This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Hirvonen, Onni

Title: The Problem of the First Belief : Group Agents and Responsibility

Year: 2020

Version: Published version

Copyright: © 2020 Onni Hirvonen, published by De Gruyter

Rights: CC BY 4.0

Rights url: https://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Hirvonen, O. (2020). The Problem of the First Belief : Group Agents and Responsibility. Journal of Social Ontology, 6(1), 1-20. https://doi.org/10.1515/jso-2019-0029

Article

Onni Hirvonen*

The Problem of the First Belief: Group Agents and Responsibility

https://doi.org/10.1515/jso-2019-0029

Abstract: Attributing moral responsibility to an agent requires that the agent is a capable member of a moral community. Capable members of a moral community are often thought of as moral reasoners (or moral persons) and, thus, to attribute moral responsibility to collective agents would require showing that they are capable of moral reasoning. It is argued here that those theories that understand collective reasoning and collective moral agency in terms of collective decision-making and commitment – as is arguably the case with Christian List and Philip Pettit's theory of group agency – face the so-called "problem of the first belief" that threatens to make moral reasoning impossible for group agents. This paper introduces three possible solutions to the problem and discusses the effects that these solutions have in regard to the possibility of attributing moral responsibility to groups.

Keywords: Christian List, collective beliefs, collective responsibility, group agency, moral personhood, Philip Pettit, philosophy, social ontology

1 Introduction

In our everyday practices and everyday language use, it is commonplace to find groups playing all kinds of roles. We see groups in causal-structural or explanatory roles in politics and social sciences, but, more often than not, we attribute agency and moral responsibility to them as well. States negotiate with each other and blame each other, new cultural and social movements challenge traditional institutional agencies; corporations are taken to court, and so forth.

To attribute *moral* responsibility to collective entities, we need an understanding of what constitutes moral agency or moral personhood in the first

***Corresponding author: Onni Hirvonen**, Department of Social Sciences and Philosphy, University of Jyväskylä, PO Box 35, Jyväskylä, FI-40014, Finland, e-mail: onni.hirvonen@jyu.fi. https://orcid.org/0000-0001-8493-7420

Open Access. © 2020 Onni Hirvonen, published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

place, and whether groups are the sort of entities that manage to fulfill the conditions of moral agency. It should be noted here that some of the issues that will be presented below could be taken to apply to a broader notion of collective agency, not just collective moral agency. This paper focuses on whether groups fulfill the conditions for moral agency for two reasons. (1) There are entities that possess thin forms of agency – e. g., most non-human animals – that are not moral subjects, although they can be moral "patients" who have a role in moral considerations. The philosophical problems that apply to reflective forms of agency, like moral agency, need not be issues for thinner forms of agency, or agents that are not expected to have second-order beliefs or reflective capacities. (2) Many interesting practical issues that are connected to groups are often spelled out in terms of moral agency or moral personhood.[1] In short, the general aim of the paper is to map out our moral landscape – even if some of the arguments could be used in the broader context of "reflective agency in general". This paper presents a thin view of moral personhood that emphasizes the capacity to reason and membership in a moral community (Section 2) – as opposed to thicker views, which would include an account of moral emotions or moral phenomenology. After the thin view of moral personhood is set up, the paper introduces one particular philosophical problem – the problem of the first belief – that seems to threaten the possibility of taking group agents as proper moral agents or persons (Section 3). After setting out the problem, three different solutions are presented (Section 4), all of which generate further problems. The paper finishes with a short discussion of these issues and sketches two potential philosophical routes that are available for understanding the moral standing of groups (Section 5).

2 Groups and Moral Responsibility

To be morally responsible for something often assumes that one had a part to play in bringing about a morally relevant event. In other words, it requires that one has been, at least in some very loose sense, *causally* responsible for the morally relevant event. It is equally commonplace to notice that causal responsibility does not

1 For example, recent contributions to the literature on collective responsibility and group duties (Collins 2019; Lawford-Smith 2019) come from the perspective of group agency. Related to responsibility, collective punishment can be thought of as requiring group agency (see, e. g., Malle 2010). This paper focuses mainly on List and Pettit's influential view on group persons that is presented in detail in their *Group Agency* book (2011) and which functions in the background of many contemporary views.

entail moral responsibility. We do not find lightning *morally* blameworthy even if it is the causal reason for the demise of an unlucky camper. It seems intuitive to attribute moral responsibility only to moral agents or moral persons.[2] The aim of this section is to give a general picture of what falls under the concepts of moral agency and moral personhood. This is done in order to set up the problem of the first belief, which suggests that it seems impossible for groups (or group agents) to fulfill those conditions that would make them moral agents or persons.

The first thing to note is that moral agency and moral personhood – although often used interchangeably – do not necessarily mean the same thing. In general, moral agency is attributed to a special subgroup of agents: those who are capable of conforming to some demands of morality. At a minimum, this means that one would expect that a moral agent is able to consider what is good and what is bad. That is, to be a moral agent is to be a reflective agent with, at least, some level of self-understanding and moral consciousness, though it is worth noting that there is philosophical disagreement about the exact conditions of moral agency (Haksar 2005). Talk of moral personhood, as introduced below, also focuses on what makes a moral agent but it adds an additional layer to this by analyzing the rights and responsibilities of moral agents in a social setting. In this paper, the focus is on moral personhood. This is in part for the aforementioned reason that considerations of moral personhood also include considerations for moral agency (at least if we accept that only agents can be persons) and, in part, because this article aims to problematize Christian List and Philip Pettit's account, which explicitly focuses on personhood.

