From Ontology to Morality and From Morality to Ontology: On Collins' *Organizations as Wrongdoers*

Katherine Ritchie

University of California, Irvine

Word Count: 4,663

What, if any, role does metaphysics have to play in addressing moral questions? When answering questions about moral responsibility, many theories rely on answers to questions about the nature agency and agents, the persistence of persons, and the existence and nature of free will. In recent work in social ontology, philosophers have argued for views of social categories or identities that take ethical and social-political desiderata into consideration in several ways. Some argue for views of gender identities that depend on epistemic and ethical considerations (Bettcher, 2009). Others develop views of gender categories in ways they take to be metaphysically realist, tracking robust explanatory patterns in the world. For instance, Barnes (2017) argues for this interpretation of Haslanger's views of racial and gender categories. Others argue that injustice and oppression are reflected in the nature of social categories. For instance, while it *ought* not to be this way, some gender categories are, for instance, trans-exclusionary (Ásta, 2018). Indeed, one might even argue that social categories can involve a distinctively metaphysical form of oppression (Jenkins, 2020, 2023). It is clear that many philosophers argue in ways that draw close connections between metaphysical and moral theorizing.

In Stephanie Collins' rich and rewarding new book, *Organizations as Wrongdoers: From Ontology to Morality*, she argues that "the metaphysical and moral treatment of organizations have become detached in the literature" (2023: 3-4). Her aim is to remedy this by attaching them and demonstrating "why that attachment matters" (2023: 4). In the course of the book she develops a metaphysical account of organizations (entities like states and corporations) and the relation their members bear to them. Then she argues that an account of organizations or groups as wrongdoers ought to begin from, and be grounded in, an account of the metaphysics of groups themselves. Finally, she shows how the account yields important conclusions about the ways organizations are blameworthy and how members are implicated in organizational wrongdoing. Her arguments go well beyond familiar claims that blameworthiness requires an account of agency. This less

controversial thesis is widespread in the literature on group agency and responsibility. For instance, it is adopted by those offering functional and interpretivist views (List & Pettit, 2011; Tollefson, 2015). Rather, Collins argues for a metaphysical view of groups that is situated squarely within social ontology. She argues that this not only meets metaphysical desiderata but also has important upshots when considering questions of wrongdoing and blame. As she puts it "The metaphysical picture is an important premise in the moral arguments, while the moral arguments explain why the metaphysical picture matters" (4). Here I consider this claim in more detail by examining the specifics of the metaphysical picture Collins argues for and how she motivates. In so doing, my aim is also to reflect more broadly on philosophical methods that involve meeting both metaphysical and normative desiderata.

I begin in §1 with a brief overview of the book. Then, in §2, I offer an overview of Collins' metaphysical view of organizations and membership. I show how she motivates her account and consider several points at which a metaphysical account with different motivations could diverge. In §3, I further draw out role morality and metaphysics are playing in Collins' project and connect this to a range of approaches in the methodology of social ontology.

1. An Overview of *Organizations as Wrongdoers*

Before jumping in to the main discussion, it will be useful to have a more comprehensive outline of *Organizations as Wrongdoers*. In *Part I*, Collins develops an account of organizations as structured wholes with members as material parts. I discuss this in the next section, so set that aside for the moment. In *Part II*, she argues that organizations themselves can be blameworthy not just for actions, but also for attitudes (which might be reflected in organization goals) and character traits. To avoid an overgeneration worry, Collins argues this to require a view on which organizations can have moral self-awareness. She develops an account of the phenomenology of organizations that involves inheriting features (e.g., guilt) from organization members when they have these traits in the course of playing their role in the organization (e.g., members feel guilt qua members). Here the metaphysics of organizations and memberships clearly comes into the picture. In *Part III*, she turns to how members are implicated in an organization's wrongdoing. While she holds that organizations can be blameworthy even if no individual member is, she does not deny that members are implicated. Rather, she develops a nuanced view on which members can implicated to different degrees. Costs associated with wrongdoing ought, she argues, be distributed based on proportion of

member's implication in wrongdoing. Collins argues for what she calls a 'source-tracking' model for apportioning costs to members of an organization. On this account, costs are to be distributed to members in the proportion to which they are sources of the collective's liability.

There is much more that deserves discussion, but which I cannot discuss here. Just to touch on one cluster of questions I do not have the space to address—how might the notion of first-personal awareness as an organization member fit with ordinary first-personal awareness? Are these different mental states? What, precisely, is it for some particular person to be aware of wrongdoing *as* Australia? Might there be important connections to first-personal awareness as members of other sorts of groups, like racial or gender groups?

