: An agent knows that *p* if (1) *p* is true; (2) the agent believes that *p*; (3) if *p* were true, the agent would believe that *p*; and (4) if *p* were not true, the agent would not believe that *p*.
13. Stephanie Collins (2019b, 160) claims that 'low-level' operatives outside the decision-making unit are not members of the CMA. For reasons discussed here, I think this is mistaken.
14. Note that it need not be that a member must fully understand the moral significance of the evidence in order to register the evidence. The locus of grasping of moral reasons and registering of relevant evidence may come apart in this way.
15. Collective decisions may also result in actions taken by external enactors. As mentioned, I understand these as instances of proxy agency.
16. A note on terminology: Hindriks uses normative competence, normative perspective, and normative policy as synonymous with moral competence, moral perspective and moral policy. However, it's unclear to me why he calls this 'normative'. Prudential reasons are also normative reasons. A purely prudential perspective is also a normative perspective. A prudentially rational *amoral* agent has normative competence in some sense, but no *moral* competence. I use my own terminology here.
17. This can be spelled out in terms of either Bratman's (2014) shared intentions or Tuomela's (2013) joint intentions (Hindriks 2018, 9–12).
18. Young uses this case to argue that the liability model of responsibility must be complemented with a forward-looking social-connection model.
19. Collins (2022) argues that blameworthiness presupposes the capacity to grasp one's wrongful agency and develops an account of collective moral self-awareness for group agents. Similar to group-level emotions, I don't think this is a necessary condition. We're concerned with individual moral self-awareness because this is (typically) necessary for a person to acquire moral knowledge about her potentially wrongful agency. But a group agent can acquire moral knowledge of its (potentially) wrongful agency without such collective self-awareness as long as members have the capacity for moral self-awareness and can take up the we-mode. In my view, the (controversial) move to attribute collective moral self-awareness to the group agent is not needed to defend CMA. Moral competence, which entails the capacity to acquire moral knowledge, is sufficient for qualifying as a CMA.
20. This shows that my account can provide an answer to what Hindriks calls the Problem of Amoral Agency (2018, 3). My account can differentiate between organized groups that are rational but amoral agents, and organized groups that are rational and moral agents.
21. See also Khoury (2013) for an insightful discussion of diachronic responsibility.
22. Otherwise, adults will be diachronically responsible for the bad things they did as a child. The same holds for amoral collective agents who were designed this way.

## Acknowledgments

I'm very grateful to Franz Altner, Stephanie Collins, Toni Erskine, Frank Hindriks, Lars Moen, Carlos Núñez, Grace Paterson, Herlinde Pauer-Studer, Matt Rachar, Hans Bernhard Schmid, Michael Schmitz, David Schweikard, Anne Schwenkenbecher, Paulina Sliwa, and Elisabeth Widmer for helpful comments. I'm also thankful to the audience of the MANCEPT 2021 workshop 'The Normative Profile of Collective Agents' at the University of Manchester for helpful discussion.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This project has received funding from the H2020 European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme [grant No 740922].

## Notes on contributor

**Niels de Haan** is a postdoctoral researcher at the Department of Philosophy of the University of Vienna. His research focuses primarily on normative ethics and social ontology. He currently works on collective moral agency, collective duties, collective responsibility, and normative reasons.