Moral personhood, as understood in this context, has two key aspects: it is a psychological concept and a status concept (Laitinen 2007, pp. 248–249). This is to say that we need certain psychological (or, more broadly, agential) capacities for certain types of action, but we also need recognition of our status as moral agents to really count as such in social situations. The textbook example of an agent with capacities but without status is a slave that is treated only as a tool, a means to an end, and not as a moral person in the sense of having a recognized moral status (McBride 2013, p. 62).[3] At the same time it is clear that most slave owners did not in reality consider slaves as mere tools, as any "immoral" deeds by a slave would certainly be punished. Nevertheless, all this indicates that moral personhood or

2 Though this intuition is debated in collective responsibility literature (see, e. g., Wringe 2020), this is the perspective that is adapted in this paper. As Stephanie Collins condenses it "only agents can bear duties" as "only agents can act" (Collins 2013, p. 231).

3 Peter Strawson describes a similar situation: "If your attitude towards someone is wholly objective, then though you may fight him, you cannot quarrel with him, and though you may talk to him, even negotiate with him, you cannot reason with him." (Strawson 1968, p. 79).

moral responsibility is also a status that is attributed to entities that, at least potentially, fulfill certain necessary agential conditions. Thus, if we are interested in the justifiable status attribution of moral responsibility or moral personhood, we must turn our attention to the conditions of that attribution.

Here the aim is a modest one: even if we do not have fully-fledged necessary and sufficient conditions for justified attribution of moral agency, it might be possible to give a list of *some* necessary conditions that need to be fulfilled in order to count as a morally responsible agent. Although we might not be able to offer a clear-cut definition of the full conditions of moral personhood, it is still possible to outline some minimal requirements of personifiable agency that are widely acknowledged.

What are these widely acknowledged conditions? In the context of the so-called "standard view" of personhood, these include linguistic capacities, rational capacities, communicative skills, an ability to take an intentional stance, and so forth (see e. g., Goodman 1992 for a short listing and Dennett 1976 for a detailed and influential explanation). These describe a social and rational agent that, from the perspective of morality, should also be able to act with reference to right and wrong. One way to reformulate all of this is to say that moral agency requires an ability to give moral reasons for actions. A moral agent ought to be responsive to moral reasoning and have access to a moral language of good and bad, or acceptable and unacceptable, actions. This strand of thought is alive and well in contemporary accounts of collective personhood: for Rovane (1998, pp. 85–86) it is precisely rationality and responsiveness to reasons that make an agent, while List and Pettit (2011, p. 173) emphasize the ability to function as a capable member in a system of obligations. The key idea, which is shared in the above formulations, is that moral responsibility requires, at least, the capability to offer others reasons to do so-and-so and to be responsive to reasons provided by others. To be a moral person is to be a part of the game of giving and asking for (moral) reasons. This is a highly social and relational view of morality and moral personhood, which states that, at least in principle, moral persons should be able to answer if asked for reasons for their actions. It should be noted though that this does not mean that this interaction is always necessary as in the actual world moral persons can in fact act in isolation. However, this does mean that if an agent is such that it functions in a moral realm, some sort of normative or moral community is needed, at least in an imaginary form (see, for example, McBride 2013, pp. 67–70 for the importance of virtual recognition).

The suggestion here is that moral reasoning can be used as a minimal condition for moral agency.[4] To count as a moral agent one needs to have – at least – (morally

4 This is not a claim that only those actions that result from moral reasoning would be counted as morally relevant. Rather, moral agency requires a capability or a potential for moral reasoning – even if not all of the agent's actions are subject to reasoning.

relevant) action-guiding beliefs and one needs to adhere (in a minimal sense) to constraints of rationality. As Rovane (1998, p. 130) puts it, "*the whole enterprise of agency-regarding relations presupposes that each person has a commitment to achieving overall rational unity within its own rational point of view*" [original italics]. It would be impossible to attribute responsibility justifiably to an entity that could not, even in principle, conform to any constraints of rationality in its beliefs or have the potential to be a rational agent. This is a necessary condition in the sense that if this condition is not fulfilled, it is clear that one cannot be taken as morally responsible. However, for the purposes of the current endeavor, we do not need to accept that it amounts to a sufficient condition. Moral reasoning on its own does not need to guarantee moral personhood and, indeed, quite often reasoning is complemented with capacities like self-determination and control (see, e. g., Tollefsen 2015, pp. 117–118 for a discussion of this in the context of group agents).

How does this thin picture of moral agency fit together with the idea of group-level moral responsibility? List and Pettit (2011, pp. 174–178) argue that we have a long-standing practice of using personifying phrases in the context of groups. Furthermore, there are cases where treating groups as morally responsible is the only viable option as no individual can be blamed, but a blameworthy deed has, nevertheless, been done by a group (for examples, see Pettit 2007; French 1984).

In accordance with the above conditions of moral agency, holding groups responsible would require that they are moral reasoners and agents capable of moral deliberation. Even if there are groups that have a relatively independent ontological standing – with non-reducible properties and a role in bringing about certain events in the world – this would not yet guarantee that a group is a reasoning moral agent. For example, we may think of social structures and social facts as having an effect on individuals much like any other natural phenomenon does.[5] They might have a causal connection to what we do, but if laws of sociology function largely in the same manner as laws of nature, then blaming structure-like social settings in a moral sense does not make much sense if they are understood as lacking agency. Here the possible moral blame would lie with the individuals who (more or less collectively) uphold the structures that have caused the harm.