The account Collins develops is systematic and comprehensive—drawing together several bodies of research to develop a compelling new theory. She also begins with and returns to concrete illustrations, suggesting ways the account is not just a piece of theory, but one with practical upshots that might be used in determining blame in real world cases involving states or corporations. I agree with Collins that this really does matter. The book will be of interest to a wide range of theorists especially to those working in metaphysics (in general) and social metaphysics (in particular) and those working or interested in a broad swath of social, political, and legal questions about responsibility, agency, and, in particular, wrongdoing, blame, and reparations. I turn now to the metaphysical account Collins develops and her arguments in favor of it.

2. The Metaphysics of Organizations

Collins argues for a view of organizations as realized structures with members as their material parts. The view builds on Koslicki's Neo-Aristotelian view of structured objects (2008) and my structuralist account of groups (Ritchie, 2013, 2020). While close in spirit to an account like the one I have developed of teams and committees—Collins departs in several significant ways. First, her exemplars are organizations like states and corporations, which I have not been inclined to take to be groups in the same sense as organized groups like teams. Second, she focuses on agency and decision procedures—taking organizations to be agents, whereas the account of teams as structured wholes that I developed does not address questions of agency. Finally, given her aim of providing a metaphysical account that can be put to use to answer questions about blame and responsibility, she provides a more robust set of membership conditions that can meet this goal.

On Collins' view an organization is "a collective agent that involves a large number of people who realize a structure that coordinates divided labour via rules and hierarchical command relations, guided by a collective decision-making procedure" (2023: 9). Much of the work of *Part I* of the book is to provide a clearer understanding of how members are parts of organizations by providing desiderata for the organization-member relation and conditions for organization membership. She offers seven desiderata for an organization-member relation (paraphrased from 2023: 48-49):

> *Materiality*: an analysis should render organizations locatable in the natural world
> *Internality*: an analysis should render members internal so not everyone interacting with an organization is a member
> *Overlap*: an analysis should allow for non-identical coincident groups
> *Non-transitivity*: an analysis should not make membership transitive
> *Non-circularity*: an analysis should not be circular
> *Agency*: an analysis should require the members be agents
> *Non-conventionality*: an analysis should not just be based on convention (e.g., practice or law)

I'll return to how these desiderata might be motivated below. First, let's consider how Collins develops a view that meets them.

According to Collins members of an organization play roles (or jointly realize) the positions in a structure. Drawing on Koslicki's account, she states that "all and only the members (and the members' parts) are the material parts of the organization, while the role structure and decision-making procedure are the organization's nonmaterial parts" (2023: 52). This condition meets three of the desiderata. It meets *Materiality*, as members are material parts. It meets *Internality* since others can interact with an organization without playing a role in the organizational structure. And, it satisfies *Overlap* as structural dissimilarity is sufficient to allow for non-identical but materially coincident groups. It might be worthwhile to note that here, unlike on my account, it is not clear that Collins' view allows for distinct but coincident groups which realize the same structure. For instance, two clubs might have precisely the same structure—the same roles with the same relations between them—and, due to odd circumstances, at some point in time also share all of the same members. According to Ritchie (2013, 2020) these groups are distinct if they vary in members across times or world or bear different external relations. It seems organizations are embedded in many

other structural relations and also involve variation in members across times and worlds. So these resources might be taken up by Collins as well.

In order to meet the remaining desiderata, Collins argues for three additional conditions on membership. The *Commitment Condition* requires that members to be psychologically committed to abide by the organization's decision-making procedures when relevant to their actions (2023:54). The *Inputs Condition* which requires that members are the entities an organization takes up inputs from (e.g., beliefs, preferences) and how they are processed can be affected by behavior of these entities (e.g., through voting) even if both the inputs and the processes differ from those had or used by individual members themselves (2023: 56). And finally the *Enactment Condition* which requires that "for an entity to be a member, it must be that the enactment of some of the organization's feasible decisions requires actions from that entity, where those actions are attributable to the organization" (2023: 57).

These conditions provide the resources to meet the remaining desiderata. Since these conditions require agency (having beliefs, enacting, being committed), they meet *Agency*. They do not refer to laws, conventions or practices. In the official spelling out of each, membership is defined in terms of entities, thereby avoiding circularity. Finally, while one organization can be a part of another organization, the members of the former will not necessarily meet these conditions (e.g., relating to enactment). So the account meets *Non-transitive*. Further, while Collins does not focus on this, it is worthwhile to note that even though parts of members are material parts of organizations on her view (e.g., a member's nose is a material part of an organization), these are not members given the conditions relating to commitment, inputs, and enactment.

