

Real Kinds in Real Time: On Responsible Social Modeling

Abstract. There is broad agreement among social researchers and social ontologists that the project of dividing humans into social kinds should be guided by at least two methodological commitments. First, a commitment to what best serves moral and political interests, and second, a commitment to describing accurately the causal structures of social reality. However, researchers have not sufficiently analyzed how these two commitments interact and constrain one another. In the absence of that analysis, several confusions have set in, threatening to undermine shared goals for the responsible modeling of social kinds of humans. This essay first explains the source and substance of these confusions. Then, by distinguishing different value-laden investigative questions into the classification of social kinds of humans, it sets out specific relations of dependence and constraint between empirically-driven investigations and value-driven investigations into social kinds of humans. The result is a more detailed and fruitful framework for thinking about the classification of social kinds that respects both normative interests and mind-independent causal regularities.

§1: Introduction

In the mid 2000's, as the "war on terror" ramped up, the U.S. military came to rely heavily on the use of unmanned aerial vehicles (UAVs) for both surveillance and combat missions. UAVs are controlled remotely by human drone operators. Drone operators must go through extensive specialized training. When they are ready to operate the weaponized drones, they go to work in air-conditioned ground-control stations, typically thousands of miles away from the theatre of warfare. It is a very stressful job. Empirical studies reveal that drone operators have high rates of emotional exhaustion, fatigue, stress, cynicism, and burnout (Hijazi et al. 2017). Drone operators also experience severe psychological and social "whiplash" effects that result from toggling daily between civilian and combat roles (Hijazi et al. 2017).

Drone operators are a real social kind of human. Specific social institutions, for example military training regimens and unique occupational environments, causally explain why drone operators tend to resemble one another in a variety of distinctive ways. Failing to identify and accurately describe this kind would be to know less about the world.

But who decides the precise borders, identity conditions, and characteristic properties for this kind? What about other *prima facie* social kinds of human like *women*, *men*, *millennial*, and *analytic philosopher*? Are the relevant borders, identity conditions, and characteristic properties for these groups etched into the mind-independent world, there to be discovered? Or rather, do social theorists possess a type of epistemic or moral authority that allows them to *impose* or *stipulate* these borders, conditions, and properties?

Providing a framework for how best to think about these questions is a central project of social ontology and social theory generally. Among those working on this project, there is growing consensus that the organization of social reality into social kinds of humans should be guided by background moral and political values. There is, in addition, broad agreement that the project should not be carried out in abstraction from mind-independent causal realities. Simplifying greatly: researchers generally agree that it is misguided to ignore real kinds of people and their characteristic profiles (the latter consensus), but they also generally agree that the causal constraints of social reality afford a measure of taxonomic freedom in which background values ought to steer taxonomic choice (the former consensus). We can call the embrace of these two methodological commitments the project of “responsibly modeling social kinds of humans.”

Unfortunately, the details for how these two methodological commitments are supposed to fit and interact with one another remain obscure. This obscurity has allowed a specific set of confusions to set in, largely unnoticed, potentially undercutting the responsible modeling of social kinds of humans.

The first confusion concerns the extent to which the *a posteriori* study of mind-independent causal realities must constrain value-laden claims about what social kinds of humans there are. As I explain below, social ontologists like Mari Mikkola and Sally Haslanger have overplayed the amount of stipulative freedom that social researchers have when classifying social kinds of humans. I will motivate a principle according to which, regardless of the background moral and political interests that you wish to serve, it is never acceptable to overlook or misidentify real social kinds of humans.

A second confusion is the failure to distinguish the different ways that *time* interacts with background values in the context of responsibly modeling social kinds of humans. In brief, researchers have not adequately distinguished between: (a) normative investigation into what

social kinds of humans *currently exist*; (b) normative investigation into how contemporary social kinds of humans should *change*; and (c) normative investigation into what *new* social kinds of humans should be brought into (or out of) existence. As I will explain, there are different degrees of stipulative freedom associated with each type of investigation, and the failure to separate these investigations has led social theorists to talk past one another as well as neglect important relations of normative constraint between the investigations.

A third confusion, which results from the prior two, concerns how social kinds of humans ought to be framed politically and linguistically. As I will argue, even if political recommendations compel us to use classifications that do not correspond to actual social kinds of humans as revealed through *a posteriori* investigation, these recommendations can never function as an argument against either the reality or moral importance of social kinds of humans as discovered and described through *a posteriori* investigation. In fact, political recommendations for non-kind corresponding classifications are, or at least ought to be, always premised on the causal profiles of actual social kinds of humans and the institutional mechanisms that ground them.

The goal of this article is thus to provide a richer and more fruitful analysis than is currently available for the specific ways that value-laden and empirical investigations into social kinds of humans interact and constrain one another. In §2, I examine why and in what respects this type of analysis is currently absent. In §3, I explain why value-laden investigations into social kinds are strongly dependent on empirical investigation into real social kinds of humans. §4 argues that we need to distinguish five different normative questions about the investigation of social kinds of humans. These questions have different diachronic features, affording different degrees of value-directed “stipulative freedom.” §5 concludes.

§2: The currently obscure relationship between empirical and value-driven investigations into social kinds of humans

A popular way to capture the value-laden methodological commitment sketched above is in reference to Haslanger’s “ameliorative” strategy.¹ This strategy asks us to elucidate categories

¹ In earlier papers Haslanger calls this the “analytic” strategy.

that will serve or advance specific moral and political interests, for example categories that help address the ills of racism and bigotry. As Haslanger explains, the ameliorative strategy confers a degree of stipulative authority to justice-minded theorists who order the social world:

The questions “What is gender?” or “What is race?” require us to consider what work we want these concepts to do for us; why do we need them at all? The responsibility is ours to define them for our purposes....neither ordinary usage nor empirical investigation is overriding, for there is a stipulative element to the project: *this* is the phenomenon we need to be thinking about....on this approach, the world by itself can’t tell us what gender is, or what race is; it is up to us to decide what in the world, if anything, they are.
(Haslanger 2012, p. 224).

On this view, competing social taxonomies need to demonstrate that they can pull their normative weight by advancing specific social or political goals. Haslanger is well known for advancing hierarchical classifications for gender and race categories on the grounds that these classifications purportedly best serve feminist and anti-racist values.