Are there groups that are capable of moral reasoning? Are there rational group agents with moral beliefs? Rovane (1998, pp. 137–141) makes the case that groups may constitute a rational point of view from which they reason in accordance to their commitment to pursuing certain aims and goals. This idea has also been taken up and defended by List and Pettit (2011) who state that groups need to make

5 We may here think of Durkheimian social facts or Searle's (1995) institutional facts, which are non-agential but socially constructed parts of the world that have a major impact on individuals' lives.

decisions that are consistent in relation to their goals. These goal-oriented or purposive groups acquire their beliefs with various kinds of thought aggregation processes and collective decision-making mechanisms. For example, a group may form its beliefs by committing to the decisions that are reached through, more or less democratic, voting procedures. To appear as an effective promoter of its given purpose, the commitments that guide the actions of the group need to be consistent and rational. Those groups that decide to aim for consistency and coherence by reasoning from the perspective of collective decisions and commitments are said to "collectivize" reason and they are taken as loci of rational reasoning (List and Pettit 2011, pp. 69–71; Pettit 2003, pp. 176–177). One of List and Pettit's central claims is that the nature of some of the thought aggregation processes is such that the beliefs attributed to groups are not straightforwardly reducible to the individuals' beliefs. This claim is defended with an impossibility result, which shows that there is no collective decision-making mechanism that would unproblematically retain responsiveness to individuals' beliefs and the rationality of collective commitments (see List and Pettit 2004). As the rationality of the commitments is more important in relation to the fulfillment of the purpose of a group, it is reasonable to let go of the responsiveness to the individuals' beliefs. This, in turn, makes "collective beliefs" discontinuous with their individual constituents, although group beliefs still supervene on individuals' beliefs. List and Pettit take this to show that groups have, in a sense, minds of their own and, therefore, some of the so-called mental properties should be understood as properties of a group agent. From the perspective of potential moral agency, it is not far-fetched to think that groups' beliefs could also have moral content and moral relevance in the senses that (a) it is possible for a group to formulate beliefs about moral issues and (b) that a group could be held responsible for the beliefs it holds.

List and Pettit's theory is a version of a broader family of theories of group beliefs that base group beliefs on collective decision-making, collective acceptance, and reasoning from the group's perspective.[6] However, this "collective decision-making and commitment"[7] model of group beliefs faces a challenge. If a

6 Other famous theories belonging to this line of thought are those of French (1979) and Rovane (1998). See Hirvonen (2017a) for a short analysis of the similarities between French, Rovane, and List and Pettit.

7 This is not the way that List and Pettit would describe their own account. Instead, they posit it in functionalist terms, seeing group as a functional whole. However, they also put a heavy emphasis on the decision-making mechanisms and practices which need to be – at least tacitly – accepted by the group members. In short, in the case of groups the functional whole needs the acceptance or practical commitment of its individual parts to exist. Pettit (2003, p. 177) seems to hint towards this when he talks about the necessity of the group appearing as an effective promoter of its purpose because otherwise it would lose its members and cease to exist.

newly constituted group mind is like a blank slate and the group's beliefs are formed one by one, this leads into the so-called *problem of the first belief*. Though the following argument focuses on List and Pettit's well-known view, it is meant to apply more generally to any possible theories that hold that group beliefs are formulated through decision-making procedures, one belief at a time. Although what follows might appear as a downright criticism of group agency, the main aim is to show that the problem of the first belief can be avoided, but with a cost, as the strategies used for the avoidance have a major effect on the ways in which we can hold groups to be responsible.

3 The Problem of the First Belief

The problem of the first belief, as introduced by Arto Laitinen (2014), concerns the possibility and understandability of group's beliefs and applies to all group agents that form their beliefs in a proceduralistic manner. The problem results from combining two plausible assumptions about how groups formulate their beliefs with a holistic understanding of beliefs (and meaning) in general. The first assumption is that groups form their beliefs one by one. It is believable that a group, at the moment it comes to exist, has not yet formed any beliefs. The first belief has to be collectively accepted and formed through the use of group's decision-making procedure (and this differs from individual human development where the holistic framework of meaning is gradually learned). The second assumption is the non-reducibility of group beliefs and their apparent discontinuity with individual beliefs, as emphasized by List and Pettit. If we take the discontinuity between collective and individual beliefs to be true, groups seem not to be able to straightforwardly inherit any beliefs from their members. When these two assumptions are combined with holism, the result is what is called "the problem of the first belief". This problem of having or believing only one belief can be formulated in three ways: as an *instantiation* problem, as a *conceptual* problem, and as a *psychological* problem.

In the *instantiation formulation*, the problem is that the solitary first belief is not actually a belief at all. This is based on the idea that beliefs and other mental states are not atomistic entities that could exist one by one, without the support of other mental states. The instantiation formulation relies on the holistic view of beliefs, according to which beliefs are not of the sort of entities that can be instantiated in the world by themselves or alone, without being in connection with other beliefs. While the instantiation formulation stresses the impossibility of having solitary beliefs at all, the *conceptual formulation* says that a singular belief does not have any meaning or conceptual content (Laitinen 2014, p. 44). This claim

is based on holism about meaning, which states that "the content of every belief depends to a large degree on a broad range of one's related beliefs" (Schwitzgebel 2015, section 3.2). A belief system containing only one belief does not actually have any content, because the content of a belief is tied to its relations to other beliefs. The *psychological formulation* emphasizes that a subject who has only one belief cannot apprehend the meaning of its sole belief (Laitinen 2014, p. 45). Michael Bratman condenses this last formulation in stating that "to talk of a *subject* who intends is to see that subject as a center of a more or less coherent mental web of [...] intentions and cognitions" (Bratman 2014, p. 127). Thus, if a group has only one belief, it would not be fit to be held responsible, especially if we consider responsibility to be something that requires reasoning, understanding, control, and attitudes towards one's own beliefs – that is, subjectivity.

This argument for the impossibility and senselessness of having only one belief poses a challenge for proceduralistic "collective decision-making and commitment" theories of group beliefs. For example, in List and Pettit's view, groups have their own minds and beliefs in a very literal sense, and a group's intentional states are formed through a decision-making process, and they are discontinuous with the minds of individual members. However, it seems implausible that any mechanism of formulating group beliefs which relies on collective decision-making and commitment would create a suitable holistic network of beliefs. In short, forming "beliefs" one by one through a decision-making mechanism seems to make groups vulnerable to the problem of the first belief, rendering literal talk of their beliefs nonsensical. This makes it impossible for groups to be capable moral reasoners. After all, according to our thin description of moral agency, to be a moral agent one needs to have at least beliefs about good and bad, or beliefs with moral content.