## ORCID

Niels de Haan  <http://orcid.org/0000-0002-4021-7410>

## References

- Aristotle. 2009. *Aristotle: The Nicomachean Ethics*. Translated by David Ross. Oxford: Oxford University Press.
- Björnsson, Gunnar, and Kendy Hess. 2017. "Corporate Crocodile Tears? On the Reactive Attitudes of Corporate Agents." *Philosophy and Phenomenological Research* 94 (2): 273–298. doi:10.1111/phpr.12260.
- Braham, Matthew, and Martin van Hees. 2011. "Responsibility Voids." *The Philosophical Quarterly* 61 (242): 6–15. doi:10.1111/j.1467-9213.2010.677.x.
- Bratman, Michael. E. 2004. "Shared Valuing and Frameworks for Practical Reasoning." In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, edited by J. Wallace, P. Pettit, S. Scheffler, and M. Smith, 1–27. Oxford: Oxford University Press.
- Bratman, Michael. E. 2014. *Shared Agency: A Planning Theory of Acting Together*. New York: Oxford University Press.
- Bratman, Michael. E. 2017. "The Intentions of a Group." In *The Moral Responsibility of Firms*, edited by E. W. Orts, and N. C. Smith, 36–52. Oxford: Oxford University Press.
- Chant, Sara Rachel. 2015. "Collective Responsibility in a Hollywood Standoff." *Thought: A Journal of Philosophy* 4 (2): 83–92. doi:10.1002/tht3.161.
- Collins, Stephanie. 2018. "'The Government Should Be Ashamed': On the Possibility of Organisations' Emotional Duties." *Political Studies* 66 (4): 813–829. doi:10.1177/0032321717739553.
- Collins, Stephanie. 2019a. "Collective Responsibility Gaps." *Journal of Business Ethics* 154: 943–954. doi:10.1007/s10551-018-3890-6.
- Collins, Stephanie. 2019b. *Group Duties: Their Existence and Their Implications for Individuals*. Oxford: Oxford University Press.

