

Gender as Interpretation

Alice van't Hoff

“I am a woman.” So says Ava. But what could she mean in making this claim? The *obvious* interpretation is that she has asserted a fact about her gender. Adherents of the position that I call “minimal realism” about gender affirm this obvious answer. I will call the view of their opponents “deflationism”.¹ Minimal realism and deflationism are defined in terms of their position on the following principle:

Minimal Realism (MR): Gender classifications should be accurate and they are accurate when they track the objective and universal gender kind membership facts.

Minimal realists affirm and deflationists deny this principle, which I henceforth refer to as “MR” for short.

I’m going to argue that the flight from minimal realism is a mistake. I’ll start in §1 by sketching the conceptual terrain. In §2 I’ll set out the two arguments that support deflationism; I’ll suggest that while neither argument succeeds, there is an important challenge for the minimal realist in the vicinity, namely showing just what the objective gender kinds might be without accepting normatively unpalatable conclusions. Nonetheless, there are also considerations *favoring* realism, which I explore in §3: only the realist, I’ll suggest, is in a position to fully explain why misgendering is wrong. This gives rise to a dilemma: while minimal realism is attractive, it seems unfeasible. In §4 I sketch a solution to this impasse that treats gender kinds as *interpretive*. To foreshadow: a central theme below is the tension between the demands of normativity and objectivity. Interpretivism allows us to reconcile these demands by treating normative facts as partly constitutive of the socio-metaphysical kinds.

¹ See Dembroff 2018, Dembroff 2020, p. 12, Barnes 2020, Bettcher 2009, Bettcher 2013, and Bettcher 2014. Jenkins 2023 also develops a sophisticated deflationist view that I have unfortunately been unable to discuss here. For an opposing perspective see e.g. Mikkola 2006, Bach 2016.

1 Gender deflationism

The example of Ava notwithstanding, minimal realists need not insist that all grammatically assertoric classificatory claims are best interpreted as straightforwardly descriptive. Realism is a socio-metaphysical rather than semantic thesis. (It does, of course, have semantic implications: if MR holds, then classificatory claims could feasibly be straightforwardly descriptive; for some other views, by contrast, this is not possible.) Deflationists and minimal realists can also agree about how we ought to classify people by gender: both parties may agree, for example, that it is true to say that Ava is a woman. We can instead think of the difference between deflationists and minimal realists ground-theoretically; their disagreement emerges when we consider what *grounds* the appropriateness of some classificatory practice.²

Deflationism about gender is not a monolith; there are different versions of the view with different strengths and weaknesses. One option, then, is what we might call “gender relativism”. According to this view the appropriate practices of gender classification will vary across communities.³ This approach is naturally paired with a kind of semantic contextualism, which holds that the *truth* of gender ascriptions is tied to the particular contexts to which the utterance of the ascriptive claim is indexed. An alternative form of deflationism is expressivist: expressivists treat gender ascriptions as *expressing* some kind of non-cognitive mental state (McGrath 2021). Roughly, what is expressed is a plan to treat the person(s) to whom we ascribe a gender in a certain way (or perhaps, our acceptance of a system of norms that would justify the adoption of such a plan) (Gibbard 1990). A third option is what we might call “anti-representationalism”. Views like these are somewhat analogous to error-theoretic metaethical views. As I am understanding them, anti-representationalists hold not just that gender classification is not truth-apt but further embrace this and call for a reform of classificatory talk and practices. Correspondingly, anti-representationalists will reject any programs aimed at recovering the cognitivist trappings of gender classification. As I interpret them, Bettcher 2014 and Barnes 2020 defend versions of anti-representationalism.

² Analogously, grounding-theoretic ideology helps to make sense of metaethical debates between naturalists and non-naturalists, see Rosen 2017.

³ This is one of the most popular approaches in contemporary philosophy of gender, see Stoljar 2011, Ásta 2013, pp. 722–7. Dembroff 2018, Saul 2012, Diaz-Leon 2016.

2 Arguments for deflationism

Two arguments seem to support deflationism: in §2.1 I canvass the first and in §2.2 I sketch the second. I will argue that although neither argument is dispositive as stated, there is an important truth in the vicinity, which the arguments are attempting to track; in §2.3 I call this the “squeezing argument”.

2.1 Ontological oppression

The first argument (developed in Dembroff 2018) takes the form of a *reductio* directed against the Real Gender Assumption:

(RGA): Gender classifications should track the operative gender kind membership facts (Dembroff 2018, p. 29).