What motivates these conditions? Collins says that her "argument for these three conditions consists in the fact that they allow us to satisfy" the desiderata (2023: 54). But are precisely these conditions needed to meet the desiderata? It seems not. For instance, one could hold that not all members of an organization are connected to enacting actions of the organization or that organizations could have inputs from only some members. I am not suggesting these are better accounts, but they are available from a purely metaphysical standpoint. And, if they are supported by meeting the desiderata, what motivates the seven desiderata Collins sets out? It seems they are motivated from several different sources. Some relate to intuitions about metaphysics connected to particular cases.

For instance *Overlap* is motivated by intuitions that clubs can overlap completely in members but be distinct. (These intuitions have also recently been found to hold in empirical psychological work, Noyes et al., 2023). Some of the desiderata are supported by theoretical virtues, e.g., N*on-circularity*.

What about *Agency*? Recall that this desideratum is that organization should allow only agents to be members. It does not seem, at least prima facie, that there is a *metaphysical* reason to require that organizations have only agents as members. Compare, for a moment, an argument I made in previous work that social groups—like teams and committees— have only social creatures, people, or structured groups thereof, as members (Ritchie, 2020). My motivation there was that social groups—like teams and committees, will have structures with roles defined such that only creatures like humans could play them (e.g., throwing a ball, seconding a motion). That is, I took it the constraint came from the structure, given the relations that define positions in it.

Now turn to an organization like a state or a corporation. Is it the case that the structure of a state involve roles just for agents (or for just persons or social creatures or…)? Is it the case that we take the material parts—those that determine the location of the organization—to be just agential members? At least prima facie, it seems sensical to answer a question like "Where is Australia?" by describing the location of the landmass that is referred to by "Australia." We do not need to check to see where the Australian citizens currently are (perhaps spread out all around the world) in order to answer it. In fact, saying Australia is located in the US and Sweden and Namibia and China and… since there are Australian citizens in each of those places seems at least counterintuitive. Perhaps this is because of polysemy in the term "Australia". It can be used to pick out, at least, a particular area landmass as well as a state that, Collins argues, has people as its material parts. That seems correct, but I wonder then, what motivates taking the landmass that is referred to by "Australia" to not be *part of* the material realization of Australia.

Consider another case—the corporation Apple. Where is Apple located? Cupertino, California seems like a reasonable answer (particularly for those of us who own Apple products and recognize this location from the default settings). But, of course, Apple employees (who are the members and material parts on Collins' account) are located all over the world. Saying Apple is located in many places around the world also seems reasonable, but I wonder here too whether our intuitions might

not be tracking the location of employees, but the locations of factories and stores located around the world.

In contrast, suppose we ask "Where is the history club?" And, suppose further, that the club always meet in Humanities Hall Room 14. Currently the club is not meeting and the members are in various places around town. Where is the history club? It seems that Humanities Hall Room 14 is not a very good answer. These cases suggest that there is a contrast between what we (at least intuitively, perhaps in ways that are mislead by language) take to be the material parts of states and corporations, on the one hand, and clubs, on the other. So, it seems that it is not metaphysical intuitions about what counts as a part that motivates *Agency*.

To draw what I take to be the argument in favor of *Agency*, let's look at Collins' response to the worry that organizations can have material parts that are not members (or parts of members). She says, that while buildings or machines might be importantly connected to organizations they are not material parts since "when we consider organizations as *agents who can do wrong*, it seems that this more expansive view includes too many things as the parts of organizations" (2023: 58, *emphasis original*). She goes on to say that to include these "*when considering organizations as agents of wrongdoing"* would conflate *"parts* of an organization with its *property"* (ibid., *emphasis original*). I completely agree with her contention that if we are considering organizations as potential wrongdoers, their property is not likely to be relevant to attributing blame (or, as she says, only insofar as it connects to something like negligence of members of an organization). However, Collins began her investigation by saying she was offering a metaphysics of organizations and their members that would then be used to support a moral view. Here, it seems, intuitions about moral responsibility and what sorts of things can be morally responsible or can be wrongdoers is guiding the metaphysics of membership and parthood. Given metaphysical considerations alone, why not take agents and entities like buildings, machines, and parcels of land to be material parts of organizations? Then, in offering an account of wrongdoing, one could limit *which* parts are relevant to attributing responsibility, blame, and a share of costs.