But how are ameliorative stipulations supposed to fit with the other methodological commitment sketched in §1, namely, the requirement to do justice to the mind-independent causal regularities of social reality? The following passage from Catharine Mackinnon, referenced and seemingly endorsed in Haslanger (2012, p. 84), underscores the importance of this methodological commitment:

Epistemologically speaking, women know the male world is out there because it hits them in the face. No matter how they think about it, try to think it out of existence or into a different shape, it remains independently real, keeps forcing them into certain molds. No matter what they think or do, they cannot get out of it. It has all the indeterminacy of a bridge abutment hit at sixty miles per hour. (MacKinnon 1989, 123)

Ian Hacking, discussing the moral importance of accurately describing causal regularities that bear on social kinds of humans, makes a related observation:

We want the right classification – the correct sorting of child abuse or teen-age pregnancy – so that confronted by abusive parents or pregnant teenagers we can embark on a course of action that will change them for the better and will prevent others from joining their ranks.... We want principles according to which we can interfere, intervene, help, and improve. (Hacking 1995, p. 360–361).

Interfering, intervening, helping, improving – these are all causal actions directed at individuals and institutions who are themselves bound up in specific, socially mediated causal patterns. The only way that these (or any) actions can be successful is if the models on which they are premised accurately identify and track causal reality. Whether the goal is fixing patriarchy or fixing a dishwasher, one’s instructional manual needs to correspond to the causal reality of the current system.

So: how does this constraint fit into the practice of ameliorative stipulation? Based on discussions provided by supporters of ameliorative stipulation, the answer seems to be: weakly and vaguely. Proponents of the ameliorative strategy typically address the empirical constraint by appealing to Elizabeth Anderson’s (1995) essay “Knowledge, Human Interests, and Objectivity in Feminist Epistemology” – see especially Mikkola (Mikkola 2015, 2016a, 2016b) and Haslanger (2012). The relevant idea from Anderson’s essay can be summarized as follows. Begin with the claim that (social) reality is rich and variable. Countless possible carvings present themselves, so there is no one way to carve it up. Next, claim that this plentitude affords the (social) taxonomist freedom when theorizing about what kinds or types (of people) there are. Specifically, the (social) taxonomist has a choice of which organizing principle – for example which framing investigative question – to use. This principle or framing question will highlight some particular region, depth, or organization of (social) reality as significant. Finally, claim that one’s choice of organizing principle or framing question, and thus one’s choice of which region/ depth/ organization of (social) reality to bring attention to, should be informed by carefully vetted value-commitments (e.g., the values of anti-racism and feminism).

I do not disagree with any of the above. But I still remain puzzled about the actual extent to which ameliorative stipulations are supposed to interact with, or be constrained by, empirical investigation into causal regularities. Anderson (1995) claims that a theoretical organization must be “empirically adequate” in order to be “politically adequate.” But for Anderson – and

especially Mikkola's and Haslanger's application of that view – what this amounts to is that theoretical organizations must be limited to true claims. In other words, when imposing value-directed theoretical organizations on the richness of social reality, we should not make things up.

The upshot is that Mikkola's and Haslanger's appeal to Anderson (1995) leaves us with a rather weak and vague empirical constraint on ameliorative stipulation, which is to avoid fictions. Because there is now a cottage industry of social researchers who appeal to Haslanger's ameliorative strategy to justify their own use of value-laden classifications, this is an important but neglected dialectical point; researchers generally take it for granted that the empirical-normative balance of the ameliorative strategy has been worked out, but closer inspection of that balance and its roots in Anderson (1995) raise important questions about the precise constraining role of mind-independent causal regularities as discovered through empirical investigation.

In other places, Haslanger mentions the importance of “descriptive projects” for social classifications. According to Haslanger, descriptive projects involve “relying on empirical or quasi-empirical methods. Paradigm descriptive projects occur in studying natural phenomena” (Haslanger 2012, p. 223). Haslanger goes so far as to claim that the ameliorative strategy will “depend crucially on descriptive efforts” (Haslanger 2012, p. 353, fn. 22). This would seem to give empirical investigations a greater constraining role regarding ameliorative stipulations – at least greater than the prohibition on making things up as reviewed above. Unfortunately, Haslanger (2012) provides little direction concerning what these “crucial dependencies” consist in or how they actually work in tandem with ameliorative goals. (It is perhaps telling that Haslanger's remark here about crucial dependencies – one of the few places where Haslanger explicitly addresses the relationship between ameliorative and descriptive projects – is relegated to a footnote).

Finally, there is significant confusion over whether the ameliorative strategy is supposed to apply to concepts, parts of the world, or both. Do ameliorative stipulations pertain to what we want our social concepts and terms to *mean*? Or is the goal to stipulate something about the worldly phenomena to which our concepts and terms might refer (i.e., borders, membership criteria, characteristic properties, etc.)? Is there even a difference? Haslanger is aware of the confusion:

I have obscured this in some of my published work by distinguishing different projects by reference to different sort of concepts we are investigating....and different ways of investigating concepts (conceptual, descriptive, ameliorative) This is confusing. [One] source [of] confusion is that I am an externalist about concepts and so I believe that in order to understand the content of a concept—what a given concept is a concept of—one must investigate the world, not the mind. For an externalist, [providing a theory of our concept of race (or water), and providing a theory of race (or water)] involve investigation into the (watery, racial) parts of the world.... (Haslanger 2014, p. 111; see also Haslanger 2012, p. 429 for a similar concession).

I want to put aside for the moment questions about reference, semantic externalism, and linguistic meaning. I will return to these below, in §4. The point to emphasize here is that this concept-world ambiguity internal to the ameliorative strategy only inflames the obscurity with which we have been dealing, namely, how value-driven classifications of social kinds of humans fit with or depend on primarily empirical investigations into social kinds of humans.

§3: The political and moral requirement to identify, track, and accurately describe real social kinds of humans

The foregoing shows that we are not on solid ground employing the ameliorative strategy unless we can first understand how ameliorative stipulations are supposed to “depend crucially” on descriptive efforts. The point is not that the ameliorative strategy is not worth pursuing – far from it. The point is that the political and moral interests that justify that strategy in the first place cannot be adequately served unless these questions about dependency and constraint are first sorted out.

