



Ideal Theory in Social Ontology as Ideology

ABSTRACT: *In discussions about the possibly ideological character of ideal theory in liberal political philosophy, one worry concerns the underlying social ontology and how it can serve to make important structural injustices less visible. More recently, similar concerns have begun to appear within social ontology itself, with several authors arguing for a shift from more traditional models to different forms of nonideal or critical social ontology instead. This article develops a conception of ideal theory applicable to both social ontology and political theory, and it is then argued we should take seriously worries about ideal theory playing an ideological role. One consequence of how the ideal/nonideal distinction is understood here, however, is that the line between ideal and nonideal theory is not sharp. What we have is rather a continuum ranging from the strongly ideal to the strongly nonideal, and where the balance then needs to shift away from the former.*

KEYWORDS: Social ontology, Methodology, Ideology, Nonideal theory

In an influential argument, Charles W. Mills criticized traditional political theory for how its employment of idealizations could play an ideological function in masking existing injustices, that it is ‘in crucial respects obfuscatory’ (Mills 2005: 166). He also highlighted the social ontology underpinning mainstream liberalism as problematic. Although social ontology itself does not typically have explicit political aims, Mills’ argument raises a worry about how it could still quite possibly play an ideological role, helping to maintain oppressive structures by making them appear more natural or innocent than they really are. Mills advocated nonideal theory instead.

In recent years some social ontologists have also become attentive to similar concerns, seeking to develop ways of doing social ontology that will enable a critique of problematic practices rather than potentially covering them up. The most influential theorist is probably Haslanger (2012), but authors like Ásta (2018), Brännmark (2019), Burman (2023), Jenkins (2023), and Richardson (2023) have also explored alternative approaches to social ontology, often proposing some form of nonideal theory.¹ The present article comes to this discussion with two main objectives. One is to develop a conception of the ideal/nonideal distinction (which originally comes from political theory) that is also clearly applicable to theorizing in social ontology. The other is to consider whether Mills’

¹ I am using the term inclusively here, similar to Burman (2023). There are certainly differences between these authors, e.g., Jenkins (2023) proposes an *emancipatory social ontology* and Richardson (2023) a *critical social ontology*, but given the way the distinction between ideal and nonideal social ontology will be understood here such differences will mainly be ones of emphasis: there are many more precise ways of doing nonideal work.



worry about how ideal theory risks having problematic ideological effects carries over to social ontology. Let us however start with how Mills himself draws the distinction between two kinds of theory.

1. Mills on Ideal and Nonideal Theory

Debates about the use of idealizations have been prominent in political theory (Valentini 2012; Erman and Möller 2022), with Mills' argument being a seminal contribution. But while political theory is normative in character, it is important to note that what distinguishes ideal theory is not that it is normative—nonideal political theory is normative too. 'What distinguishes ideal theory is the reliance on idealization to the exclusion, or at least marginalization, of the actual' (Mills 2005: 168). The choice between ideal and nonideal theory is a *methodological* one (Hänel and Müller 2022). If ideal theory works to make a certain normative order seem more natural than it really is, the choice to work with idealizations will in practice bear on normative matters, but the connection would be incidental. It would be a functional role that such theorizing would play, having little to do with the intentions of the theorists themselves. If we look for a potentially ideological role for ideal theory, the primary place to look for such possible effects would presumably be in areas where theorists make assumptions about how people are related to each other in forming communities or societies, and where these can then help shape societal self-perceptions. Such assumptions are not found just in political theory. Indeed, some recent calls for turns towards nonideal epistemology (McKenna 2023) and nonideal philosophy of language (Beaver and Stanley 2023) take their cue from Mills. And in her argument for nonideal social ontology, Burman (2023) does the same. Her main construct is however rather what she calls *the standard model of ideal social ontology*. She lists five categories of features (and altogether sixteen features) characterizing it. Her account is not just based in Mills but also in Guala's (2007) conception of a standard model of social ontology, which includes as one of its tenets that collective intentionality is a basic building block of the social realm. She then takes Gilbert, Searle, and Tuomela as paradigmatic theorists for the standard model of ideal social ontology. The approach taken in the present article is different in that the focus is simply on ideal theory. This means that someone like Guala, who rejects the standard model of social ontology, can still count as doing ideal theory (and arguably should, since his approach is game-theoretical and relies heavily on idealizations). The contrast to nonideal theory becomes cleaner if one works with a distinct conception of ideal theory rather than a more complex notion that combines both methodological choices and substantive ideas about the nature of the social.²

One challenge in discussing the role of ideal theory is that it is typically not an explicitly stated methodology but rather an implicit methodological practice.

² In a similar vein, while Gilbert, Searle, and Tuomela might do ideal theory and have a focus on collective intentionality, it would seem possible to work on collective intentionality without a high level of idealization. This is another reason for keeping the ideal/nonideal distinction separate from what might be classified as standard or non-standard social ontology.

We have theorists working in ways *similar* to other theorists in using idealizations/abstractions. There are patterns that can be articulated. But given its implicit character, it seems unrealistic to expect that we can explicate much more than a family-resemblance conception of ideal theory. Indeed, what Mills (2005: 168–169) identifies is precisely like that. He points to the following typical characteristics of ideal theory: (i) *an idealized social ontology*, where the human beings who make up a society are typically portrayed as ‘abstract and undifferentiated equal atomic individuals’; (ii) *idealized capacities*, where, at least for the purpose of theorizing, individual agents are typically understood in terms of ‘completely unrealistic capacities’; (iii) *silence on oppression*, where little or nothing is said about ‘actual historic oppression and its legacy in the present, or current ongoing oppression’; (iv) *ideal social institutions*, conceptualizing fundamental social institutions ‘with little or no sense of how their actual workings may systematically disadvantage women, the poor, and racial minorities’; (v) *an idealized cognitive sphere*, where a ‘general social transparency will be presumed, with cognitive obstacles minimized as limited to biases of self-interest or the intrinsic difficulties of understanding the world, and little or no attention paid to the distinctive role of hegemonic ideologies and group-specific experience’; and (vi) *strict compliance*, where at least theorists in the Rawlsian vein, in articulating their principles of justice, presume that everyone will ‘act justly and to do his part in upholding just institutions’ (Rawls 1999: 8, cited by Mills). This is a list of features where Rawls certainly stands as the paradigmatic example of ideal theory, but it is not supposed to merely be about Rawls.