A strategy like this might be needed anyway. Part of Collins's argument is that, e.g., an organization does not always count as a wrongdoer if its building causes harm. Rather, she argues, it is only if the building causes harm *and* it is connected to the organization in the right sort of way via, e.g., "intent,

recklessness, negligence" connected to the "organizations 'structure, procedures or members" (58). That seems completely correct to me. Now recall that Collins argues that parts of members are parts of an organization. It seems, however, that a part of a person—their kidney or pinky finger, say—causing harm should only entail that the organization is a wrongdoer if that part is connected up to the organization in the right way. Collins has a compelling account about how only some actions of members of an organization connect to the organization. It seems like a parallel sort of argument might be made for parts of members—which Collins takes to be parts of an organization—and to things like buildings and documents—which she does not take to be parts of an organization.

Finally I want to consider how *Inputs*. This condition partially determines who is a member of organizations like democratic states. On Collins's view some citizens can fail to be members of a democratic state, because they do not have sufficient *capacity* for inputs into the decision procedure of the state. Collins rightly notes that for some voting might be extremely difficult. From this, given her commitment to *Inputs*, she argues "voters who are systematically suppressed—who legally may register to vote, but who face large obstacles to doing so—do not count as members [of the state]. As with low-level employees, this is as it should be: these people are shielded from the individual implications of organizational wrongdoing" (2023: 77). Why think metaphysically these citizens are not members of the state? This metaphysical conclusion also strikes me as worrisome from a moral standpoint, as it seems to stand on the side of oppressive states denying membership to those who have been systematically oppressed. Rather, one might think these individual citizens are members of the state, but should not be apportioned a share in the blame of actions of the state. These metaphysical and moral claims can, and I think should, come apart.

3. Ontology, Morality, Methodology

Collins begins *Organizations as Wrongdoers* with the idea that discussions of organizational wrongdoing ought to hew more closely to metaphysics. As quoted above, she says "the metaphysical picture [developed in the book] is an important premise in the moral arguments, while the moral arguments explain why the metaphysical picture matters" (2023: 4). I agree that metaphysics matters and that it can play an important role in moral, social-political, and other projects. My question here is whether the order is as straightforward as Collins states here. Regardless of the answer to this question, drawing together work on the relationship between organizations and their members (their

metaphysics) and moral projects about organizational wrongdoing is important and has practical upshots.

As I considered above, there is reason to think that at least some of the metaphysical desiderata and conditions on membership Collins argues for are partially motivated by her aim to develop a picture that can do certain things in the moral and social-political realm. *Agency* is partially motivated by the aim to have a metaphysical picture that can neatly fit with the project of theorizing about the wrongs of organizations and their members. A ramification of *Inputs* is argued to be desirable due to its shielding oppressed individuals from being blameworthy for state actions by taking them not to be members of those states. Other arguments in Chapter 2 rely on relations like supervenience or grounding failing to provide the resources to enable a "transfer" of properties from members to organization, which Collins argues is required for a theory of organizational wrongdoing (2023: 29).

In motivating her realized structures view of organizations Collins supports it in several ways—including ways related to metaphysical explanation and meeting metaphysical desiderata like *Overlap*. She also adds that the view "can do useful work in moral and political philosophy" and "For that reason alone, it is a view worth defending" (2023: 30). This line nicely brings out the way Collins's aim is to have a clear and robust metaphysics of organizations that can also do important work in moral and political arenas. The ability to do that work is an important part of the motivation for the view.

I want to close by returning to some examples from social ontology that I opened our discussion with. Considering these in a bit more detail will help to draw out ways metaphysical and moral considerations can stand in different priority relations in theory-building.

Consider Haslanger's view of race and gender as social positional categories (2000, 2012). She offers the following as an initial first-pass definition of what it is to be a woman

> "S is a woman iff S is systematically subordinated along some dimension (economic, political, legal, social, etc.) and S is 'marked' as a target for this treatment by observed or imagined bodily features presumed to be evidence of a female's biological role in reproduction" (2000, 39).

Let's construe this as a metaphysical view of the social category *women*.[1] As Haslanger notes, since oppression is the key condition on membership in the category, non-oppressed persons (even if they are taken to be female, women, …) are not women on this definition. Haslanger notes that she does not find these cases to be troubling since her aim is to "capture a meaningful political category for critical feminist efforts" (2000: 46). Feminist efforts simply do not need to be aimed at liberating non-oppressed persons, so failing to categorize these people as women does not undermine that aim. Interpreted as a metaphysical analysis, the account could be seen as attempting to capture the (real, metaphysical) contours of structural forms of gender oppression. The metaphysics of gender categories on this sort of theorizing could differ in significant ways from how people standardly think of gender categories—in terms of their nature, extension, and so on.