However, there is a significant obstacle for getting clear on these dependencies. To see this, we need to return to Mikkola’s and Haslanger’s appeal to Anderson’s epistemological framework. The problem is the way that Mikkola and Haslanger, in order to make room for ameliorative stipulation, frame the mind-independent world of facts. Consider the following passages:

Good theories are systematic bodies of knowledge that select from *the mass of truths* those that address our broader cognitive and practical interests. (Haslanger 2012, p. 226; emphasis added)

Given time and inclination, I could tell you many truths – some trivial, some interesting, many boring – about my physical surroundings. But *a random collection of facts* does not make a theory; they are a *disorganized jumble*. (ibid. p. 226; emphasis added)

On my view, the world is replete with things, and parts, fusions, sets, collections, and properties of things. Differences abound. There are more divisions than we could ever note or care about. (ibid. p. 188)

Contra *the bare accumulation of truths*, however, Anderson notes that theoretical inquiry ‘aims at some organized body of truths that can lay claim to significance’—theoretical inquiry aims at elucidating an organised body of significant truths. If the goals of our inquiry go beyond *the mere accumulation of truths* though, we will have ‘multiple grounds for criticizing, justifying, and choosing theories besides truth’. The door is thus opened for moral, social and political values to enter the context of justification because now theory choices must be made relative to a broader set of aims: what counts as significant? (Mikkola, 2015, pp. 783-784; emphasis added)

The problem is that these and similar passages encourage an equivocation between two senses of “disorganized jumbles.” On the one hand, there are the disorganized jumbles that epistemic agents might delineate or conceptually arrange. Any gerrymandered list or randomly collected set is an example. Given the availability of relational properties and an intent to gerrymander, it is always possible to produce such random collections e.g., producing a list of the first nouns from the first clues from each Jeopardy! television program that aired on a Wednesday. No one should doubt the possibility of there being “disorganized jumbles” in this sense.

On the other hand, there are disorganized jumbles that might exist mind-independently. These might be the states of affairs that correspond to a gerrymandered list, for example the set of actual objects that are the referents of the gerrymandered Jeopardy! list mentioned above. If

one has a liberal view about the ontology of sets, then perhaps such disorganized jumbles exist in some limited mind-independent, metaphysical sense.

But crucially, there is an important difference between it always being possible for epistemic agents to produce disorganized jumbles (the former sense), on the one hand, and the mind-independent world consisting only in disorganized jumbles (the latter sense), on the other. In his *An Essay Concerning Human Understanding*, Locke offers a nice description of what this latter scenario, involving ubiquitous mind-independent jumbles and the saturation of logical space, would amount to:

in all the visible corporeal world, we see no chasms or gaps. All quite down from us the descent is by easy steps, and a continued series of things, that in each remove differ very little one from the other. There are fishes that have wings, and are not strangers to the airy region: and there are some birds that are inhabitants of the water, whose blood is cold as fishes, and their flesh so like in taste that the scrupulous are allowed them on fish-days. There are animals so near of kin both to birds and beasts that they are in the middle between both: amphibious animals link the terrestrial and aquatic together; seals live at land and sea, and porpoises have the warm blood and entrails of a hog; not to mention what is confidently reported of mermaids, or sea-men. we shall find everywhere that the several species are linked together, and differ but in almost insensible degrees (Locke, *Essay*, III. vi. 12)

What is helpful about invoking Locke's description here is that we know that it is dead wrong. As Millikan (2010) vividly explains:

As a matter of empirical fact, logical space is almost entirely empty. Huge gaps exist separating tiny clumps or clots of actual individuals or carving out miniature humps, peaks, or ridges on the surfaces of these clumps. There are, for example, well over two million separate extant species of animals, but they do not generally fade into each other, let alone into shoe horns and alarm clocks. (Millikan 2010, p. 58)

And why are there humming birds and Kodiak bears but not, say, shoe horn-coyote hybrids? Millikan continues:

The clotting is a contingent matter of empirical, not logical, fact.... [there are] clotting mechanisms that determine certain individuals within specific space-time regions to be caused, each closely to resemble the next, the restriction on displayed property mixes depending also on historical contingency. (ibid.)

One set of “clotting mechanisms” are the natural laws that make H₂O but not HO₂ a stable molecule. Another set of clotting mechanisms are the biological, ecological, and evolutionary laws that explain why *Canis familiaris* (but not *Canis familiaris-lava lamp hybrids*) are a stable species, where studying one member will inform you about unencountered members.

As Millikan (1999), Root (2000), Mallon (2002), Bach (2012), Epstein (2014), and several others have made clear, there are also clotting mechanisms in the social world. It is in virtue of these mechanisms – mostly social institutions that wield significant causal influence – that specific sets of social properties are currently and mind-independently clustered together (as opposed to being disorganized jumbles).² These property-mixes or property-organizations – call them “real kinds” (or even “natural kinds,” if you are willing to look past misleading connotations about intrinsicity and necessity) – obtain in the social world, even if they do not exhibit the same degree of stability and uniformity as, say, microstructural or biological real kinds. Importantly, the relevant clotting mechanisms and the real social kinds of humans that they fashion are subject to *a posteriori* discovery and investigation.

Consider again the example of drone operators. Drone operators share a variety of characteristic properties, and for the same reason. The reason is the clotting mechanisms to which they are each related, which include shared historical conditioning regimen and occupational environment. In contrast, there are currently no clotting mechanisms that systematically sustain a real kind “drone operator-Inuit hybrids.” The point here is not that Inuit

² There are interesting differences among the above theorists with respect to their precise construals of the “clotting mechanisms” (Millikan), “anchoring schema” (Epstein), “homeostatic clustering” (Boyd, Mallon), or “replicative machinery” (Bach) that produce real social kinds. Below I focus on the replication-based explanations of clotting from Millikan (1999, 2017) and Bach (2012), but the arguments to follow are compatible with other accounts. I also think that Haslanger’s more recent work on “social structures” has shifted in the direction of a more full-throated realism that more closely aligns with these authors, but that is a topic for another occasion.

cannot be drone-operators. The point is that there are currently no social institutions causally molding individuals to reliably exemplify social properties that are characteristic of being both a drone operator and an Inuit. Such are the yawning gaps and chasms in the social world, situated between real social kinds to which social scientists direct investigative efforts. Returning to the mandate from Hacking and Mackinnon to intervene effectively, there is a moral priority to understand these real kinds, via *a posteriori* investigation, for the kinds that they are.

Social science done well identifies these clustered social properties and, ideally, provides causal explanations for the clustering (the clotting mechanisms). But it is also useful to have general theories of the clotting mechanisms, as these can provide unificatory power and inform social scientists what to look for. Millikan emphasizes various levels of copying – institutional, social, psychological – to explain why members of a real historical kind non-accidentally share properties. One set of copying mechanisms is the institutional environments that mold individuals so that they are similar (e.g., drone pilots). Another set is differential forms of socialization and associated conditioning methods. In Bach (2012), I argued that we should construe what Boyd and Richerson term “prestige bias” and “conformist bias” – which is (roughly) people’s tendency to copy individuals who are common or prestigious – as important replication-based clotting mechanisms. On this view, institutional forces help determine which traits are common and which are prestigious (in the case of gender, these forces would include sex segregation and withholding power and status from gender non-conforming individuals), and then prestige-based and conformist-based copying mechanisms further explain the clustering of specific social properties (e.g., the clustering of properties that are characteristic of historical women and men).