In explaining the kind of theorizing that he finds viable, Mills then relies on a distinction between *idealization* and *abstraction*, referring to O’Neill (1988). The idea is that there are two main ways in which we can simplify things in order to facilitate theorizing. We have (i) *simplification by subtraction*, i.e., by removing elements we get something that is easier to grasp, and are hopefully left with some core features of the phenomenon that we are interested in. But we also have (ii) *simplification by substitution*, where we replace factors that are difficult to handle in theorizing with simpler and more computationally manageable ones. The second form of simplification introduces distortions in a way that the first does not, and both O’Neill and Mills suggest that abstraction is therefore the better option, making theory more grounded in reality.

Mills (2005: 173) recognizes that a shift to working with abstractions does not provide any guarantees. For instance, merely *omitting* features like race, gender, and class would not be much of an improvement from working with, say, standard Rawlsian idealizations, at least not if we are interested in capturing the dynamics of those features. But one might also worry about whether the difference between idealization and abstraction really holds up as a load-bearing distinction. To begin with, idealization as such hardly precludes attention to the relevant features of our societies. One can use idealized formal models in seeking to understand the dynamics of important inequalities and injustices, which is indeed something some theorists do (e.g., O’Connor 2019; Molander 2022). Maybe such models have their problems, but it is far from clear that those problems would be about *substitutions* having been made rather than *subtractions*. Neither is it obvious that idealization necessarily

involves lower fidelity. As Cartwright (1989: 187–188) has argued, when we idealize, or reason by *substitution*, it can make sense to reason about how closely the simplification approximates the real phenomenon, since our model will still be a complete picture of how certain factors hang together, whereas in the case of *subtraction*, we are left with something that is essentially incomplete, and which need not give us a picture, not even approximately, of the dynamics between different factors in the relevant real-world phenomenon. Additionally, omitting certain things from how we model something can affect how the remaining parts are understood. A standard rational-choice model clearly involves idealizations, but even just by abstracting we can get an overly rationalistic conception of agents and their deliberations, *e.g.*, by removing biases and heuristics from the picture. Such a model would arguably still provide a skewed picture of human agency.³

Given the above, there seems to be a need for additional conceptual work in articulating a helpful distinction between ideal and nonideal theory. For the present article, this would have been needed anyway, in order to make the distinction more readily applicable to social ontology (where Rawls is hardly a paradigmatic theorist), but given these two needs it seems reasonable to aim at a tweaked conception of the distinction that (i) does not rest on the problematic distinction between idealization and abstraction, and (ii) is applicable to both social ontology and political theory (and perhaps epistemology and philosophy of language as well, although there is not room to consider this here). Let us turn to this task.

2. Ideal Theory in Social Ontology

In trying to say something about ideal theory in social ontology, it seems reasonable to first say something about just what counts as *social ontology*. In one sense, it is something that has been around for a long time. Already Plato provides a kind of social ontology in the *Republic*, and many political philosophers since have rested their arguments on ideas about the nature of things like *the state*, *society* or *institutions*. There is also an even looser sense in which basically everyone has a social ontology: we all have some vague ideas about how the communities and societies that we are part of are built up. For instance, some people take entities like *nations* very seriously, while others view such larger wholes with suspicion (even if they might never have heard the term *methodological individualism*).

The discipline now typically referred to as *social ontology* is however relatively young. Indeed, in an early entry, Tuomela (1984: ix) noted that '[i]t is somewhat surprising to find out how little serious theorizing there is in philosophy (and in social psychology as well as sociology) on the nature of social actions or joint actions'. Presumably, *serious theorizing* meant theorizing in the vein of analytic philosophy, and it could certainly (at least then) be argued that analytic philosophy had a history of focusing on individuals and individual action. Apart from Tuomela, theorists like Gilbert (1989), Bratman (1992), and Searle (1995) corrected this lack of attention, making important contributions which helped establish the field of social ontology,

³ For instance, the standard Kantian picture of human agency might be seen as overly rationalistic. Even in his later works, Mills (2018) did however see promise in a revisionary Kantianism.

as well as bringing attention to the phenomenon of *collective intentionality*. Other theorists, such as Guala and Hindriks (2015) and Hédoin (2019) have more recently drawn on game theory in order to throw light especially on the nature of institutions as solutions to collective-action problems, making contributions that would surely also count as *serious theorizing* by Tuomela's standards.

It is this more particular sense of *social ontology* that will be in focus here, although the wider and looser sense is ultimately also relevant for the argument. The idea is to articulate a conception of ideal theory applicable to theorizing in the field of social ontology, but to do so in a way that retains its applicability to Mills' own object of analysis as well: mainstream liberal political theory. If we worry about ideal theory playing an ideological role, that worry will be greater if social ontology and political theory can be understood as possibly working together, supporting the same kind of looser everyday understanding of our societies which we have as members of them, an understanding that can play an obfuscatory role in relation to certain injustices. If social ontology plays an ideological role, the hypothesis is that it does so in alignment with political theory. In building on Mills' view but trying to broaden the applicability, here are four main features that can be said to characterize ideal theory.⁴

(1) *Transcendentism (about inquiry)*. This first feature is a backgrounding one. It is about a *lack* of reflection on how one as a theorist is always located somewhere. Philosophers rarely, nowadays at least, explicitly pronounce that they are taking *the view from nowhere*, or *the point of view of the universe*, or looking at things *sub specie aeternitatis*. Yet many texts arguably read *as if* this is what their authors are taking themselves as doing, *as if* their authors think themselves capable of transcending their attachments, allegiances, biases, and vested interests. The classic Rawlsian approach of reasoning in terms of a veil of ignorance and the original position clearly builds on a faith in our capacity for such transcendence. Mills (2005: 175) distances himself from transcendentism, pointing out that among the kinds of theorists (Marxists, feminists, critical race theorists) with whom he is aligned,

[t]he crucial common claim—whether couched in terms of ideology and fetishism, or androcentrism, or white normativity—is that all theorizing, both moral and nonmoral, takes place in an intellectual realm dominated by concepts, assumptions, norms, values, and framing perspectives that reflect the experience and group interests of the privileged group (whether the bourgeoisie, or men, or whites).