4 Solving the Problem

I argue here that there are at least three possible solutions to the problem of the first belief, all of which are based on relaxing some of the assumptions that cause it. Thin functionalism (*a*) lets go of holism about beliefs and embraces a functionalist conception of beliefs. Social externalism (*b*) expands holism beyond singular minds in such a way that the meaningfulness of groups' beliefs is attained through a shared holistic framework of meanings. The bottom-up inheritance model (*c*) loosens the discontinuity and separateness of group beliefs from individual beliefs and argues that groups acquire a holistic web of beliefs from their members. All three options manage to solve at least some versions of the problem of the first

belief. However, either they do not manage to solve all instances of the problem, or they raise other philosophical problems.

a. Thin Functionalism

As the instantiation formulation states, having a singular belief alongside a commitment to holism about beliefs is contradictory as, according to holism about beliefs, the singular belief is not really a belief at all. To solve the contradiction, one would either need to have a suitable web of beliefs from the get-go or let go of holism about beliefs. Here the focus is first on the latter option as it is readily present in the collective agency literature in the form of endorsing a thin functionalist view of beliefs (and agency). If beliefs are simple functional states that do not need to relate to other such states, a group with only one belief may still have a genuine belief. For example, in List and Pettit's (2011, p. 20) account, beliefs are defined as representational states, while agency follows a simple belief-desire-action model where representational states (beliefs) are combined with motivational states (desires) and an ability to act according to these. As beliefs and desires can be defined as functional properties or dispositions of a system, the problem of the first belief does not arise because there is no requirement for these states to be related to any other states of the same kind.

However, the thin functionalist solution creates several concerns. First, there are general doubts about the functionalist theory of beliefs (or mind) itself (see, for example, Block 1980 for a summary of problems of functionalism). Though functionalism gives us simple standards to see if some being has a belief or not, the definition itself can be criticized for being too liberal, as it forces us to accept that all sorts of entities from simple animals to robots and computers have beliefs (or minds). On its own, it does not seem to be able give us the necessary tools to separate the interesting cases like higher animal and human minds from simple mechanical constructs, bacteria, and plants. This, in turn, makes the concept of belief (or mind) so broad that it becomes useless in separating the interesting cases of self-understanding mindedness, which is needed for moral agency, from simple functionalist cases of having a mind.

Even if we grant, for the sake of the argument, that simple representational states are beliefs, and that the demarcation between "simpler minds" and "more complex minds" can be drawn in a meaningful way, thin functionalism still runs into problems. Groups with singular beliefs might have simple minds, but at the level of human agency and language, commitments to holism seem more commonplace. For example, at this level Pettit also argues for holism (see Pettit 1996, chapter 4, *For Holism, Against Atomism*, or Pettit 1998, *Defining and Defending Social Holism*). In other words, even if agency in its simplest forms could

be defined functionally, conceptual understanding and linguistic skills are often taken to require holism. Because List and Pettit's ultimate goal in their *Group Agency* book is to argue for the personhood of groups, the thin functionalist approach, which allows groups to be simple minded agents, seems inadequate: personhood is something that is taken to require membership in a system of obligations which, in turn, requires complex capabilities for language and communication.

The *instantiation* version of the problem of the first belief does not necessarily arise with thin functionalist minds, but the *conceptual* and especially *psychological* versions become prominent with personifiable group agency, given that one cannot be a person if one is not a relevant member of a broader framework of meaning and understanding. Though groups may well have beliefs in the thin sense, this does not yet show that they would be fully-fledged members of a linguistic community, a developed system of obligations, or a part of the game of giving and asking for reasons. In other words, thin functionalism does not manage to solve the issue of (self-)understanding. Even if a group's one belief is a genuine belief, the group still does not have any other beliefs to understand its sole belief (as stated in the psychological formulation), nor any broader conceptual framework within which its sole belief would make sense for itself or any others (as stated in the conceptual formulation).

To conclude the discussion on thin functionalism, let me offer a brief aside. Namely, what would it require to think that the thin functionalist mind is enough for moral personhood? Clearly this would force us to endorse a much broader conception of moral agency than the one presented by the standard view: moral responsibility could be attributed to beings that lack self-reflective capacities. This goes strongly against the commonly held intuition that moral persons must be at least in principle capable of reflecting on their deeds and understanding them. After all, that is exactly what is taken as the difference-maker between quite complicated animals and reflective moral human beings. For example, this difference is precisely what Christine Korsgaard makes clear in saying that: "We cannot expect the other animals to regulate their conduct in accordance with an assessment of their principles, because they are not conscious of their principles. They therefore have no moral obligations." (Korsgaard 2004, p. 87.) Abandoning some forms of understanding and self-reflection in an account of moral personhood seems to be a price that we are not willing to pay. Thus, while a charitable reading of thin functionalism solves the instantiation problem of the first belief, attributing moral responsibility to group agents on this basis might be too hasty. Indeed, solving the conceptual and psychological formulations of the problem seems like a more fruitful way to go than relaxing the standard conditions of moral personhood.

b. Social Externalism

The second solution accepts holism (about beliefs and about meaning) and suggests that a group's solitary belief gets its conceptual content from the social surroundings of culturally shared meanings. According to this kind of "social externalism", the holistic network of beliefs need not be within the mind of a single agent but, instead, any belief only makes sense as part of a collectively shared cultural web of meanings. As long as an agent's beliefs are connected to this holistic cultural network of meanings, it does not matter if it has only one belief. The instantiation version of the problem of the first belief does not arise, as the belief is connected to other beliefs in the commonly shared web and same applies, *mutatis mutandis*, to the conceptual formulation. The cultural web of meanings is understood as a shared public background that is needed in any communication. The meanings might be contestable, but they are not empty. While List and Pettit together lean towards thin functionalism, on his own Pettit also comes close to social externalism when he states that "the contents of [a group's] judgments and intentions will inherit determinacy from the presumptively determinate words that are used by its members to express those contents" (Pettit 2003, p. 183).