It is worth repeating here that *to assert the mind-independent reality of real social kinds is not to deny that mind-dependent processes and institutions play a significant or even exhaustive role in the creation and maintenance of such kinds*. Human minds, with all of their contingent and often malformed interests, are of course responsible for the social processes and institutions that causally explain the development and on-going maintenance of real social kinds of humans. Nonetheless: it is *not* “up to us” what these causally efficacious processes currently *are* or *have been*. Nor is it up to us what their resulting human kinds currently are and have been. These points should not be confused with things that *are* to a greater degree up to us, for example: what these processes and their resulting human kinds might *become*, what they might

be *replaced* with, how we want to *label* real social kinds or other groupings, and how we want to *politically represent* real social kinds of humans (these different types of investigation will be the topic of §4). For example, in the case of socially gendered kinds of humans, mind-driven contingent processes of differential socialization have played a central role in the causal production and maintenance of the kinds men and women. These processes could have been other than they are, and it is in our power to change or replace these processes in the future. Nonetheless: what these processes have been, and what are currently the social kinds that they have fashioned, are just as much a part of the empirical world, there to be discovered and better understood, as are chemical kinds.

Let's take stock. In their attempt to motivate the justificatory role of background values in the classification of human kinds, we have seen how Mikkola and Haslanger overplay the formless character of the empirical world: it is true that there are an infinite number of potential groupings (gerrymandered or otherwise) that social reality makes available to would-be carvers, but it is not true that mind-independent social reality is a “disorganized jumble” or a “random collection of facts.” On the contrary, there are very specific and very real organizations of facts or properties that have non-randomly developed – what I am calling “social kinds of humans” – and these kinds have characteristic properties and causal-explanatory features independent of social theorists and the values of those theorists. Such real kinds are there to be discovered or neglected, described accurately or inaccurately.

I started this section with an attempt to clarify how value-driven/ameliorative investigations might “crucially depend” on primarily empirical/descriptive investigations into social categories. With the investigative importance of real social kinds of humans now established, I can assert the following dependency, put in the form of a principle for responsible social modeling, between ameliorative and descriptive investigations:

Principle for the Moral-Political Importance of Real Kinds (PRK): For any set of moral and political interests (*s*) that require intervening with or changing causal features of social domain (*d*), satisfying (*s*) will require identifying, tracking, and accurately describing mind-independent social kinds of humans in (*d*).

The basic idea here is that, if one's political and moral goals include predicting, explaining, or changing the social world, then it is never morally or politically advantageous to know less about the world and its real kinds and particularly its real social kinds of humans.

Consider opioid addicts, which are a real social kind of human in this sense. There are contingent but causally real social institutions and processes – medical and insurance companies that over-prescribe pain medication, economic institutions that frustrate job and lifestyle expectations for would-be members of the middle class, social expectancies surrounding gender role, and so forth – that causally mediate a set of characteristic properties and especially opioid-directed behavior for individuals who have been sufficiently exposed to these causal forces. If we want to intervene through social and legal policy in effective and morally responsible ways, then we need to defer to what we learn through *a posteriori* investigation into the characteristic properties and clotting mechanisms of this kind. If we do not see this mind-independent organization for what it is – or if we instead defer to armchair ameliorative stipulations about what it is – then we undermine our ability to interfere, change, and communicate successfully about the kind.

PRK is further supported by reference to the *interconnectedness* of social environments and how this impacts normative interests. For example, there is mounting evidence that masculine norms (e.g., norms against admitting and openly discussing internal suffering) have played a role in the current opioid epidemic.³ Very likely then, bringing morally positive change to the social kind “opioid addict” will in part rely on the accurate empirical investigation, and subsequent intervention with, distinct social kinds of humans like “men.” The point generalizes. But these types of connections are difficult to foresee from an ameliorative armchair, whether one is stipulating classifications in the interests of drug recovery or feminism.⁴ One needs to study real kinds and discover their interactions so that one can plot how to intervene effectively.

³ See, e.g., Bilsker et al. (2018). See also British Columbia Coroner's Service (2018) “Illicit Drug Overdose Deaths in BC January 1, 2008 – February 28, 2018” (<https://www2.gov.bc.ca/assets/gov/public-safety-and-emergency-services/death-investigation/statistical/illicit-drug.pdf>)

⁴ The opioid epidemic is clearly not in the interest of feminist values – but how is a definition of gender like Haslanger's, stipulated to consist strictly in hierarchal relationships, equipped to address this? The definitions seem to obstruct us from drawing empirically correct and politically important connections that describe the often surprising interrelationships between social kinds of humans. I interpret Mills (2014) as raising a related concern

While PRK does not explicitly rule out the possibility that imprecise or even inaccurate descriptions of real social kinds of humans could meet political ends, it does restrict both the venue and grounding for such descriptions. For example, to manipulate social meanings (Lessig 1995) in order to promote politically desirable social change, there may be cases in which public policy level construals of social kinds of humans lack empirical precision but nonetheless serve vetted political interests (perhaps if such descriptions interact favorably with well-known cognitive biases). But even here, in order for the non-kind corresponding classification to achieve its political goal non-accidentally, it must be premised on, or calculated in reference to, our most empirically accurate models of contemporary social kinds of humans, their clotting mechanisms, and what it takes for these mechanisms and associated kinds to change. In other words, the instrumental value of a political framing will be dependent on the accuracy of the empirical models of the mind-independent kinds and clotting mechanisms that one is seeking to manipulate. Thus PRK's requirement to identify, track, and describe real social kinds of humans accurately still applies. In §4, I explore in greater detail the different ways, including the above, that this requirement functions relative to different value-laden investigative questions.

This point also helps situate PRK with a previous critique of ameliorative analysis. Both Saul (2006) and Mikkola (2016b) claim that Haslanger's ameliorative analysis of race and gender in terms of hierarchical concepts is unlikely to be effective in the fight against gender and race-based injustice. Addressing Haslanger's prediction that her hierarchical concepts will cause a positive shift in peoples' gender self-conceptions, Saul observes:

But there are many uncertain empirical matters. Haslanger takes it that the shift would be a positive one, causing us to question our identities in a productive manner that furthers the cause of social justice. But would this happen?Whether [this] would occur is a matter of human psychology. Gambling on the positive responses Haslanger expects is risky. (Saul 2006, p. 138)

about the epistemic priority of ameliorative claims relative to empirical investigation in the context of race. As Mills puts it, important questions about race "cannot be settled by conceptual fiat. So our categories should not beg the question in favor of a particular interpretation's being true in advance of an examination of the historical evidence" (2014, p. 96).

Mikkola (2016b) similarly expresses skepticism that Haslanger's categories will affect "ordinary social agents" (p. 88) in the way that Haslanger suggests.