Philosophy is a subject very much constituted by an ongoing dialogue with what has come before. There are few (if any) other subjects where historical figures are as alive

⁴ In drawing on Mills (for the ideal theory part of *the standard model of ideal social ontology*), Burman (2023: 14n9) chooses to 'downplay or exclude two features—idealized capacities and strict compliance—from the characterization of ideal social ontology.' This is because she finds them largely missing in social ontology. In the present article, the idea is instead to articulate a unified conception of ideal theory applicable to *both* social ontology and political theory (and possibly more).

and present in contemporary discussions. Philosophical inquiry often also takes its starting-point in various tenets, positions, and arguments that are widely accepted or at least taken seriously within the discipline. It is a dialogical way of proceeding that crucially depends not just on cold hard logic but also on what is already regarded as *relevant* or *important* within the tradition. Additionally, many arguments involve appeals to intuitions or the reader's sense of what is reasonable – again, aspects that go beyond cold hard logic and which certainly can be influenced by one's place, identity, and accumulated experiences. Positionality matters. What is here called *transcendentism* involves a lack of reflection about positionality. Certainly, in some cases such reflection might have taken place without ending up in the relevant texts as something visible. But for a lot of work in the analytic tradition, it is difficult to shake the impression that this is not the case.

If we look at social ontology, then apart from the above kind of general transcendentism there is potentially also a more specific type of idea in play, about the relation between *social ontology* and *the social sciences*. Both aim at better understanding the social realm, or at least parts of it. How do they relate to each other? Different theorists might take different stances here, but to the extent that this relation is commented on by social ontologists, it often seems that they take what they are doing to be such that it does not require any greater engagement with the social sciences. One example here is Bratman (2022: xx):

[M]y aim is not to chart in detail the many ways in which different kinds of organized institutions function. We can expect much variation in such details, variation that is the concern of social scientific investigation. My aim is rather to articulate an abstractly specified infrastructure that is common to and important for a broad range of human organized institutions, despite wide variability across those institutions.

Alternatively, one might take the stance that social ontology is prior to the methodological and theoretical choices made in the social sciences, one example here being Searle (2009: 9):

I believe that where the social sciences are concerned, social ontology is prior to methodology and theory. It is prior in the sense that unless you have a clear conception of the nature of the phenomena you are investigating, you are unlikely to develop the right methodology and the right theoretical apparatus for conducting the investigation.

Such comments at least signal an idea about there being a rarefied altitude at which philosophical inquiries into social phenomena take place. In a sense they do involve a certain level of self-reflection, but it is not a critical one: rather, it is one that justifies not engaging with the messy details of the social sciences. Still, while one can find passages like these, the main argument here is about just how difficult it is to find the relevant kind of critical self-reflection outside of works in social ontology that fall on the nonideal side of things—with Jenkins (2023) being a good example of such reflection.

(2) *Atomism (about agents)*. A main criticism of Rawlsian political philosophy has had to do with its conception of persons (or selves or agents). In the so-called *liberal-communitarian* debate of the 1980s, Sandel (1984) characterized this atomistic conception of the person as an *unencumbered self*. In Rawls' theorizing, the fundamental choice faced by people concerned the very conditions under which they are supposed to enter into voluntary association with others, *i.e.*, a choice faced by a kind of pre-social or at least pre-societal selves. While this kind of approach can provide explanations of how separate individuals can come to be joined in a social arrangement, it can also be seen as creating the very problem it then solves by first removing the actual glue in our social relations. To reason about what people would choose in something like the original position, we need to abstract from the messiness of their concrete and specific roles and commitments. But actual selves are always already social, where various social roles and commitments define who people are, and partly what they want as well. Even in cases where we actually need to coordinate with strangers, our attempts at doing so will still start from pre-existing scripts and schemas that we have been brought up with and socialized into so that they are second nature to us. What we do now, and what we consider as viable options to begin with, will be grounded in our histories of previous interactions.

To what extent do social ontologists operate with a similarly atomistic picture? It is difficult to generalize here. Social ontology is more varied in character than Rawlsian political theory, but two important clusters are (i) theorizing in terms of collective intentionality and (ii) approaches using game theory. If we start with the latter, game theory is an obvious case of working with a highly idealized conception of the agent. Since game-theoretical approaches hinge on *us as theorists* being able to determine what will maximize utility for the relevant agents, this means that we must assume a very simplified picture of those agents and their aims, and where these agents also strictly operate according to certain principles (axioms of rational choice). We need to regard preferences as fixed, and we need to severely limit the available actions, typically to a binary choice. Of course, it is a further question whether these simplifications are worth doing because of how they facilitate theorizing, but at least it seems clear that a strongly atomistic picture is being presumed.

As for theorizing in terms of collective intentionality, this approach was for a long time dominant within social ontology, with influential authors like Tuomela, Gilbert, Searle, and Bratman developing different accounts. While there are many important differences between their approaches, they can arguably all be said to deliver a form of *how-possibly explanations* of how we can get from individual agents to things like social groups, institutions, and societies—offering logical or rational reconstructions where one starts with certain basic building blocks and then shows how more complex phenomena can be built from these. For instance, Searle (2010: 6–7) proposes that '[t]he enormous complexities of human society are different surface manifestations of an underlying commonality' and that he is 'in search of a single mechanism' for creating institutional facts. Similarly, in describing his theoretical project, Bratman (2022: xvii) declares that he seeks to 'contribute to our understanding of organized institutions by *constructing* them', using the

interpersonally shared intentions that he has developed a theory for ‘as building blocks’. To get the needed generic building blocks one must abstract from the concrete and specific features always already characterizing actual people as the socially embedded individuals that they are. This theoretical engineering project of working with building blocks and showing how complex social wholes can be constructed from these is made workable by starting from an atomism about agents, by assuming unencumbered selves.