On initial inspection the RGA seems to bear a strong resemblance to MR. The critical difference is that this principle singles out “operative gender kinds”, whereas MR focuses on “objective and universal gender kinds”. Later I will argue that this difference is crucial.

The argument against the RGA goes as follows:

- (i) Gender classifications should track the operative gender kind membership facts
(RGA, assumption for reductio)
- (ii) Some operative gender kind membership facts are unjust **(Ontological Oppression)**
- ∴ (iii) Gender classifications should sometimes track unjust gender kind membership facts **(i, ii)**
- (iv) Gender classifications should never track unjust gender kind membership facts
(Moral Assumption)

By conjunction introduction, premises (iii) and (iv) give rise to a contradiction. We are assuming premise (i) for the sake of *reductio*, and premise (iv) is a methodological precept for which I will not argue (cf. Jenkins 2016, p.396). Only (ii) thus requires substantial defense.

Before presenting such a defense, we must clarify what exactly is meant by “operative gender kinds”. The terminology is Haslangerian: as defined by Haslanger, the operative gender concept

is the concept that “actually determines how we apply [gender] terms to cases” (Haslanger 2012b, p. 92). A gender “kind” is a way of grouping people by reference to their gender. Operative gender kinds, then, are classes of persons, delineated by the concepts that actually guide our classificatory behavior. But, as Dembroff points out, the social kinds that predominate in a given context are sometimes oppressive. So-called ontological oppression of this form comes in two main variants. First some social kinds are unjustly exclusionary. In 1672, for example, the Test Acts established that only those who partook of communion in the Church of England were able to hold public office in that kingdom. These acts imposed constraints on the kind of *public officials* that unjustly excluded Catholics and non-conformists. In other cases structural injustice prevents the recognition of various social kinds. This may in turn contribute to ongoing *hermeneutical injustice* (Fricker 2007). Bisexual erasure presents an example of this phenomenon. Bisexual people are *erased* when the category *bisexual* is not recognized and, accordingly, bi men are classified as gay and bi women as straight (Dembroff 2018, p. 27).

Gender kinds are not exempt from ontological oppression. Dominant practices in the West categorize people by virtue of their natal genitalia. This constructs the kind “man”, for example, in ways that exclude trans men. This is unjust because the assignment of sex to a person is morally arbitrary, and morally arbitrary categorizations do not affect a person’s social entitlements (Rawls 1971, p. 72). A different category of possible gender kind membership facts is operative in other, non-dominant environments, in which persons are assigned gender kinds on the basis of how they identify, where to identify as a certain gender is to hold certain beliefs about oneself. This implies that those who lack the cognitive wherewithal to believe that they are women could not be counted as such. This seems unjust since it excludes severely cognitively disabled women (Barnes 2022). Alternatively, some maintain that gender is determined by one’s relations to social patterns of oppression that are fixed by one’s perceived reproductive role. This excludes non-passing trans people who may not be perceived to occupy a reproductive role that corresponds to their gender. This too seems unjust. Premise (ii) and thus the falsity of the RGA seems to follow.

Nonetheless, there is something puzzling about the argument. Given the definition of “operative gender kinds” we can parse the RGA as follows:

(RGA*): Gender classifications should track the classes delineated by our actual practices of classification.

Assuming that “track” is understood as something like “be correlated with”, RGA* looks analytic or nearly so. It’s hard to see how classificatory procedures could fail to produce anything other than the classes that they *do* produce. But if $RGA = RGA^*$ and if RGA* is analytic, then the denial of RGA that was supposed to follow from Dembroff’s argument should be contradictory.

Something must have gone wrong. The solution, I think, is to recognize a shift in the implicit agential parameter between the first and second occurrences of “classification”. The principle thus comes out as:

(RGA+): *Our* gender classifications should track the classes delineated by *socially prevalent* practices of classification.

If this is the right interpretation, then the target of Dembroff’s argument is definitely *not* analytic. Moreover, RGA+ is clearly falsified by the argument from ontological oppression: if we should avoid injustice then we should not blindly follow social custom, since that custom is not invariably just. RGA+ does not, however, seem to stand in any particularly close connection to MR. The combination of $\neg RGA+$ and MR looks, on its face, perfectly consistent. This is not a problem for Dembroff’s argument *per se*, but only for those who seek to use the conclusion of that argument to rule out MR. But, given this, the possibility of ontological oppression is not, enough to establish deflationism.