If we understand groups' beliefs in the sense suggested by social externalism, at least others can take an intentional or communicative stance towards the group and understand its beliefs as meaningful. However, even if a group's beliefs are meaningful, this does not guarantee that the group can understand its own beliefs. While the conceptual problem is solved as others can understand the group, the psychological problem remains, as a group with a singular belief does not have the means to understand its sole belief.

Following social externalism would make groups, as Laitinen (2014, p. 47) puts it, like pawns in a game of chess. They have a meaningful role in social practices – they are centers or carriers of commitments – but they are not players in the sense that they would create or track those commitments themselves (see also Hindriks 2008 for a view that corporate agents owe their existence to external recognition). In this picture, we can attribute beliefs (with moral content) to groups but they lack any self-understanding of those beliefs. Though this seems to fit with our everyday practice of attributing statuses and mental states to groups, the nature of groups' agency would be quite different from that of individual self-conscious agency. The psychological capacities for self-understanding are not present in the case of groups.

In relation to responsibility, how far does solving the instantiation and conceptual versions of the problem of the first belief get us? If group attitudes and beliefs are understandable within a culturally shared moral sphere, it is plausible that we can attribute moral beliefs to groups as well. These are singular beliefs of a simple agent that make sense in the broader social setting. However, it is highly

unlikely that a mere belief (with some presumably moral content) is enough to constitute moral agency or moral personhood – at least if we stick to the standard view. Although there seems to be no conceptual problem in attributing moral beliefs to collective entities, the psychological issue remains relevant: the group does not understand its belief, and despite the possibility of holding a belief that has moral import (for those who understand it), the group does not seem to have the required psychological properties of a robust moral agent. Korsgaard's point on animals and their lack of consciousness of moral principles is relevant here. It seems that we are again at a philosophical crossroads: if groups are to be morally responsible, we can look for a further solution to the problem of the first belief, or we can instead try to loosen some of the standard conditions of moral agency and/ or moral personhood.

c. Bottom-Up Inheritance

The third solution to the problem defends a more robust group agency by claiming that group also has its own "personal" web of beliefs that it inherits, "bottom-up", from its members. Where thin functionalism and social externalism failed to solve the psychological problem and show how groups could have personal understanding, the bottom-up inheritance model is from the outset more promising: it retains an internalized notion of a web of beliefs that enables both meaningful beliefs and self-understanding. However, there are certain philosophical moves that need to be made to achieve this. Namely, we need to relax the groups' reliance on explicit decision-making in their belief formation and also relax the claims of discontinuity between individuals' and groups' beliefs.

How does a group "inherit" its web of beliefs from its members? A straightforward answer is to rely on common knowledge and shared beliefs within the group.[8] If all the members of the group believe that the Earth is flat, then the group believes that the Earth is flat. When the set of members' shared beliefs is large enough – that is, it includes enough beliefs for the group to function normally in practical contexts of moral action – the group will count as an agent with its very own holistic web of beliefs. If this is the case, none of the versions of the problem of the first belief will arise, as the group's beliefs are part of a holistic web of beliefs

8 An even broader option would be to rely on even wider shared beliefs available in the population as suggested by an anonymous reviewer for this journal. However, here I want to limit the discussion to the beliefs of group members for two reasons: one, if beliefs are shared in a wider population, it can be assumed that they are also shared within the group and, two, if we agree that group beliefs are "holistically supervenient" (see List and Pettit 2011, pp. 69–72) on their members' beliefs, this seems to foreclose the possibility of inheriting beliefs from the wider population.

and thus related to other beliefs (contra the instantiation formulation) through which they get their meaning (contra the conceptual formulation), and as the group itself is the locus of its web of beliefs, it has the conceptual resources to understand its own beliefs (contra the psychological formulation).

One consequence of the bottom-up inheritance model is that self-reflective agency (and thus also genuine group personhood and group responsibility) will be limited to those groups that have a suitably large set of shared beliefs. If the inherited beliefs need to be shared by all (or at least most) group members, only groups that are uniform enough can be held to be responsible. In the case of considerations for collective responsibility, this might leave out some of the interesting cases because as the number of group members grows, it becomes less likely that a belief that is shared by everyone (or most) in the group. For example, many groups that are of interest for social theory, like cultural groups or states, are potentially unfit to be taken as responsible. This worry, however, might be exaggerated as we could as well assume that there exists a culturally shared background, which includes a broad range of beliefs. Furthermore, it might be a good choice to leave large non-organized groups like cultures outside of responsibility attributions as they do not function as centers of collective decision-making and it is questionable if they are agential entities at all.[9]

Thus, limiting the set of potentially personifiable group agents only to certain kinds of groups seems defensible. However, inheriting beliefs from members also leads to a further philosophical worry, namely, the conflation of shared beliefs and group beliefs. List's categorization of aggregate, common, and collective/corporate attitudes is useful for seeing the problem.[10] According to him, aggregate attitudes are "shorthand summaries of the underlying individual attitudes and need not generally be action-guiding for the collective or its members" (List 2014, p. 1606). These are "summative" or "survey" attitudes that can be easily reduced back to the individual attitudes on the matter at hand. Any group arguably has all kinds of aggregate attitudes they but need not endorse them or commit to them as a group. Aggregate attitudes may not be even known by the group. A common attitude, in turn, comes close to what has been referred to above as common knowledge or shared beliefs. According to List (2014, p. 1609) it is "an attitude held by all individual members of the collective, where their holding it is a matter of

9 This is the line that is often taken in collective responsibility debates. See, for example, Lawford-Smith (2019) or Collins (2013 2019).