(3) *Isolationism (about situations)*. In the contractualist tradition which forms the background to Rawlsian political theory, a key figure of thought is *the state of nature*: a situation where a collective of individuals stands without society and without institutions and faces a problem of how to coordinate or cooperate. Of course, this type of modeling will presuppose the kind of unencumbered selves that we have already discussed, but there is also another important feature to it: the strongly *demarcated* character of the situations being considered. Again, Rawls (1999: 7) provides a clear example, when stating that he is working with ‘the basic structure of society conceived for the time being as a closed system isolated from other societies’, suggesting that ‘once we have a sound theory for this case, the remaining problems of justice will prove more tractable in the light of it.’ Part of the messiness of everyday social life is that different situations overlap with each other and spill over into each other. The state of nature, or the original position for that matter, is demarcated in two main ways. One concerns ignoring the existence of an ‘outside’ potentially impinging on what happens in the situation under consideration, as well as how it can be difficult to draw a line between ‘inside’ and ‘outside.’ These are complications from which ideal theorists abstract away. The other demarcation is temporal. The state of nature has no history, it just *is*. And yet actual institutions and social structures always have a history, often being strongly path-dependent in how they evolve.⁵

A tendency to theorize using sharply demarcated situations can also be found in much of social ontology, perhaps most clearly in what kinds of examples are used in developing theories and explanations of social phenomena. Theorists of collective intentionality often work with simple, small-scale, one-off situations. For instance, in describing the kind of examples he focuses on, Bratman (1992: 327) notes how ‘[y]ou and I might sing a duet together, paint a house together, take a trip together, build something together, or run a give-and-go together in a basketball game.’ And drawing on remarks in Simmel about temporary social interactions, Gilbert (2006: 97–98) suggests that ‘we can explore the nature of social groups in general by investigating such small-scale temporary phenomena as those of two people going for a walk together, dancing together, working on a project together, and so on.’

⁵ For instance, when political parties emerged, the precise ways in which this happened in different countries were shaped by the institutions already in place. If one had single-member districts and a plurality voting system (as was the case in the US), the continued path naturally led to a two-party rather than a multi-party system—and then these parties naturally became invested in the preconditions of that two-party system. The point here is not that processes like this are arational, but rather just that we need historical detail in order to understand why specific institutions end up the way they do, and that the timing of when changes occur can be crucial for how such processes unfold (for more on this, see Lipset and Rokkan (1967) on what they call *critical junctures*).

By focusing on small-scale cases we get a demarcation from larger societal contexts, and by focusing on temporary phenomena we get a temporal demarcation. These demarcations can certainly help in facilitating theorizing, but at least if we accept that a lot of social life is really *not* that well demarcated, they also amount to a risky methodological bet.

If we turn to game-theoretically oriented social ontology, the first theorist to develop such an account of institutions, Schotter, suggested (1981: 20) that we ‘start our analyses in a Lockean state of nature in which there are no social institutions at all, only agents, their preferences, and the technology they have at their disposal to transform inputs into outputs.’ More recently, Hindriks and Guala (2021: 2029) have explained how ‘[t]he point of departure of game-theoretic accounts of institutions is a game form that specifies the preferences that agents have and the actions or strategies that are open to them. Within a particular game, agents maximize their utility by taking into account how others are likely to behave.’ Again, there is no inside or outside to how such situations tend to be conceived. The agents being considered are the only relevant ones. There is no history that explains why certain options are being considered rather than others. Possible workings of power that place the relevant players in their respective situations are also typically left out of the picture. There is just a strictly limited number of available moves and possible outcomes.

(4) *Voluntarism (about explanations)*: In traditional contract theory, once we have construed the agents and the situation which they face, the key question that we ask is whether they would, knowing all the relevant facts, come to an agreement on certain principles. And typically, the agents and the situations they are in have been set up so that the answer is that it would be rational or reasonable for these people to join together in a certain scheme of cooperation. Of course, similar to other types of idealized modeling (such as in economics) the idea is then that we have thereby thrown light on actual societies, even if these are obviously messier and more complex. This kind of approach lets us identify the conditions under which agents would come together in a spontaneously voluntary way, *i.e.*, without there already being an authority in place that coordinates their actions by telling them what to do, or without there already being social practices in place which can be extended or reinterpreted to address new issues that appear. The idea seems to be that when we have shown how an agreement (or a harmonizing of attitudes and behaviors) could come about under such circumstances, an important explanatory task has been completed.

In political theory, contractualist thought from Hobbes to Rawls has very clearly had this conception of the explanatory task in mind. In social ontology, voluntarism does not revolve around the idea of an explicit contract, but it is still there. Collective-intentionality theorists typically describe how we can move from being mere individuals to a collective by how we bind ourselves to each other through commitments. For example, Tuomela (2007: 13) highlights the centrality of *collective commitment*, ‘that is part and parcel of the we-mode can be regarded as a central “glue” in social life. It concerns the group members’ collectively binding themselves, for example, to an idea, action, or to the group itself.’ Gilbert instead emphasizes the notion of *joint commitment*, arguing that it ‘both results from and creates something plausibly

thought of as a real unity of the parties: they co-create a commitment of them all that ties each to all in a web of rights and obligations' (Gilbert 2018: 225). This is a picture of social glue as coming into existence through clear-eyed acts of committing ourselves to certain social arrangements. While actual social life can often be murky or opaque, ideal theory assumes a transparency of (possible) social arrangements in order to get theorizing going.

In game-theoretical accounts, the whole point is of course to identify equilibrium solutions, and these do not require any deeper coming together as a 'we'—agents stick to the equilibrium because it maximizes their own utility. However, even when the games being modeled involve larger groups, potential heterogeneity tends to be left behind, with the population typically split into just two camps where all members in each camp have the same beliefs and preferences. A strong level of intra-group agreement or harmony is accordingly being presumed in order to facilitate modeling in terms of what will essentially be standard two-player games, like hawk-dove (e.g., Guala and Hindriks 2015). Transparency is also typically presupposed in such modeling, with players knowingly adopting certain strategies in light of all the relevant facts. Such theorizing can certainly provide insights into some types of social dynamics, but will by necessity have to keep many aspects of the social realm out of the picture, e.g., how pre-existing power relations often influence the ways we approach various problems or how we might have been socialized into seeing different alternatives as viable or not. And again, in actual social life things can often be murky or opaque, with people struggling to see what is really at stake.