PRK clarifies, strengthens, and generalizes the empirical component of this line of objection. It clarifies and strengthens that component by making clear that the "empirical matters" that bear centrally on whether a given ameliorative stipulation will have its intended outcome are the real social kinds of humans and associated clotting mechanisms described earlier. Such real kinds are more focused and richly organized than the generic empirical categories "human psychology," "ordinary social agents," and "ordinary speakers" referenced by Saul and Mikkola. It is a central tenet of the explanatory approach to natural kinds that maximally predictive and explanatory categories are ones that latch onto the actual causes of homeostasis or clotting for the properties one seeks to predict or explain.⁵ Thus if social researchers eschew empirical models of real social kinds of humans in favor of models that target generic or superordinate categories, then they leave predictive and explanatory power on the table. Per PRK, to do this is to impede the satisfaction of intervention-based normative interests. Returning to Haslanger's projected shift in self-conception, we improve the accuracy of our predictions about that shift – and we can better determine whether Haslanger's hierarchical categories are the ideal instruments for causing it – by consulting our most accurate empirical models of the specific *kinds* of social agents and psychologies currently fashioned by historical gender systems, e.g., empirical models for historical men and women. Finally, PRK generalizes this empirical constraint on ameliorative stipulation by applying it to all instances of social modeling.

§4: Different investigative questions – and different degrees of stipulative freedom – for responsible social modeling

PRK indicates a specific and important dependency of ameliorative classification on empirical investigation. However, some justice-minded social theorists will be reluctant to embrace PRK, especially if they believe that PRK upholds the social status quo as a normative

⁵ See especially Griffiths (1997) and Boyd (1999). Both authors also discuss how this tenet applies to the explanatory and predictive practices of the biological and social sciences.

ideal and further cements harmful forms of marginalization.⁶ I submit that this type of concern rests on a confusion. The confusion results from a failure to distinguish between the different temporal aspects of real social kinds of humans and how these different aspects impact different value-laden investigative questions. Specifically, it is crucial that we separate at least five value-laden investigative questions in the context of social modeling. These are questions about:

1. which social kinds of humans *currently* and mind-independently exist
2. how contemporary social kinds of humans ought to *change*
3. which *new* social kinds of human ought to be brought into (or out of) existence
4. which *political construals* we should adopt for the classification of social kinds of humans
5. which *political labels* we should adopt for the classification of social kinds of humans.

Failure to separate these five ameliorative questions has led social researchers to talk past one another and draw unwarranted conclusions. This section disentangles and explores the five questions. It also clarifies the application of PRK. PRK informs us that satisfying intervention-based normative interests will require the accurate empirical modeling of real social kinds of humans, but it does not specify exactly how that requirement is met in the context of different value-laden investigative contexts. This section provides that missing information. As we will see, for some value-laden investigations into human kinds there is very little “ameliorative freedom” for the appeal to background values, and for other types of value-laden investigations there is substantially more freedom.

PRK makes relevant at least five value-laden investigative questions (see fig. 1 below). The first three questions pertain to real social kinds of humans, and they are each distinguished by a different diachronic feature. I discuss each question in turn.

Ameliorative Question 1 (AQ1): On the Investigation of Historical and Contemporary Social Kinds of Humans: What social kinds of humans historically and currently exist, the accurate tracking of which is important for serving current moral and political interests?

⁶ See Mikkola (2016b), especially §4.4.2. See also the “normativity problem” as discussed in Haslanger (2000).

Figure 1. Relationships of dependence and normative constraint between value-laden and empirical investigations into social kinds of humans. Row *D* indicates whether answering the target Ameliorative Question must take into account the outcomes of distinct Ameliorative Questions (the “+” indicates when this dependence is particularly strong).

		1	2	3	4	5
Topic	A	Historical and Contemporary Social Kinds of Humans	Changing Social Kinds of Humans	New Social Kinds of Humans	Political Construals of Social Kinds of Humans	Labeling of Social Kinds of Humans
Ameliorative Question	B	<i>What social kinds of humans historically and currently exist, the accurate tracking of which is important for serving current moral and political interests?</i>	<i>If contemporary social kinds of humans can survive change, how should our best moral and political theories direct change to the underlying anchoring or clotting mechanisms that fashion kinds?</i>	<i>Given current moral and political interests, what new social kinds of humans, and thus what underlying anchoring or clotting mechanisms, do we want to bring into (or out of) existence?</i>	<i>Given current moral and political interests, how should we politically frame or represent historical, contemporary, projected, and new social kinds of humans?</i>	<i>Given current moral and political interests, with which terms should we refer to historical, contemporary, projected, and new social kinds of humans?</i>
Ameliorative Freedom	C	Very Minimal	Moderate	Moderate	Moderate +	Moderate +
Normatively Constrained by:	D		B1+	B1+, B2	B1+, B2, B3	B1+, B2, B3, B4
B1-B5 Applied to Remote Drone Operators	E	Clotting/anchoring mechanisms: historical training environment, occupational settings, etc. Characteristic properties: stress, emotional fatigue, moral conflict, burn out, whiplash (Hijazi et al. 2017)	e.g., Engineer changes to kinds referenced in (E1) to reduce the incidence and severity of undesirable characteristic properties (acute stress, etc.)	e.g., Engineer a new social kind of human that oversees the moral decision making and mental health of drone operators e.g., Eliminate the kind	e.g., Politically frame remote drone operators in terms of the mind-independent social kind of human referenced in (E1). e.g., Politically prioritize a disjunctive category comprised of kinds believed to perpetrate morally unjustified military aggression	e.g., “drone operators,” “UAV operators,” referencing the mind-independent social kind from (E1). e.g., “perpetrators of military aggression,” highlighting the disjunctive class from (E4)

B1-B5 Applied to Men and Women	F	e.g., Clotting/anchoring mechanisms: differential socialization, prestige bias; conformist bias; historical binary gender system; etc. (Bach 2012) Characteristic properties, e.g.: positions in social hierarchy, bodily presentation, division of labor, psychological characteristics, etc. (Bach 2012).	e.g., Engineer changes to kinds referenced in (F1) to eliminate oppressive aspects of gender role, secure equal opportunity, secure fairness, etc. (Bach 2012)	e.g., Historical gender kinds outside the man/woman binary	e.g., Represent political gender as historical gender kinds referenced in (F1). e.g., Represent political gender kinds as hierarchical categories (Haslanger 2012) e.g., Represent political gender kinds in terms of the disjunctive gender identity-or-class (Jenkins 2016)	e.g., “men” and “women” used to refer to historical gender kinds as referenced in (F1). e.g., “men” and “women” used to refer to privileged males and subordinated females (F4) e.g., “men” and “women” used to refer to categories of gender-identity inclusive of transgender persons-or-gender as class (F4)
--------------------------------	----------	--	--	--	---	---

As we saw earlier, answering this question requires that we defer to what the mind-independent empirical word tells us about underlying anchoring or clotting mechanisms and the characteristic properties that they fashion for kind-members. The researcher’s freedom to advance ameliorative stipulations when addressing this question is thus severely limited. At this very moment, the social world and its kinds are what they are. If social theorists want to produce models of this world to inform legal and social policy – to facilitate the interventions that Hacking highlights – then these models need to map real social kinds accurately.