- Collins, Stephanie. 2022. "I, Volkswagen." *The Philosophical Quarterly* 72 (2): 283–304. doi:[10.1093/pq/pqab032](https://doi.org/10.1093/pq/pqab032).
- Copp, David. 2006. "On the Agency of Certain Collective Entities: An Argument from "Normative Autonomy"." *Midwest Studies in Philosophy* 30 (1): 194–221. doi:[10.1111/j.1475-4975.2006.00135.x](https://doi.org/10.1111/j.1475-4975.2006.00135.x).
- Copp, David. 2007. "The Collective Moral Autonomy Thesis." *Journal of Social Philosophy* 38 (3): 369–388. doi:[10.1111/j.1467-9833.2007.00386.x](https://doi.org/10.1111/j.1467-9833.2007.00386.x).
- de Haan, Niels. 2021a. "Collective Culpable Ignorance." *Thought: A Journal of Philosophy* 10 (2): 99–108. doi:[10.1002/tht3.484](https://doi.org/10.1002/tht3.484).
- de Haan, Niels. 2021b. "On the Relation Between Collective Responsibility and Collective Duties." *Philosophy (London, England)* 96 (1): 99–131. doi:[10.1017/S0031819120000364](https://doi.org/10.1017/S0031819120000364).
- Donaldson, Thomas. 1982. *Corporations and Morality*. New Jersey: Prentice-Hall, Inc.
- Erskine, T. 2003. "Assigning Responsibilities to Institutional Moral Agents: The Case of States and 'Quasi-States'." In *Can Institutions Have Responsibilities? Collective Moral Agency and International Relations*, edited by T. Erskine, 19–40. London: Palgrave Macmillan UK.
- Feinberg, Joel. 1968. "Collective Responsibility." *The Journal of Philosophy* 65 (21): 674–688. doi:[10.2307/2024543](https://doi.org/10.2307/2024543).
- French, Peter. 1984a. "A Principle of Responsive Adjustment." *Philosophy (London, England)* 59 (230): 491–503. doi:[10.1017/S0031819100067930](https://doi.org/10.1017/S0031819100067930).
- French, Peter. 1984b. *Collective and Corporate Responsibility*. New York: Columbia University Press.
- French, Peter. 2017. "The Diachronic Moral Responsibility of Firms." In *The Moral Responsibility of Firms*, edited by E. W. Orts, and N. C. Smith, 53–65. Oxford: Oxford University Press.
- Gilbert, Margaret. 2002. "Collective Guilt and Collective Guilt Feelings." *The Journal of Ethics* 6 (2): 115–143. doi:[10.1023/A:1015819615983](https://doi.org/10.1023/A:1015819615983).
- Greco, John. 2009. "Knowledge and Success from Ability." *Philosophical Studies* 142 (1): 17–26. doi:[10.1007/s11098-008-9307-0](https://doi.org/10.1007/s11098-008-9307-0).
- Hess, Kendy. 2014. "Because They Can: The Basis for the Moral Obligations of (Certain) Collectives." *Midwest Studies in Philosophy* 38 (1): 203–221. doi:[10.1111/misp.12024](https://doi.org/10.1111/misp.12024).
- Hills, Alison. 2016. "Understanding Why." *Noûs* 50 (4): 661–688. doi:[10.1111/nous.12092](https://doi.org/10.1111/nous.12092).
- Hindriks, Frank. 2018. "Collective Agency: Moral and Amoral." *Dialectica* 72 (1): 3–23. doi:[10.1111/1746-8361.12215](https://doi.org/10.1111/1746-8361.12215).
- Jaster, Romy. 2020. *Agents' Abilities*. Berlin: De Gruyter.
- Khoury, Andrew. 2013. "Synchronic and Diachronic Responsibility." *Philosophical Studies* 165: 735–752. doi:[10.1007/s11098-012-9976-6](https://doi.org/10.1007/s11098-012-9976-6).
- Lawford-Smith, Holly. 2015. "What 'We'?" *Journal of Social Ontology* 1 (2): 225–249. doi:[10.1515/jso-2015-0008](https://doi.org/10.1515/jso-2015-0008).
- List, Christian. 2005. "Group Knowledge and Group Rationality: A Judgment Aggregation Perspective." *Episteme; Rivista Critica Di Storia Delle Scienze Mediche E Biologiche* 2 (1): 25–38. doi:[10.3366/epi.2005.2.1.25](https://doi.org/10.3366/epi.2005.2.1.25).
- List, Christian, and Philip Pettit. 2011. *Group Agency: The Possibility, Design and Status of Corporate Agents*. Oxford: Oxford University Press.
- Maier, John. 2018. Abilities." *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), URL=<https://plato.stanford.edu/archives/spr2018/entries/abilities>.
- May, Larry. 1992. *Sharing Responsibility*. Chicago: UCP.
- Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge: Harvard University Press.
- Pauer-Studer, Herlinde. 2014. "A Constitutive Account of Group Agency." *Erkenntnis* 79: 1623–1639. doi:[10.1007/s10670-014-9632-y](https://doi.org/10.1007/s10670-014-9632-y).
- Pettit, Philip. 2007. "Responsibility Incorporated." *Ethics* 117 (2): 171–201. doi:[10.1086/510695](https://doi.org/10.1086/510695).
- Pettit, Philip. 2009. "Corporate Responsibility Revisited." *Netherlands Journal of Legal Philosophy* 38 (2): 159–176.
- Pettit, Philip. 2017. "The Conversable, Responsible Corporation." In *The Moral Responsibility of Firms*, edited by E. W. Orts, and N. C. Smith, 15–35. Oxford: Oxford University Press.
- Sliwa, Paulina. 2017. "Moral Understanding as Knowing Right from Wrong." *Ethics* 127 (3): 521–552. doi:[10.1086/690011](https://doi.org/10.1086/690011).

- Sosa, Ernest. 2015. *Judgment and Agency*. Oxford: Oxford University Press.
- Strawson, Peter F. 1962. "Freedom and Resentment." In *Proceedings of the British Academy, Volume 48*, edited by G. Watson, 1–25. Oxford: Oxford University Press.
- Tuomela, Raimo. 2013. *Social Ontology: Collective Intentionality and Group Agents*. Oxford: Oxford University Press.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.
- Whittle, Ann. 2010. "Dispositional Abilities." *Philosopher's Imprint* 10 (12): 1–23. <http://hdl.handle.net/2027/spo.3521354.0010.012>.
- Young, Irish Marion. 2011. *Responsibility for Justice*. Oxford: Oxford University Press.