So: we now have what is hopefully a workable conception of *ideal theory*. It is somewhat different from Mills' own conception, but should serve as a tightened-up account developed from his. It is not as Rawls-focused, but as should be clear it still classifies someone like Rawls as an ideal theorist. When it comes to the four features, (i) *transcendentism* captures Mills' emphasis on reflexivity and overlaps with his item of an *idealized cognitive sphere*, (ii) *atomism* is something he explicitly refers to in characterizing his item of *idealized social ontology*, (iii) *isolationism* should capture much of what he is after in his item of *ideal social institutions* (which is about a narrowness in focus), and (iv) *voluntarism* connects to his items of *idealized capacities* (in presuming knowledge and transparency) as well as *strict compliance* (in working with an idea of everyone being onboard with an arrangement). In being less Rawls-centric, the proposed account should be more general than Mills' and accordingly more readily applicable to other areas. Rather than building *silence on oppression* explicitly into the characterization, the proposed account is intended to be explanatory in this regard: helping us understand why certain approaches tend to end up being silent on oppression.

Now, while our focus here will be on the possible ideological function that ideal theory can play, it should be clear from the above discussion that one might in any case also worry about the descriptive adequacy of ideal theory. Such concerns have occasionally already been raised in the literature. For instance, in discussing Bratman, Shapiro (2014: 258) notes that 'theorists have largely eschewed giving analyses of activities involving authority structures. They have confined themselves to egalitarian activities such as the aforementioned duet singing and house painting, where neither participant has normative power over the other.' And Kutz (2000: 11)

worries about ending up with an ‘account unsuited to the depersonalized, hierarchic, bureaucratic, but nonetheless collective institutions that characterize modern life.’⁶ We will not explore such doubts about descriptive adequacy here, but their existence is worth keeping in mind. In the cost-benefit analysis that we perform when making theory choices, the descriptive benefits of ideal theory should not be taken for granted—even apart from worries about ideology.

3. Ideal Theory as Ideology

When it comes to the problematic role (potentially) played by ideal theory, Mills (2005: 172) makes the following key claim:

Ideal theory, I would contend, is really an *ideology*, a distortional complex of ideas, values, norms, and beliefs that reflects the nonrepresentative interests and experiences of a small minority of the national population—middle-to-upper-class white males—who are hugely *over-represented* in the professional philosophical population.

That professional philosophy has had and still has a problem with representativeness is hardly disputable. The more contentious claim is that ideal theory plays a *distortional* role (which then also benefits the group whose interests and experiences are being reflected). It should however be said that Mills is far from unique in making this kind of point. For instance, similar arguments have been made by Raymond Geuss as well. He finds the main fault with ideal theory to be that it abstracts from the workings of power, which is in line with Mills’ view, with my characterization above, and with Burman’s (2023) analysis of the standard model of ideal social ontology, which also highlights power as a crucially missing component. Geuss (2008: 52–53) then understands ideology in terms of how a particular configuration of power ‘brings it about that certain contingent, variable features of our human mode of existence (which are in fact maintained in existence only by the constant exercise of that power) appear to be universal, “natural,” or necessary or spontaneously arising features.’

A theory playing an ideological role is mainly about which ideas are *present* in the theory, and that certain ideas explicitly or implicitly make the current order of things appear as natural or reasonable in a way that legitimizes it. But it can also be about which ideas are *absent*: how the theory does *not* provide adequate conceptual tools for articulating the political concerns of subordinated, oppressed or marginalized groups. As such, absences are not harmful, but if theories built on certain absences dominate a discussion, they will risk institutionally crowding out possible alternative approaches that could have served as the basis of a more systematic critique of the current order. Alternative ideas become marginalized. Of course, effects like these are difficult to trace empirically since they concern very subtle mechanisms, where big-

⁶ Burman (2023: Chapter 2) also raises similar worries, and Ritchie (2020: 100) argues that ‘the methodological assumption of starting with small group one-off interactions and scaling up to large group long-term interaction should be abandoned.’

picture ways of seeing the world might seep into general awareness and then do much of their work in the background. Certain ways of seeing the world can also already be established, and the philosophical underpinnings that are then produced would be more like an affirming reflection in the mirror. It does however not seem outlandish to think that even such reflections could potentially play a reinforcing role, leading to a strengthened sense of the relevant ways of seeing the world being warranted, also among people who do not read philosophy. Even people just having a vague sense of their everyday thinking ultimately being backed up by serious inquiry could be impactful.

Are these kinds of *possible* effects too nebulous to be taken seriously? No. The best way to think about them is arguably in terms of *risk* and *intellectual responsibility*. In most of life, we often take precautions even when we are not certain about some ill effects necessarily following from our actions. If we can identify plausible mechanisms through which ill effects might happen, then there is ground for taking them seriously. And especially if we as social ontologists hope for and tend to highlight possible positive real-world effects, *e.g.*, that our theorizing about collective agency and responsibility will help humanity address collective-action problems, it seems fitting to also consider possible ill effects. Typically, those who put forward ideology worries do not think that people consciously produce ideology, but in the face of ideology worries it would certainly be a conscious choice if one dismisses these concerns without seriously reflecting on them. The suggestion here is accordingly that it would be intellectually irresponsible to just brush aside questions about how the social positions, experiences, and interests of participants in academic theory production within social ontology might influence and distort theorizing, and hence in the long run help legitimize unjust societal arrangements. It would be irresponsible not to consider how social ontology does not just stand outside of society, coolly observing and dissecting, but rather potentially holds up a trick mirror to current society in a way that risks obscuring its problematic aspects.