10 A similar distinction is found in Gilbert's (2013, Chapter 8) discussion on shared values. She states that summing up beliefs in a group – even with added element of common knowledge – is not enough to bind participants together or to constitute a social group, whereas sharing values in a more collective sense would require a joint commitment to believe as a body that something has value (Gilbert 2013, pp. 191–192).

common awareness". Individuals share beliefs on a certain matter and this is known by everyone in the collective. Lastly, there are collective/corporate attitudes that are held by a group as an intentional agent. This requires a supporting theory of collective agency and an account of how the beliefs are formed.

As previously stated, some accounts, like the one given by List and Pettit, emphasize conscious commitment and group decision-making mechanisms. This leads back to forming beliefs one by one (or in very limited sets), which, in turn, leads back to the problem of the first belief. The bottom-up inheritance model solves the problem by including the category of common attitudes into the category of collective/corporate attitudes. What is problematic about this is illustrated by Gilbert's (1992, p. 273) example of two groups, the Library Committee and the Food Committee of a residential college. Both groups consist of the same members and most of the members believe that the college students consume too much starch. This is also common knowledge amongst the committee members. Yet the two groups differ in that the Food Committee has in its meeting formed a belief that the students consume too much starch, whereas the Library Committee has no opinion on the matter. The example aims to highlight that even if there are two groups with the exactly same shared background of beliefs, not all of them can be attributed to both of the groups as group beliefs. This points towards the conclusion that groups do not automatically inherit the beliefs of their members.

If we consider the conflation of shared beliefs and group beliefs from the perspective of moral responsibility, shared or aggregate beliefs do not seem to have the same standing as group beliefs. The group itself does not formulate its shared or aggregate beliefs, they are accidental to it. Can it still be held responsible for them? One option is to think that group can be held responsible only for the corporate/group beliefs and these beliefs get their meaning through the shared cultural background, but this solution leads us back to social externalism, which makes the moral agency of groups metaphorical at best.

The second option[11] is to say that the group's beliefs about the *world* are determined only by its decision-making procedure. However, the common knowledge of the members is inherited in the sense that it determines the meaning of the group's beliefs but not whether the group has the belief or not. This meaning-inheritance can be interpreted in two ways. First, we can understand this inheritance as a limited version of social externalism. Here the reference point for the meaning is not the broader cultural context but rather the members of the group. However, it seems clear that this version faces the same problems as broader social externalism.

This second interpretation of inheritance would make the common beliefs about meanings literally the group's beliefs and, thus, the group would have a

11 Thank you for an anonymous reviewer for this journal for suggesting this option.

whole range of beliefs about meanings of concepts that it did not produce through its decision-making procedure. If we dismiss the worry that this is a category mistake, conflating common and group beliefs, the meaning-inheritance does not seem as problematic as inheriting all the common knowledge. At least the group would not have beliefs about states of affairs that it did not get through its decision-making mechanisms. Although, to make this claim, one needs to subscribe to a view that beliefs about meanings are somehow distinct from other beliefs. It is also clear that, as groups can form explicit beliefs about meanings, the meaning-inheritance model should include a clause that, unless the group explicitly decides otherwise, it thinks that any concept means what common knowledge dictates.

There is also the odd consequence that a group's understanding of its own beliefs would become largely accidental. This might not be as bad as the psychological formulation of the problem of the first belief, but it leaves the group with a limited control over its own understanding. A group's beliefs about meanings can change without it actively trying to change them if the common understanding of a certain meaning changes amongst group's members. Meaning-inheritance is perhaps the most viable version of bottom-up inheritance, but it also comes with costs – the conflation of common and group beliefs, and an odd psychology – that might be too much for some to accept.

So far, it is clear that none of the presented solutions to the problem of the first belief are completely unproblematic. Despite the possible internal coherence of the positions, attributing moral agency to group agents on the grounds of thin functionalism or social externalism clashes with the existing background intuitions about the nature of moral agency. Bottom-up inheritance, in turn, has problems in showing what beliefs can be justifiably counted as group's beliefs: as things currently stand, we do not have a fully convincing story available of how the inheritance could work in such a way that the categories of shared and corporate beliefs would not be conflated.

Despite its problems, the bottom-up inheritance model comes closest to making groups morally responsible in terms that are analogous with individual moral responsibility. However, we can also cast doubt on the idea if we really need to argue for group responsibility in the exact same sense as individual responsibility. As stated by Smith (2018) it is commonplace to accept that there is a disanalogy between individual and group agents (see also Kusch 2014). And indeed, the whole purpose of discussions around group responsibility is to give an accurate description of the moral landscape and moral agents – including their similarities and differences. It is equally clear that we want to track the moral aspects of collective and shared actions – especially to "ensure that there is as much blame delivered as, on the face of it, there is blame deserved" (List and Pettit

2011, p. 167; see also Tollefsen 2015, pp. 114–115). However, what is not clear is whether this is best achieved through attributing moral agency or moral personhood to groups.

The bottom-up inheritance model pushes the analogy between individuals and groups the furthest. In comparison, thin functionalism and social externalism both rely on a looser sense of similarity between group agents and individual agents. It is also notable that thin functionalism and social externalism are not mutually exclusive options as it is possible to have functionalist description of minds within a broader holistic setting. Nevertheless, even combined, they are unable to solve the psychological formulation of the problem of the first belief. This makes group agents crucially different in comparison to individual agents, not in the least because "moral psychology" – in the sense of having self-reflective capacities and self-control – has traditionally been one of the key intuitions behind discussions on moral agency. Thus, if one is willing to hold onto groups' moral responsibility, this would require a complementary argument to show how certain psychological conditions are not central for moral agency. In other words, thin functionalism and social externalism can be taken to point either towards to impossibility of morally responsible group agents or towards the idea that we should perhaps reconsider some of the standard conditions of moral personhood.