To be clear, our empirical knowledge of real kinds, social or other, is never complete, and as researchers form hypotheses and choose what to investigate, background values play a role. In addition, real kinds can overlap and nest (consider the case of biological species), affording further opportunity for guidance from values. Nevertheless, the overlapping and nested kinds are neither disorganized jumbles nor property classes the extensions of which are tethered to armchair or ameliorative stipulation. They are real kinds subject to *a posteriori* discovery and investigation. To treat such real kinds epistemically *as* real kinds requires deferring epistemically to what we learn about them, which (as discussed earlier in the context of the interconnectedness

of social kinds) can often be surprising or incompatible with prior ameliorative stipulation. In sum, there is very minimal ameliorative freedom to stipulate outcomes to AQ1.⁷

AQ1 should be distinguished from the following, prospective question:

Ameliorative Question 2 (AQ2): On the Investigation of Changing Social Kinds of

Humans: If contemporary social kinds of humans can survive change, how should our best moral and political theories direct change to the underlying anchoring or clotting mechanisms that fashion these kinds?

If answering AQ1 tells us what type of hand has been non-accidentally dealt to a group of people, answering AQ2 aims to inform us how that hand might be played to best address vetted political and moral interests. The future is open, so there is more ameliorative freedom to stipulate answers to this question. But even here, the outcomes of AQ1 impose a significant degree of normative constraint. A social kind's causal environment consists in many other social kinds, each with competing interests and causal dispositions. Bringing about desired change is always a matter of navigating this social environment of real kinds the causal regularities of which are best understood through the types of empirical investigations that correspond to AQ1. In addition, empirical investigation into a kind's clotting mechanisms and characteristic properties may indicate that there are particular changes that it cannot survive, in which case moral prescriptions for change are more effectively construed as recommendations for the creation of new social kinds of humans (or the elimination of contemporary kinds).

Consider the case of gender. There are good grounds for understanding *men* and *women* as social kinds for which kind-members share characteristic properties on account of historical processes (e.g., differential socialization, prestige bias, etc.) that make sexed individuals reproductions of either historical men or historical women (Bach 2012). Now, in the same way that a historically unified biological kind can undergo changes to its characteristic properties and

⁷ For additional discussion of the epistemic difference between *classes* and *real kinds*, see especially Millikan (2010). For discussion of why this epistemic difference undercuts the strategy of appealing to Dupre's "promiscuous realism" to support the ameliorative strategy, see Bach (2016).

still be the same kind, there are morally prescribed changes that historically unified social kinds like *men* or *women* can undergo – for example, gaining equal powers and opportunities relative to one another – and still be the kinds *men* and *women*. But determining which actual material changes and interventions are required to bring about this change – and not just changes to the historical kinds men and women, but also changes to medical professionals, educators, and other social kinds of humans that impact gender – requires understanding these kinds (their clotting mechanisms, their characteristic properties, their capacity for change) as the kinds that they are.

Our moral and political desire for change may often exceed what is possible or plausible concerning the trajectory of contemporary social kinds of humans. For example, rather than (or in addition to) engineering social change to the kinds *men* and *women*, we may want to engineer fundamentally new social kinds of humans in the context of gender. This motivates an additional prospective ameliorative question:

Ameliorative Question 3 (AQ3): On the Investigation of New Social Kinds of Humans:

Given current moral and political interests, what new social kinds, and thus what underlying anchoring or clotting mechanisms, do we want to bring into (or out of) existence?

It was only recently that advanced militaries decided to use unmanned aerial vehicles for combat and surveillance purposes. Under direction of a specific set of background values, they plotted how to bring into existence a distinct social kind, the drone operator.

How much stipulative freedom applies to AQ3? On the one hand, the future is open, so there is reasonable ameliorative freedom when stipulating what new kinds we want to bring into (or out of) existence. On the other hand, the future is inevitably crowded by the causal trajectories of other contemporary real kinds. This means that, as was the case for AQ2, new kinds can only be brought into existence by navigating an empirical route through the world as delineated by AQ1 (and to some extent AQ2). In this respect, the outcomes of AQ2 and particularly AQ1 normatively constrain the outcomes of AQ3.

I turn now to two value-laden questions that at first glance appear quite untethered from the *a posteriori* study of mind-independent kinds. But as we will see, these questions are also normatively constrained to an important degree by the outcomes of empirical investigations relevant to AQ1 (and to some extent AQ2 and AQ3).

Ameliorative Question 4 (AQ4): On Political Construals of Historical, Contemporary, Changing, and New Social Kinds of Humans: Given current moral and political interests, how should we politically frame or represent historical, contemporary, projected, and new social kinds of humans?

The world might tell us what kinds of humans there are at a given time (AQ1), but it does not dictate how we must politically, socially, and legally frame or represent (or fail to represent) those kinds. For example, the category *people who were born on an even-numbered day of the month* is not a social kind of human; there are no distinguishing characteristic properties for *people who were born on an even-numbered day of the month*, and this is because there are no social clotting mechanisms causally structuring similarities between *people who were born on an even-numbered day of the month*. Nonetheless, we are free to frame this as an important political class. That is, it is ultimately “up to us” whether to designate *people who were born on an even-numbered day of the month* as a political class with special privileges.

Surely, politically highlighting *people who were born on an even-numbered day of the month* will not serve vetted background moral and political interests. But there are other classifications that while not in a one-to-one mapping relationship with mind-independent social kinds of humans might still serve important normative interests. Call these “nominal political kinds of humans.” One example of nominal political kinds is disjunctive classes of real social kinds of humans. A *posteriori* investigation tied to AQ1 reveals that *U.S World War II veteran* and *U.S Iraqi War veteran* are distinct social kinds of humans, with different sets of characteristic properties on account of different historical clotting mechanisms. But there are a variety of moral and political reasons to highlight a more inclusive disjunctive category – *U.S War Veteran* – that include centralizing veteran services, recognizing heroism, and recognizing patriotic sacrifice.⁸ But note that the category is still rooted in and has its political importance in virtue of what we know about the causal-historical reality of the category’s disjuncts, which are social kinds of humans as discoverable through AQ1. Moreover, how best to bring about change

⁸ Over time, such efforts can contribute to historical clotting mechanisms that anchor a distinct social kind of human. I further discuss this type of possibility below (but the general dynamic is a familiar one – see especially Hacking’s writings).

to this disjunctive category will depend on how best to bring about change to the social kinds of humans that are its disjuncts (AQ1 and AQ2).