Since the argument here is framed in terms of *risk*, it might very well be the case that on closer inspection these worries about our theories playing an ideological role are overblown. Risks can be so insignificant that they are not worth the effort of avoiding. But that would be a conclusion, not a starting-point. The responsible thing initially is to take the relevant risks seriously. To this point, and tied to the four features of ideal theory identified above, here are some possible mechanisms that are causes for concern:

(1) *Effects of transcendentalism*. Transcendentalism is not about the particular substance of one's theory, so it is primarily problematic in how it involves not taking certain risks seriously. In worrying about ideological effects of dominant models in social ontology, we are worrying about whether social ontology potentially contributes to making highly contingent social arrangements seem more natural or reasonable than they really are. If someone is then doing a certain form of social ontology as if it simply is *the* obvious approach to things, rather than being *a* product of what really are non-obvious methodological choices, then surely that would only exacerbate the problem. Choosing to model something in a certain way will inevitably sharpen the contours of certain aspects of the social realm more than others, and at the very least it seems reasonable to demand reflection on what

possibly gets lost or rendered less visible in doing so, and how one's own perspective and position can play a role in understanding the explanatory tasks to be addressed. This is not to say that personal self-reflection automatically solves things. We are all prone to wishful thinking, and there is always also the risk that even critical reflection grows stale, so that one more or less mechanically points out certain obvious issues and then moves on. It would also be difficult to get anywhere if every single paper had to start by asking fundamental questions about what one is doing. Still: if something you are doing has problematic effects and you do not reflect on how that might be so, it is very likely that you will just keep on doing things in the same bad way.

(2) *Effects of atomism.* Which actual groups that are privileged or oppressed will be the product of concrete historical processes, and will be entwined with specific social positions and identities. If we consistently abstract from things like people having genders, being racialized, or occupying class positions, we risk implicitly rendering these features secondary concerns—even though for our everyday agency they might take primacy in how we understand our choices and navigate the societies in which we live. This is a type of concern that Mills (2005: 168) strongly emphasizes:

An idealized social ontology (...) will abstract *away* from relations of structural domination, exploitation, coercion, and oppression, which in reality, of course, will profoundly shape the ontology of those same individuals, locating them in superior and inferior positions in social hierarchies of various kinds.

When abstracting from our prior entanglements, then apart from the thinness of such accounts, the relevant preferences are typically also treated as *exogenous*, i.e., they are simply assumed to be there and are not part of the explananda that our theorizing is supposed to make sense of. And yet the problems that we face in the real world might have to do not just with coordinating in order to get what we want, there might also be problems with *what we want*—what we have been socialized into desiring. Ideal theory renders such matters less visible. To the extent that this kind of picture seeps into general awareness, it could accordingly support a practice of not taking such concerns seriously. Additionally, working in that way could also restrict us in developing relevant critical tools.

(3) *Effects of isolationism.* A main characteristic of isolationism is to set the historical context of existing practices and institutions aside. It also involves bracketing possible wider implications of behavior in the situations under consideration. One can certainly be an isolationist and consider more complex situations in steps and one can also bake certain external factors into how one describes what is at stake in a situation. Still: ultimately one needs a closed system. And perhaps this will sometimes be fine. But it is another matter to consistently settle for analyses of social matters that zoom in on and isolate a few features as the relevant ones, rather than taking in various background factors, wider implications, and historical contexts. Certain matters then risk being consistently pushed to the margins. Obvious examples here are inequalities and injustices related to gender and race. If we look at institutions simply as rules that facilitate coordination, the institutions that we have now will presumably be assessed as largely egalitarian. But

what effects do such rules have in the wild, so to speak, when they are interacting with various other factors that we abstracted away from in our theorizing? Working with a narrow and thin description of society, we risk underpinning an analysis of current social issues where unequal outcomes are understood not in terms of deeper structural problems but in terms of certain groups, such as women and racialized persons, not making use of the opportunities that are openly available to them. By helping to make messy and more subtle problems less visible, the thinness of an isolationist approach risks doing real damage to the prospects of addressing actual social ills.

(4) *Effects of voluntarism.* In trying to account for how social arrangements can come into place *ex nihilo* we arguably set ourselves a harder explanatory task, compared to starting from an idea of human beings as always already social. While there is nothing wrong in taking on a hard explanatory task, the worry here is that a persistent focus on this type of explanation could help create or maintain an impression that social arrangements, at least when not being overtly oppressive, are typically based in consent or that they are at root cooperative or mutually beneficial (and hence: legitimate). Indeed, this tendency sometimes even comes as programmatic pronouncements, *e.g.*, when Tuomela (2013: 229) suggests that institutions tend to solve or dissolve ‘conflicts between collective and individual rationality—and give cooperative, collectively beneficial solutions to these problems in the face of the chaos and conflict that unfettered individual action tends to lead to.’ If one works with this kind of Hobbesian contrast, then we can basically expect any set of institutions to be counted as an improvement. For our actual societies, however, the choice is rather between what we have now and something somewhat different—not about what life without any institutions would look like.⁷ This is not to say that voluntarist explanations, with their focus on the problem-solving character of institutions, are without value. They can potentially highlight dynamics of which we should be aware. But if this type of theorizing is the dominant model, we risk becoming less sensitive to how institutions can also sometimes be oppressive (or typically: beneficial for some, but oppressive for others).

We have now identified some *possible* mechanisms through which ideal theory might work to reinforce a problematic *status quo* by casting a light that makes the practices and institutions that are already in place seem more natural, reasonable, and egalitarian than they really are—as well as casting problematic aspects of our societies in shadow, making them less visible. We are not talking about any necessary consequences here, so saying that none of this follows from ideal theory would be an irrelevant response. The point is that our models and theories *risk* causing skewed perceptions and interpretations about how our societies are organized and function, and that this kind of risk matters for how we should theorize in social ontology.

⁷ Especially in political theory there is a larger discussion to be had about how to make comparative assessments of different alternatives in terms of whether these are just or not. For an argument that such assessments can (and should) be made without having hypothetical and idealized constructions as our measuring rod, see Sen (2009).