5 In Conclusion: Collective Responsibility Reconsidered

Does the above analysis mean that group agents have no place in discussions on responsibility? This conclusion consists of a consideration of certain theoretical pathways that will aim to conceptualize groups and responsibility without a need to attribute *full* moral personhood to group agents. One way (A) is to reduce group responsibility to aggregated responsibility, while the other (B) suggests a differentiated account of moral personhood that could perhaps do justice to the disanalogy between individual and group agents.

Let us start with option (A). In the case of highly structured groups, one could argue that irreducible group responsibility is an unnecessary concept. The claim is that in the interesting cases where groups have well-defined decision-making mechanisms in place, they also have a well-defined system of tasks and rights for their members (as suggested by Tuomela 2007, p. 24; see also Tuomela 2003, pp. 221–232). These systems define and distribute responsibilities within collective entities – like corporations – in such a fashion

that there is no need to attribute an additional layer of responsibility to the group itself. The question of robust collective responsibility does not need to come up at all, as the relevant responsibilities can be attributed on the individual level, according to membership roles and responsibilities that come with them.[12]

But if this is the case, how do we make sense of the apparent discontinuity between collective decision-making and collective rationality and individual level decision-making and rationality, which is at the center of List and Pettit's argument? One plausible option is to conceptualize group-level rationality as partly externalized thinking instead of group agency-constituting collectivized rationality. In other words, collective decision-making mechanisms are merely a part of every individual's repertoire of tools for deciding what to do. This is a reductionist reading that emphasizes individuals' role as the active realizers of so-called group goals. A group would get nothing done if its members did not strive for its ends, and neither would it ever decide anything if the individuals did not agree to follow its decision-making mechanisms. Thus, on this interpretation, the ultimate responsibility for carrying out the acts that follow from group level decisions would lie at the individual level. This is not to say that moral considerations and moral reasoning would not happen in the collective context: even if groups themselves were not considered as independent moral agents, we should still require shared reasoning (or reasoning in collective contexts) to be as morally responsible as private individual reasoning. Reductionist accounts can still agree that individual (moral) reasoning happens in collaboration with others, and in relation to shared cultural background.

The "aggregated responsibility" line of thinking highlights that we need not attribute singular distinct moral agency to a group in a way that would be "discontinuous" with its individual members. After all, from the complete separateness of a group mind, it follows that a group ought to be somehow able to communicate its intentions to those individuals who do its bidding and it remains a mystery what this mechanism could be. However, to be fully plausible,

---

12 An anonymous reviewer for this journal suggested that this option does not seem quite right, because allowing corporations to attribute responsibility to individuals seems to require that corporations have beliefs. Yet this is exactly what the problem of first belief led us to deny. So we cannot allow the group to distribute responsibility. In reply to this objection, I would say that membership roles do not rely on corporate beliefs, but are instead something that individuals attribute to each other and which can be stated in the rules and organizational documents of a corporation. Therefore, there is no need to posit a group agent, which would be attributing these roles. They are individual obligations, the content of which is defined in shared practices within (and partly without) the corporation. Of course, this opens up various questions relating to power of defining these roles and obligations, which cannot be discussed here.

reductionist accounts would need to do two things, which are far too complex to be fully analyzed in the remainder of this paper. First, they would need to show how moral group agency could be reduced back to individual responsibilities. This kind of reductionism can work either by showing that groups are not agents at all – and therefore obviously not moral agents – or by allowing that groups are agents but not agents of the relevant moral kind and thereby reducing merely the moral features of groups back to individual level. The second, further, task for reductionist accounts is to analyze how group-related phenomena like coercive structures, power disparities, and domination affect individual responsibilities. These are demanding tasks but nothing that could not, in principle, be done. After all, there is no "knock down argument" for or against group agency – at least not to my knowledge – and thus it seems reasonable to keep the option of aggregated individual responsibilities open.

Now let us move to option (B). The discussion in the previous sections showed that the full attribution of moral responsibility is problematic and, perhaps, not necessary in the context of groups. However, instead of reducing all group responsibilities to individual responsibilities, there is the option to claim that groups are indeed agents but rather different agents from individual human agents. For example, Smith (2018, p. 16) claims that personhood includes a performative element – as in List and Pettit's account – but groups have different performative roles to individuals. Thus groups can be taken as persons who have different rights and responsibilities to those of individual human beings. Similarly, I have previously suggested (Hirvonen 2017b) that personhood should be understood as including different aspects and different "enabling conditions", and thus it could be possible that groups could fulfill some of these conditions. They could potentially be (moral) persons in some relevant sense but not necessarily in the same sense as individuals are.

An approach that acknowledges the possibility of different performances and partial moral personhood is perhaps more open to attributions of (partial) moral agency to groups. However, there is the lingering worry that whatever the conditions of partial moral personhood are, these conditions will nevertheless include developed moral psychology, and would require groups to be moral reasoners in themselves. If not, then this option will perhaps stretch the concept of moral responsibility far beyond its current intuitive limits (which include agency and moral reasoning). Thus the challenge of the "partial personhood" approach would be to show that these limited attributions of moral personhood are such that they can avoid the three forms of the problem of the first belief that full-fledged "decision-making and collective acceptance" accounts of group agency faced.

As a concluding afterthought, even though it is unclear whether or not moral responsibility can only be attributed to individuals, collectives can nevertheless

figure in our moral sphere as objects of moral wrongdoings. As Korsgaard (2004, pp. 95–96) states, non-rational animals may not be sources of obligations in themselves but there is still a sense in which they can obligate us. We can track harms done to animals that are not moral subjects themselves and we can track harms done, for example, to nature. Similarly, we may think of groups having some moral standing in the eyes of moral agents. In this sense, if we agree that there are objective collective entities, they can be sources of moral claims even if they would not be capable of participating in moral reasoning themselves. Examples of this could include any oppressed minorities like cultures that might be in some sense claimants for rights to retain certain cultural practices and so forth. Thus, even if the stronger claim that groups are robust responsible moral agents turns out to be untrue, groups do not lose their moral significance altogether. And even if groups could be moral agents, it may be helpful to make distinctions between groups that are morally relevant in different senses – some as subjects, some as objects, and some not at all.