Another example of nominal political kinds is classes unified on the basis of a particular social property. Consider Haslanger's hierarchical definitions of gender groups: the property "subordinated on the basis of perceived/imagined female role in biological reproduction" determines the extension of the term *women*; the property "privileged on the basis of perceived/imagined male role in biological reproduction" determines the extension of *men*. The motivation for this classification is political: the account of *women* in particular is supposed to keep us focused on "the ones that matter," claims Haslanger (2012, p. 240). Putting aside here whether this classification does in fact serve these normative interests (see the discussion of Saul's objection from §3), I want to focus on how its potential to do so must be (or ought to be) premised on the types of *a posteriori* investigation that are relevant to AQ1 and AQ2. Indeed: the potential value of Haslanger's proposal, I suggest, requires it being reworked in reference to PRK, AQ1, and AQ2 in the following way. Begin by recognizing that current gender categories are real social kinds of humans, causally sustained by historical gendered institutions, and only discoverable through *a posteriori* investigation. Next, observe problematic moral and political features of these mind-independent kinds and prescribe morally appropriate changes. Finally, claim that an important way to bring about the prescribed change is to focus social theorists' attention on a particular *property* of the contemporary kind – subordination/privilege – and then call *that* "men" and "women." Whether or not we agree that this maneuver would in fact help bring about the prescribed change to the underlying real social kinds of humans (I am skeptical), it is important to understand how proposals like Haslanger's are implicitly premised in this way, and thus normatively constrained by, the empirical realities of real social kinds of humans. Failure to understand this makes it likely that social theorists with shared political goals will talk past one another.

So far, I have steered clear of questions about what the terms associated with social kinds of humans – "women," "men," "drone operator," "millennial," "analytic philosopher," "continental philosopher," etc. – "really mean" or refer to. This is not an accident. The fulcrum of the analysis on offer (and what set the stage for AQ1-5) is the political and moral importance of accurately identifying and describing mind-independent social kinds of humans, regardless of what these kinds are called (or not called). That is, what these kinds *are* – their clotting

mechanisms, their characteristic properties – is distinct from what terms we (folk or social scientists) use to refer to them. For example, even if we did not have a concept or term “opioid addict,” there would still be a mind-independent social kind of human that is of great moral and political significance.

Of course, there are still many important normative issues connected to how we use our social kind language. In particular, how we use social-kind terms contributes to the clotting mechanisms that sustain a kind’s characteristic properties. Less appreciated, I submit, is how our use of social kind language, especially in the context of productive social scientific investigations, has an important impact on our capacity to gain knowledge and effectively communicate about real social kinds of humans (thus satisfying the morally important PRK). These observations motivate a further value-laden investigative question:

Ameliorative Question 5 (AQ5): On the labeling of Historical, Contemporary, Projected, and New Social Kinds of Humans: Given current moral and political interests, with which terms should we refer to historical, contemporary, projected, and new social kinds of humans?

As we did earlier, let’s suppose that social scientific research into how social institutions have historically organized individuals’ different roles in biological reproduction provide strong grounds for identifying real gendered kinds of humans in terms of something like historical lineages (Bach 2012). Should we then call *these* mind-independent kinds “men” and “women?” Or should we call them something else? Should we reserve the terms “men” and “women” to refer to Haslanger’s hierarchy-based property classes? Or should the terms “men” and “women” be used to refer to classes considered more fully inclusive of transgender persons?⁹ Some of the different options are given in cells (F4) and (F5) of fig. 1. While these issues are especially salient in discussions about gender-related social kinds, they generalize to discussion about almost any politically important social category.

I will not attempt full answers to these questions here. Instead, I want to raise several overlooked considerations about these questions that follow from the preceding discussion about

⁹ Both Haslanger’s and Bach’s accounts have been challenged on the grounds that they might exclude some transgender persons. Mikkola (2016b) argues this way against the account of gender given in Bach (2012), and Jenkins (2016) argues this way against Haslanger’s (2012) account of gender.

the relationship between value-laden and empirical investigations into social categories. These considerations are meant to inform a future and more in-depth analysis of AQ5.

First, even if real social kinds are not in the ambit of folk terms (referentially or descriptively), these mind-independent kinds still possess their moral/political significance in the context of responsible social modeling. Thus, the verdicts of conceptual analysis and ordinary language analysis do not function as arguments against PRK or as arguments against the framework for ameliorative questions as set out above.

Second, for the purpose of characterizing historical and contemporary social kinds (as well as the prospective kinds relevant to AQ2 and AQ3), it is possible to introduce variables, existential quantifiers, and sentences that describe causally (functionally, theoretically, etc.) relevant relations as determined by *a posteriori* investigation. This use of “Ramsey sentences” allows the removal of familiar and contested social kind-terms like “men” and “woman.” Social theorists could use this method to express the outcomes of value-driven investigations AQ1, AQ2, and AQ3. Then, in reference to these Ramsey sentences, theorists could isolate questions pertaining to linguistic practice, for example: “which of these kinds, if any, has historically accommodated past folk and/or scientific referential use of the social term *P*?”; “Given current moral and political interests, with which terms should we refer to these kinds?”. This exercise helps demonstrate how value-laden investigative questions about ontology come apart from value-laden investigative questions about linguistic practice. It also demonstrates why debates about the latter questions should not function as arguments against outcomes to the former questions.

Third, political arguments about what label to use for both real social kinds and nominal political kinds need to consider the coordinating function of language. This is often overlooked by researchers who endorse stipulative classifications. Suppose that the expression “drone operator” was used to refer to mind-independent drone operators as well as McDonald’s employees. This would impede effective information gathering and communication about both kinds, obstructing PRK and AQ1, thus subverting moral and political interests. This concern is amplified in cases where we have good reason to think that social scientists are tracking mind-independent social kinds of humans with specific concepts and associated terms. These are generally cases in which the research program exhibits considerable predictive and explanatory success (see, e.g., Richard Boyd’s “accommodation thesis” in Boyd 1999). Saul (2006) and

Mikkola (2016b) raise a concern that Haslanger's ameliorative analysis of gender might create miscommunication between feminists and ordinary speakers. I think this is right, but I also think that the problem is more general (and more worrisome). There is a division of labor among social scientists (as well as researchers generally, and even laypersons) with respect to the epistemic project of tracking and accurately describing real kinds. Language plays a vital role in coordinating this project, indicating to researchers whether they are gathering information about the same thing (Millikan 2017 explains this function of language in detail). There would need to be significant political benefits for accepting semantic proposals about social categories if those proposals would disrupt this coordinating function.