4. Moving Towards Nonideal Theory

When Rawls (1999: 215–217) introduced his distinction between ideal and nonideal theory, he conceived of them as two distinct but complementary approaches—they each have their own tasks and we need both of them.⁸ However, many recent contributors to this debate in political theory often see the two as competing approaches and argue for abandoning ideal theory in favor of nonideal theory, what Erman and Möller (2022) call *the strong position*. The above argument leans more in this latter direction, but to the extent that there is reason to take the relevant ideological mechanisms seriously, they primarily have to do with balance between ideal and nonideal theorizing. The issue would be with ideal theory as a *dominant approach* in social ontology (especially if it goes hand in hand with ideal theory as a dominant approach in political theory). What is needed is arguably then not a complete abandonment of ideal theory, but rather a shift of balance within the field towards (much) more nonideal theory. One thing to emphasize then is that nonideal theory can come in many different forms. This is partly because it is primarily negatively defined, but also because there are four features and they all admit of degrees. What we face is accordingly something more like a multi-dimensional continuum. Theorists who clearly fall on the nonideal side of things will reject all four -isms identified in section 2. In contrast, theorists who clearly fall on the ideal side of the continuum might make occasional concessions to how the actual social realm is messy and complicated, but will still largely proceed at least *as if* they accept the four -isms. The map is however much more complex than what you would get from a simple dichotomy. In order to give at least somewhat more structure to this map of the theoretical landscape, let us briefly look at the key characteristics of nonideal theory in social ontology that we get just from the contrast with ideal theory.

(1) *Rejecting transcendentalism*. Few philosophers have transcendentalism as an explicit methodological stance. The challenge is rather to not work *as if* one accepts it, especially in philosophy, since its high level of abstraction sets it apart from other disciplines that might study similar phenomena. Social ontology in particular might even seem to rest on being able to transcend at least the social sciences. Yet even if we have to move upwards in abstraction in doing philosophical work, we can still reflect on *how* it is done. Rejecting transcendentalism is about embracing *positional reflexivity* as imperative. While different nonideal theorists might differ in terms of how this is done and the extent to which it is made visible in their texts, reflection on the role of social ontology within larger society and with respect to social change will be a key element in nonideal social ontology. How far should this be taken? One possibility is to simply opt for *immanent critique* (Diehl 2022), which relies on normative standards from within the society or community under evaluation. One potential worry, however, is that this leaves us with a degree of relativization that could even threaten the political projects that many nonideal theorists find important. For instance, Haslanger (2021: 55) cautions against simply

⁸ On the whole, Rawls obviously prioritizes ideal theory, but he occasionally also goes into nonideal theory; for a good overview, see Simmons (2010).

relying on immanent critique, arguing that we know certain moral truths, and that ‘[t]hese are not truths we learn from theory; theorizing is guided by these truths.’ Some of the moral judgments that she considers, *e.g.*, that ‘[t]he Atlantic slave trade was wrong. Nazi genocide was wrong’, are surely ones that all her readers will agree with. But we also know that these are sentiments not shared by literally everyone, even today. In this insistence on a form of Archimedean point, one could possibly see certain trace elements of transcendentalism remaining in Haslanger’s account. But maybe that is even how it should be, and at any rate it comes after careful methodological consideration of the available options.

(2) *Rejecting atomism.* One thing that should be noted is that *atomism* (as used here) is overlapping with but not the same as *methodological individualism*. Some theorists in the nonideal camp, *e.g.*, Haslanger (2022), strongly reject methodological individualism (and certainly then also atomism). But others, like Ásta (2023) and Brännmark (2021), seem to accept at least some kind of individualism as a working assumption. Atomism involves modeling the individual agent as an independently functioning unit, and then looking at what happens when such a unit meets other such units under circumstances which make coordination, cooperation or conflict possible. In contrast, a non-atomistic individualist will understand individuals as always already embedded in existing social arrangements and relations. Social structures will play an important role in such approaches, it is just that they will be understood in terms of systematic patterns in how individuals relate to each other. Examples and modeling will accordingly feature not just rational agents with a couple of preferences between them, but richer descriptions of existing interrelations between agents and the power structures within which their capacities for agency are embedded. Of course, compared to the complexity of actual individuals and their relations, such descriptions will still be relatively thin in character, so we are mainly talking about differences in degrees between levels of abstraction.

(3) *Rejecting isolationism.* In a way, isolationism is a form of atomism, but with respect to situations. Of course, no-one believes that in actual life, we simply go through a series of well-demarcated situations where it is completely clear what is at stake, as well as who is part of this or that situation and who is not. Rather, there are unclear boundaries, both over time and in the moment, and agents are typically part of both smaller and larger social contexts, with a host of different factors influencing their behavior. Theorists isolate situations in order to facilitate theorizing. And just as the rejection of atomism does not mean giving up on abstraction altogether, nonideal theorists cannot avoid simplifying and purifying the cases they use in order to drive their theorizing. But what can be avoided is to work with cases where agents are located in some version of a state of nature. Rather than isolating features and considering them in turn, nonideal theorizing will tend to look at situations where several different dimensions are simultaneously in play, and where there are connections to what has come before and what happens or can happen in other situations. This might then throw light on social features that cannot be properly captured by an isolationist approach. For instance, in her analysis of different types of power, Burman (2023: 215–219) identifies *spillover power* as something that can empower an agent in certain situations but which rests on what holds in other situations.

(4) *Rejecting voluntarism.* Voluntarism presupposes that we can, from the philosophical armchair, determine what agents would choose under certain circumstances. This exercise hinges on having clear rules for how those choices happen, and atomism and isolationism enable that. These features are accordingly mutually supporting. Once we admit more messiness into the picture and consider why people want the things they do, explanatory models focused on how individuals *would* join together in various ways *given* that they want certain things that are obviously incomplete. In contrast, nonideal theorists will tend to start from how human agents are always already social. The goal is then rather to make sense of more concrete social behaviors by showing how actions can be reasonable for agents in the light of the social structures in which they are embedded. This will mean that nonideal theorists, in looking at specific cases or examples, will typically take more of an *interpretative* approach, where relevant concepts and theories will be used in order to throw light on those cases, similar to how empirical researchers might use qualitative analysis in order to make sense of actual situations, events, and processes. One task that fits well with the project of doing nonideal theory is to provide explanatory narratives that serve as alternatives to the standard stories typically being told (Haslanger 2015: 9–12).

While I hope that these brief comments provide some substance to what is involved in practicing nonideal theory in social ontology, it is still an approach mainly characterized by what is rejected rather than what is embraced. Accordingly, there can be many different ways of doing nonideal theory, and while there is already some nonideal work being done in social ontology, we should be careful about treating particular authors as paradigmatic exemplars. Doing so risks foreclosing other viable approaches, and ultimately the argument here points towards expanding the kind of work that is done: towards more pluralistic theorizing. Within the (very) old discipline of philosophy, social ontology is a surprisingly young subdiscipline, one where it should not be surprising if much remains to be explored.