2.2 The freestanding argument

Deflationists may then turn to what I have called the “freestanding argument”. The thought here is that gender classification stands free of and is largely unconstrained by the objective facts about gender, whatever these might be. An argument of this kind is given in Barnes 2020. In support of her proposal, Barnes points to a (realist) trend in contemporary metaphysics that allows that how things are in reality may not fix what we can truly say in natural language (Barnes 2020, p. 711). Her citations suggest that what she has in mind here are projects that seek to characterize

metaphysical reality in a joint-carving or non-superficial language, a kind of *ontologese*, or alternatively those that regard metaphysical inquiry as an investigation into the *truthmakers* or *grounds* of more quotidian truths.⁴ Conversely, truth in natural language seems to come fairly easily: broad agreement about the use of a term is sufficient or near-sufficient to fix the truth of many claims articulated by using the relevant term (Barnes 2020, p. 712). Call approaches that separate natural language and metaphysical reality in this way “metaphysicalist”. Applied to the case of gender, metaphysicalism favors a *pluralist* picture of gender classification, which I take to be the view that there is no *one* concept picked out by our gender-talk, but something like an overlapping family of polysemous notions. Our folk concept of gender is a “many-splendored thing” (Barnes 2020, p. 722), trying to forcibly identify this multifaceted practice with a unitary metaphysical foundation prevents us doing justice to either.

Recent discussion in the philosophy of gender has focused heavily on the extent to which various proposed analyses are exclusionary and how exclusion could be avoided (e.g. Jenkins 2016; Barnes 2022). But if we are pluralists then the pursuit of inclusion is chimerical: if there is an irreducible multiplicity of gender kinds then searching for a single all-encompassing kind is a fool’s errand (Barnes 2020, p. 720). Moreover, if classificatory practices float free of the metaphysical substructure, then we should not expect folk use of gender terms to have a significant bearing on the actual metaphysics. Inclusiveness is a normative desideratum for various social institutions, not necessarily a constraint in our explication of the metaphysics. Realist presuppositions, the idea seems to go, have led philosophers astray. Instead of investigating what ultimately explains gender, they have become mired in unproductive quarrels about the extensions of gender terms (Barnes 2020, p. 706). The freestanding argument suggests that to advance the debate we need to give up a misguided commitment to principles like MR.⁵

There are two important claims here: first that classificatory practices float free of the metaphysical underpinnings and second that the pursuit of inclusiveness pertains to the free-floating superstructure and not to the metaphysics. Both claims are misleading. *Pace* Barnes, I am skepti-

⁴ See n.27 p. 726 Barnes 2020.

⁵ It’s worth noting that gender pluralism, as I understand it, largely comes apart from the question of whether we need an ontology of gender for the purposes of feminist politics. I agree here with Mikkola 2016, pp. 105–17 who argues for a negative answer to this question.

cal that questions about the membership of gender kinds can be fully divorced from the broader project of giving a philosophical explanation of gender structures. Although I am sympathetic to metaphysicalism, it is worth noting that, even when it comes to the questions of traditional metaphysics, it has come in for formidable criticism (Burgess 2008). More strikingly still, as Barnes herself has suggested, accounts that identify metaphysics with the study of what is fundamental—a standard gambit of metaphysicalists—do not provide us with an appropriate characterization of feminist metaphysics (Barnes 2014). This generates a problem because once we concede that feminist metaphysics doesn't track fundamental reality,⁶ we deprive ourselves of many of the resources that made the metaphysicalist move plausible. Simplifying significantly, the metaphysicalist proposal allows us say that the everyday truths may come apart from the facts comprised in deep metaphysical reality. The challenge for the advocate of this idea is to explain why it doesn't give rise to inconsistencies. Metaphysicalism purports to avoid the problem either by distinguishing the language of metaphysics from our everyday discourse or by arguing that the facts of metaphysics are about something different from the banalities of common sense—the grounds or truthmakers of our reality. But, when it comes to the metaphysics of gender this move won't work. Unlike in traditional metaphysics, ordinary talk of gender and the phenomena of interest to feminist metaphysicians share a common subject matter; it is not implausible that when I say there are three beers in the fridge I am not making a claim about platonic abstracta. It is rather harder to insist that those who deny that Elliot Page is a man are controverting the view that accords people first-person authority over their gender.⁷

That is not to say that the metaphysics of gender should embroil itself in semantic or conceptual analysis. Rather the point is that debates about the inclusiveness of various proposals *were never about language*. It is objectionable for an analysis of gender to be exclusionary *not because* this falsifies some true claim about a person's gender, but because it indicates that we are no longer tracking the right phenomenon. To take an example, the analysis of gender given

⁶ Barnes 2014, p. 349, see also Barnes 2020, n.31 p. 726.