Drawing from Hacking's notion of "looping effects," we may want to use politically stipulated terms (AQ5) to refer to a nominal political kinds (AQ4) the highlighting of which we think will be instrumental to bringing about morally prescribed change (AQ2) to contemporary social kinds (AQ1). Above, this is how I recommended that we rework Haslanger's account of gender, and I flagged how such manipulations of social meaning, if we want them to be effective, are still beholden to the causal profiles of real social kinds of humans. Here, I am flagging the additional point that any purported political gains for the stipulative manipulation of social-kind language must always be weighed against the moral and political risk of impeding knowledge-gathering and knowledge-sharing, thus undercutting the moral and political contributions of PRK and AQ1.

§5: Conclusion

Social researchers agree that the project of modeling social kinds of humans should be guided by background normative interests. They further agree that models of social kinds should comport with mind-independent facts. Unfortunately, the alleged fit and potential constraints between these two commitments have not been sufficiently interrogated and analyzed. As a result, several confusions have set in, largely unnoticed, undercutting widely shared goals for the responsible modeling of social kinds of humans. This article has sought to address this situation by providing a more detailed exploration of the relationship between value-laden and primarily empirical investigations into social kinds of humans.

Several themes emerged. There are real, mind-independent, contemporary social kinds of humans, discoverable through *a posteriori* investigation. The reality and causal profiles of these kinds have great moral and political significance, and they have this significance for any set of values that requires changing or intervening with social reality. We do not have stipulative freedom in reference to our moral interests to say what these kinds are. We can prescribe ways that we want these kinds to change, or for new kinds to develop, but for these recommendations to serve normative interests effectively, they need to be based upon the nature of contemporary social kinds of humans as revealed through *a posteriori* investigation.

If there are good moral grounds for advancing nominal (non-social kind) classifications and associated labels, such grounds can never function as arguments against either the mind-independent reality or moral importance of real social kinds of humans as revealed through *a posteriori* investigation. In fact, many arguments in favor of nominal classifications are better cast as arguments about how best to engineer change to contemporary social kinds, in which case stipulative claims about nominal classifications are really ways of manipulating causal regularities that currently maintain contemporary social kinds of humans. The benefits of such manipulations would need to be weighed carefully against the risks of disrupting the coordinating function of social kind-term language, especially in the context of productive social scientific research programs.¹⁰

REFERENCES

Anderson, Elizabeth (1995): "Knowledge, Human Interests, and Objectivity in Feminist Epistemology". *Philosophical Topics* 23. No. 2, p. 27–58.

Bach, T. (2012). Gender is a natural kind with a historical essence. *Ethics*, 122(2), 231-272.

Bach, T. (2016). Social categories are natural kinds, not objective types (and why it matters politically). *Journal of Social Ontology*, 2(2), 177-201.

¹⁰ I want to thank the editors and anonymous reviewers for their comments on earlier drafts of this article. The article was much improved as a result. Of course, any remaining mistakes or oversights are completely my own.

- Bilsker, D., Fogarty, A. S., & Wakefield, M. A. (2018). Critical Issues in Men's Mental Health. *The Canadian Journal of Psychiatry*, 0706743718766052.
- Boyd, R. (1999). 'Homeostasis, Species, and Higher Taxa', in R. A. Wilson (ed), *Species: New Interdisciplinary Essays* (Cambridge: MIT Press)
- British Columbia Coroner's Service (2018) "Illicit Drug Overdose Deaths in BC January 1, 2008 – February 28, 2018" (<https://www2.gov.bc.ca/assets/gov/public-safety-and-emergency-services/death-investigation/statistical/illicit-drug.pdf>)
- Epstein, B. (2014). How Many Kinds of Glue Hold the Social World Together?. In *Perspectives on social ontology and social cognition* (pp. 41-55). Springer, Dordrecht.
- Griffiths, P. (1997). *What emotions really are: The problem of psychological categories*. University of Chicago Press.
- Hacking, Ian (1995): "The Looping Effects of Human Kinds". In: Dan Sperber, David Premack, and Ann James Premack (Eds.): *Causal Cognition: A Multidisciplinary Debate*. Oxford: Clarendon Press, p. 351–394.
- Haslanger, S. (2000). Gender and race: (What) are they? (What) do we want them to be?. *Noûs*, 34(1), 31-55.
- Haslanger, S. (2012). *Resisting reality: Social construction and social critique*. Oxford University Press.
- Haslanger, S. (2014). Race, intersectionality, and method: a reply to critics. *Philosophical Studies*, 171(1), 109-119.
- Hijazi, A., Ferguson, C. J., Ferraro, F. R., Hall, H., Hovee, M., & Wilcox, S. (2017). Psychological Dimensions of Drone Warfare. *Current Psychology*, 1-12.
- Jenkins, K. (2016). Amelioration and inclusion: Gender identity and the concept of woman. *Ethics*, 126(2), 394-421.
- Lessig, L. (1995). The regulation of social meaning. *The University of Chicago Law Review*, 62(3), 943-1045.
- Locke, John, *An Essay concerning Human Understanding*, ed. Peter H. Nidditch (Oxford: Clarendon Press, 1975)
- MacKinnon, C. A. (1989). *Toward a feminist theory of the state*. Harvard University Press.
- Mallon, R. (2003). Social construction, social roles, and stability, in Schmitt, F. F. (Ed.). *Socializing metaphysics: The nature of social reality*. Rowman & Littlefield.

Mikkola, M. (2015). Doing ontology and doing justice: What feminist philosophy can teach us about meta-metaphysics. *Inquiry*, 58(7-8), 780-805.

Mikkola, M. (2016a). Feminist Metaphysics and Philosophical Methodology. *Philosophy Compass*, 11(11), 661-670.

Mikkola, M. (2016b). *The Wrong of Injustice: Dehumanization and Its Role in Feminist Philosophy*. Oxford University Press.

Millikan, R. G. (1999). Historical kinds and the “special sciences”. *Philosophical Studies*, 95(1-2), 45-65.

Millikan, Ruth (2010): “On Knowing the Meaning; With a Coda on Swampman”. In: *Mind* 119. No. 473, p. 43–81.

Millikan, R. G. (2017). *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford University Press.

Mills, C. W. (2014). Notes from the resistance: some comments on Sally Haslanger’s Resisting Reality. *Philosophical Studies*, 171(1), 85-97.

Root, M. (2000). How we divide the world. *Philosophy of Science*, 67, S628-S639.

Saul, J. (2006, June). II—Jennifer Saul: Gender and Race. In *Aristotelian Society Supplementary Volume* (Vol. 80, No. 1, pp. 119-143). Oxford, UK: Oxford University Press.