JOHAN BRÄNNMARK

DEPARTMENT OF PHILOSOPHY, STOCKHOLM UNIVERSITY

johan.brannmark@philosophy.su.se

References

- Ásta (2018) *Categories We Live By*. Oxford University Press.
- Ásta (2023) 'Categories We Live By: Reply to Alcoff, Butler, and Roth'. *European Journal of Philosophy*, 31, 310–318.
- Beaver, D. & Stanley, J. (2023) *The Politics of Language*. Princeton: Princeton University Press.
- Bratman, M. (1992) 'Shared cooperative activity'. *The Philosophical Review*, 101, 327–341.
- Bratman, M. (2022) *Shared and Institutional Agency*. New York: Oxford University Press.
- Brännmark, J. (2019) 'Institutions, ideology, and nonideal social ontology'. *Philosophy of the Social Sciences*, 49, 137–159.
- Brännmark, J. (2021) 'Patriarchy as institutional'. *Journal of Social Ontology*, 7, 233–254.
- Burman, Å. (2023) *Nonideal Social Ontology*. New York: Oxford University Press.
- Cartwright, N. (1989) *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press.

- Diehl, S. (2022) 'Why immanent critique?' *European Journal of Philosophy*, 30, 676–692.
- Ermann, E. & Möller, N. (2022) 'Is ideal theory useless for nonideal theory?' *The Journal of Politics*, 84, 525–540.
- Geuss, R. (2008) *Philosophy and Real Politics*. Princeton: Princeton University Press.
- Gilbert, M. (1989) *On Social Facts*. Princeton: Princeton University Press.
- Gilbert, M. (2006) *A Theory of Political Obligation*. New York: Oxford University Press.
- Gilbert, M. (2018) *Rights and Demands*. New York: Oxford University Press.
- Guala, F. (2007) 'The philosophy of social science: Metaphysical and empirical'. *Philosophy Compass*, 2, 954–980.
- Guala, F. and Hindriks, F. (2015) 'A unified social ontology'. *Philosophical Quarterly*, 65, 177–201.
- Haslanger, S. (2012) *Resisting Reality*. New York: Oxford University Press.
- Haslanger, S. (2015) 'Social structure, narrative and explanation'. *Canadian Journal of Philosophy*, 45, 1–15.
- Haslanger, S. (2021) 'Political epistemology and social critique'. *Oxford Studies in Political Philosophy*, 7, 23–65.
- Haslanger, S. (2022) 'Failures of methodological individualism: The materiality of social systems.' *Journal of Social Philosophy*, 53, 512–534.
- Hédoin, C. (2019) 'Institutions, rule-following and conditional reasoning'. *Journal of Institutional Economics*, 15, 1–25.
- Hindriks, F. and Guala, F. (2021) 'The functions of institutions: Etiology and teleology'. *Synthese*, 198, 2027–2043.
- Hänel, H. and Müller, J. (2022) 'Non-ideal philosophy as methodology'. *Theoria: A Journal of Social and Political Theory*, 69, 32–59.
- Jenkins, K. (2023) *Ontology and Oppression*. New York: Oxford University Press.
- Kutz, C. (2000) *Complicity*. New York: Cambridge University Press.
- Lipset, S. M. and Rokkan, S. (1967) 'Cleavage structures, party systems and voter alignments: An introduction'. In S. M. Lipset and S. Rokkan (eds.), *Party Systems and Voter Alignments* (New York: Free Press), pp. 1–64.
- McKenna, R. (2023) *Non-Ideal Epistemology*. New York: Oxford University Press.
- Mills, C. W. (2005) '"Ideal theory" as ideology'. *Hypatia*, 20, 165–183.
- Mills, C. W. (2018) 'Black radical Kantianism'. *Res Philosophica*, 95, 1–33.
- Molander, P. (2022) *The Origins of Inequality*. Cham: Springer.
- O'Connor, C. (2019) *The Origins of Unfairness*. New York: Oxford University Press.
- O'Neill, O. (1988) 'Ethical reasoning and ideological pluralism'. *Ethics*, 98, 705–722.
- Rawls, J. (1999) *A Theory of Justice, rev ed.* Cambridge, Mass.: Harvard University Press.
- Richardson, K. (2023) 'Critical social ontology'. *Synthese*, 201, 1–19.
- Ritchie, K. (2020) 'Minimal cooperation and group roles'. In A. Fiebich (ed.), *Minimal Cooperation and Shared Agency* (Cham: Springer), pp. 93–109.
- Sandel, M. J. (1984) 'The procedural republic and the unencumbered self'. *Political Theory*, 12, 81–96.
- Schotter, A. (1981) *The Economic Theory of Social Institutions*. Cambridge: Cambridge University Press.
- Searle, J. (1995) *The Construction of Social Reality*. New York: Free Press.
- Searle, J. (2009) Language and social ontology. In C. Mantzavinos (ed.), *Philosophy of the Social Sciences* (Cambridge: Cambridge University Press), pp. 9–27.
- Searle, J. (2010) *Making the Social World*. New York: Oxford University Press.
- Sen, A. (2009) *The Idea of Justice*. Cambridge, Mass.: Harvard University Press.
- Shapiro, S. (2014) 'Massively shared agency'. In M. Vargas & G. Yaffe (eds.), *Rational and Social Agency* (New York: Oxford University Press), pp. 257–293.
- Simmons, A. J. (2010) 'Ideal and nonideal theory'. *Philosophy & Public Affairs*, 38, 5–36.
- Tuomela, R. (1984) *A Theory of Social Action*. Dordrecht: D. Reidel.
- Tuomela, R. (2007) *The Philosophy of Sociality*. New York: Oxford University Press.
- Tuomela, R. (2013) *Social Ontology*. New York: Oxford University Press.
- Valentini, L. (2012) 'Ideal vs. non-ideal theory: A conceptual map'. *Philosophy Compass*, 7, 654–664.