⁷ The metaethical analogy is again suggestive: one way to interpret the critique of deflationary metaethical views in Dworkin 1996 is as a denial that there is conceptual space to separate first-order and metaethical questions. The move here is analogous: I am claiming that the question of who counts as a wo/man cannot be separated from the substantive metaphysical issues.

in Haslanger 2012a has been criticized for excluding certain trans people. The criticism here is not that if Haslanger were right some trans women (say) could not truly describe themselves as women. Instead, it is that since these people *are* women, if our account isn't tracking them then it won't give the right explanation of what gender is.

2.3 The squeezing argument

So far, then, I have given fairly short shrift to the arguments that purport to motivate deflationism. *Nonetheless*, there is an important truth in the vicinity. Put very simply, the challenge for the opponent of deflationism is to say what the "objective and universal gender kinds" referenced in MR *are*. The deflationist bets that no satisfactory answer to this question can be given.

By and large, theorists of gender have thought that gender kinds are *social* rather than *natural* kinds.⁸ Disagreement about the details here is pervasive, but the general idea is clear: natural kinds are joint-carving, they support inductive explanation, and allow us to type entities by their intrinsic or essential properties; social kinds are contingent, extrinsic, and are typically *constructed*.⁹ Whereas natural kinds correspond to putatively scientifically respectable distinctions, to recognize something as a social kind is frequently to debunk or undermine its epistemic credentials. This helps to explain why feminists have generally agreed that gender kinds are social: in doing so, they take themselves to be exposing purportedly natural hierarchies as human creations and thus creating space for oppositional politics (Haslanger 2012c).

If MR holds, then there are objective and universal gender kinds. Given the foregoing, these seem to be social kinds, which I am understanding here to imply a degree of social construction. One problem is that, given the fact of socio-cultural variation, being socially constructed seems inimical to the ostensible universality and (to a slightly lesser degree) objectivity of gender kinds. But the possibility of accounts like Haslanger 2012a, which analyzes gender kinds as generic positions in a hierarchy of subordination in accordance with perceived reproductive role, seems to show that this challenge can be successfully negotiated. A related problem, however, emerges

⁸ Bach 2012 is a notable exception. One way to escape the squeezing argument is to follow him in treating gender as a natural kind; for reasons of brevity I won't explore this further here.

⁹ The most influential analysis of natural kinds is given in Kripke 1980, pp. 117–28, pp. 134–40. For one account of the construction of social kinds see Ásta 2018.

if we take seriously the threat of ontological oppression. I argued earlier that accepting that our gender classifications need not defer to the socially prevalent gender kinds does not rule out MR. The challenge then is to find a gender kind that is a). objective and universal b). distinct from the oppressive socially prevalent gender kinds.

3 A dilemma for deflationists

3.1 Miscategorization

In this section, I will suggest that there is an important regard in which deflationists are at a disadvantage relative to minimal realists. Deflationists of all stripes are vulnerable to a dilemma that arises when they are asked to say what gender ascriptions *should* track. If they answer “nothing”, then they fail to provide adequate normative guidance: to draw certain distinctions on the basis of age or race, for example, would clearly be *morally* objectionable even if we are no longer realists. If though, they give a positive answer instead, then by virtue of their rejection of MR they are barred from giving the requisite justification for this answer (or at the very least, a necessary part of the justification), namely that alternatives *miscategorize* persons.