References

Block, N. 1980. "Troubles with Functionalism." In *Readings in Philosophy of Psychology*, Vol. *1*, edited by N. Block, 268–305. Cambridge: Harvard University Press.

Bratman, M. 2014. *Shared Agency. A Planning Theory of Acting Together*. Oxford: Oxford University Press.

Collins, S. 2013. "Collectives' Duties and Collectivization Duties." *Australasian Journal of Philosophy* 91 (2): 231–48.

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

Dennett, D. 1976. "Conditions of Personhood." In *The Identities of Persons*, edited by A. Oksenberg Rorty, 175–96. Berkeley: University of California Press.

French, P. 1979. "The Corporation as a Moral Person." *American Philosophical Quarterly* 16 (3): 207–15.

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

Gilbert, M. 1992. *On Social Facts*. Princeton: Princeton University Press.

Gilbert, M. 2013. *Joint Commitment: How We Make the Social World*. Oxford: Oxford University Press.

Goodman, M. F. 1992. "A Sufficient Condition for Personhood." In *The Personalist Forum, Spring 1992 Supplement: Studies in Personalist Philosophy. Proceedings of the Conference on Persons*, Vol. *8*. No. 1, 75–81. University of Illinois Press.

Haksar, V. 2005 "Moral Agents." In *The Shorter Routledge Encyclopedia of Philosophy*, 692, edited by E. Craig. London, New York: Routledge.

Hindriks, F. 2008. "The Status Account of Corporate Agents." In *Concepts of Sharedness: New Essays on Collective Intentionality*, 119–44, edited by B. Schmid, K. Schulte-Ostermann, and N. Psarros. Frankfurt: Ontos Verlag.

Hirvonen, O. 2017a. “Group Personhood in the Contemporary Social Ontology.” In *Mind, Collective Agency, Norms: Essays on Social Ontology. Berichte aus der Philosophie*, 80–3, edited by P. Salis and G. Seddone. Germany: Shaker Verlag GmbH.
Hirvonen, O. 2017b. “Groups as Persons? A Suggestion for a Hegelian Turn.” *Journal of Social Ontology* 3 (2): 143–65.
Korsgaard, C. 2004. “Fellow Creatures: Kantian Ethics and Our Duties to Animals.” *Tanner Lectures on Human Values 24*: 77–110.
Kusch, M. 2014. “The Metaphysics and Politics of Corporate Personhood.” *Erkenntnis* 79 (9): 1587–600.
Laitinen, A. 2007. “Sorting Out Aspects of Personhood. Capacities, Normativity and Recognition.” *Journal of Consciousness Studies 14* (5–7): 248–70.
Laitinen, A. 2014. “Group Minds and the Problem of the First Belief.” *Balkan Journal of Philosophy* 6 (1): 43–8.
Lawford-Smith, H. 2019. *Not In Their Name. Are Citizens Culpable For Their States’ Actions?* Oxford: Oxford University Press.
List, C. 2014. “Three Kinds of Collective Attitudes.” *Erkenntnis* 79 (Supplement 9): 1601–22.
List, C., and P. Pettit. 2004. “Aggregating Sets of Judgments: Two Impossibility Results Compared.” *Synthese* 140 (1): 207–35.
List, C., and P. Pettit. 2011. *Group Agency. The Possibility, Design and Status of Corporate Agents*. Oxford: Oxford University Press.
Malle, B. 2010. “The Social and Moral Cognition of Group Agents.” *Journal of Law and Policy 19* (1): 95–136.
McBride, C. 2013. *Recognition*. Cambridge: Polity Press.
Pettit, P. 1996. *The Common Mind. An Essay on Psychology, Society and Politics*. Oxford: Oxford University Press.
Pettit, P. 1998. “Defining and Defending Social Holism.” *Philosophical Explorations* 1 (3): 169–84.
Pettit, P. 2003. “Groups with Minds of Their Own.” In *Socializing Metaphysics. The Nature of Social Reality*, edited by F. F. Schmitt, 167–93. Oxford: Rowman & Littlefield Publishers, Inc.
Pettit, P. 2007 “Responsibility Incorporated.” *Ethics* 117 (2): 171–201.
Rovane, C. 1998. *The Bounds of Agency. An Essay in Revisionary Metaphysics*. Princeton: Princeton University Press.
Schwitzgebel, E. 2015. “Belief.” In *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), edited by E. N. Zalta. https://plato.stanford.edu/archives/sum2015/entries/belief/ (accessed March 29, 2018).
Searle, J. 1995. *The Construction of Social Reality*. New York: The Free Press.
Smith, L. 2018. “The Curious Case of Ronald McDonald’s Claim to Rights: An Ontological Account of Differences in Group and Individual Person Rights.” *Journal of Social Ontology* 4 (1): 1–28.
Strawson, P. 1968. *Studies in the Philosophy of Though and Action*. Oxford: Oxford University Press.
Tollefsen, D. 2015. *Groups as Agents*. Cambridge: Polity Press.
Tuomela, R. 2003. *The Philosophy of Social Practices. A Collective Acceptance View*. Cambridge: Cambridge University Press.
Tuomela, R. 2007. *The Philosophy of Sociality. The Shared Point of View*. Oxford/New York: Oxford University Press.
Wringe, B. 2020. “Global Obligations, Collective Capacities, and ‘Ought Implies Can’.” *Philosophical Studies*, 177: 1523–38. https://doi.org/10.1007/s11098-019-01272-6.