Given the aims of the account there are other cases that pose a more significant challenge. Jenkins (2016) argues that Haslanger's account excludes some trans women. She notes that "Marginalizing trans women is importantly different from marginalizing nonoppressed (prima facie) cis women. This is partly because trans women are a severely oppressed group, and much of that oppression is closely tied to denials of the legitimacy of their genders" (Jenkins, 2016: 401). Feminist projects require, she argues, resources that are trans inclusive in order to address the full range of gender oppression. On this view, again understood as metaphysical, we ought to have an account of gender that includes trans women.[2]

Alternatively, one might argue that while social categories *ought* to be inclusive, due to various forms of oppression and the ways in which social practices involve categorizing people into social kinds, people can be constructed as members of categories with which they do not identify. For instance, a trans woman might be conferred the status of being a man in some context. On Ásta's view, the person is a man in that context (Ásta, 2018). But, she is careful to note, that while this is the metaphysical account she offers, it is not the way the world *should* be.

[1] The project could also be offered as a conceptual or linguistic one aimed at defining an expression or revising a concept. Some things Haslanger initially says about "asking us to use an old term in a new way" (2000: 48) sound much more like this. Her view is often taken on as a metaphysical view, including by, e.g., Barnes (2017) and Ritchie (2020). For our purposes, what matters is that the view can and has been understood as a metaphysical account of gender.
[2] Jenkins (2016) argues for a view of gender *identities* that is trans inclusive.

In all of these arguments, moral and metaphysical considerations are intertwined. Some offer a description of what are proposed to be real patterns in the world—actual social structures, the ways in which practices put in place social statues, the metaphysical conditions on the member-organization relation. Call these *metaphysics-first approaches*. Others start with morality and argue that facts about value and what people should do ought to guide metaphysical theorizing as in Chan's account of *ethics-first metaphysics* (Chan, ms). Finally, some begin with an aim to provide a metaphysical account that can be used for social justice projects or to account for organizational wrongdoing. This last strategy is partially guided by a moral aim and partially by metaphysical considerations. It seeks a metaphysical account that can *do* something in the moral or social-political realm. It can be judged, in part, on how well it meets that aim and in part on how well it satisfies constraints guiding metaphysical theorizing (e.g., theoretical virtues of simplicity, cohering with metaphysical intuitions). In this way it can't neatly be categorized as putting one element first in its approach. It involves justice and metaphysics together guiding the project. I see Collins' development of a metaphysics of organizations fitting this third approach. It is guided, in part, by metaphysical norms like capturing judgments about materiality and coincidence. The metaphysical account is also motivated by an aim to provide a view of wrongdoers. This aim explains certain features of the account (e.g., buildings aren't members of organizations), by theorizing *from* morality to ontology, rather than the other way around.

Works Cited

Ásta (2018). *Categories We Live By: The Construction of Sex, Gender, Race, and Other Social Categories*. Oxford University Press.

Barnes, Elizabeth (2017). Realism and social structure. *Philosophical Studies* 174 (10):2417-2433.

Bettcher, Talia Mae (2009). Trans Identities and First-Person Authority. In Laurie Shrage (ed.), *You've Changed: Sex Reassignment and Personal Identity*. Oxford University Press.

Chan, Rebecca (ms.). Ethics-First Metaphysics.

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

Haslanger, Sally (2000). Gender and race: (What) are they? (What) do we want them to be? *Noûs* 34 (1):31–55.

—— (2012). Resisting Reality: Social Construction and Social Critique. New York, US: Oxford University Press.

Jenkins, Katharine (2023). Ontology and Oppression: Race, Gender, and Social Reality. New York, US: OUP.

—— (2020). Ontic Injustice. Journal of the American Philosophical Association 6 (2):188-205.

—— (2016). Amelioration and Inclusion: Gender Identity and the Concept of Woman. *Ethics* 126 (2):394-421.

Koslicki, Kathrin (2008). *The structure of objects*. New York: Oxford University Press.

List, Christian & Pettit, Philip (2011). *Group agency: the possibility, design, and status of corporate agents*. New York: Oxford University Press.

Ritchie, Katherine (2020). Social Structures and the Ontology of Social Groups. *Philosophy and Phenomenological Research* 100 (2):402-424.

—— (2013). What are groups? *Philosophical Studies* 166 (2):257-272.

Tollefsen, Deborah (2015). *Groups as Agents*. Malden, MA: Polity.