The second horn requires clarification: the thought is that deflationists do not have an adequate explanation of why we ought to follow the practices of gender ascription which they urge upon us. That’s not to say that we are gripped by the thought that we should *not* follow such practices, but rather that even when we recognize that this is what we ought to do we lack a justification for acting as we ought. Consider, for example, the kinds of ontological oppression surveyed above. Cases where social kinds are unjustly exclusionary seem to be straightforwardly objectionable: constructing the kind “public officials” in a way that excludes people on the basis of their religion violates a right these individuals have to participate in the government of their community. But the second case, where social structures unjustly inhibit the recognition of social kinds, is more complicated. What is it to recognize a social kind? Presumably, what happens is some antecedent kind is acknowledged as such and comes to be an operative category within that society. Since this is an injustice, something about the antecedent kind must ground some kind of entitlement

to social recognition. Yet for a social kind *s* to ground a duty to establish or respect an *operative* category of all and only those who are *s* requires that there are social kinds that are antecedent to their *actual* social uptake. Consider again the case of bisexual erasure: the wrong here is not that “gay” and “straight” are intrinsically oppressive kinds, but that “social structures and practices fail to recognize the kind *bisexual*” (Dembroff 2018, p. 27). But this explanation seems to require an analog of MR for the case of sexualities.

We could argue that the kind “couples” is constructed in exclusionary ways that marginalize bisexual people. But this mislocates the wrong: the problem of bisexual erasure isn’t primarily that there are social and political barriers to bi people being in fulfilling and adequately recognized relationships (not that there aren’t such barriers, just that the problem that they give rise to isn’t primarily one of erasure), rather it is that on entering into such relationships, bi persons are not recognized as bisexual. In other words, the issue is not that certain forms of romantic or sexual conduct are socially disfavored, but that these forms are never recognized as bisexual. The wrong in these examples *is not* that some kinds have unjust membership conditions, but that a person’s membership in some kind is not appropriately respected. This is precisely what minimal realism calls our attention to: there are some (respect-worthy) social kinds that our classification practices are failing to track accurately. In rejecting MR we lose ability to draw our attention to *these* wrongs.

When we ask what is the wrong involved in misgendering a trans woman it doesn’t *just* seem to lie in the fact that we have construed the category “woman” in an exclusionary fashion. If all we are doing is insisting that she is “really” a man, then we would seem thereby to be endowing her with the rights and privileges of manhood. Men are (generically) materially advantaged by being men. Thus, even as we exclude her from one category, we place her in a different privileged category. Likewise if we insist that nonbinary persons must be men or women, we don’t automatically thereby deny them the rights accorded to cis persons. In this regard, what’s wrong in these cases seems relevantly analogous to bisexual erasure. Just as erasure may wrong bi people without directly depriving them of rights, in misgendering people, we ostensibly accord them with some materially adequate set of social entitlements, yet we nonetheless wrong them.

Of course, this is a simplification: part of the problem is that those who misgender rarely treat trans women as men, but regard them as defective or deviant. Moreover, misgenderings often express multiple other claims that may well be *false* (Barnes 2020, p. 722). Nonetheless, adding these details does not help the deflationist. Saying false things is not always ethically wrong. Even if misgenderings have false expressive concomitants, we still lack an explanation of their ethical wrongness. But moreover, such expressive content does not seem to be necessary or even universal. Ultimately, the disrespectful effect of misgendering springs from a deeper wrong: in (intentionally) misgendering persons, we deliberately disrespect them precisely *because* we disregard the entitlements connected with their actual gender. Even if misgendering is accompanied by no further consequences, it remains wrong. Although the nature of this wrong is beyond the scope of the present paper the idea is, roughly, that people are owed reasonable access to the social bases for self-respect in their identity (Rawls 1971, p. 62) and that in societies like ours in which gender is an important component of our identity misgendering denies persons such access.¹⁰

3.2 Autonomy

There is, however, a response that deflationists could give: trans inclusive practices, they might say, are superior because they defer to the autonomy of those persons who are the objects of classification. They could argue as follows: taxonomical systems imposed from without frequently function as means of controlling, surveilling, and exercising power over those who are classified (see for example Foucault 1978). When we instead grant persons authority over their own classification, we avoid the erection of an edifice of control and thus eliminate an axis of oppression from our society. The elimination of oppression is a good; trans inclusive practices are superior to alternatives since they serve this good.

Yet, while some people experience their gender identity as something that is autonomously chosen, for many it is not a thing over which they can exercise autonomous agency. This is particularly true for many trans people. By their lights, the goodness of trans inclusive practices has

¹⁰ Here I disagree with Bettcher 2014, p. 392 and Dembroff and Wodak 2018, where it is argued that the wrong of misgendering lies in the *invalidation* of an identity, rather than (as I affirm) in the absence of positive affirmation of such an identity.

nothing to do with autonomy and everything to do with their fidelity to persons' *actual* genders. The argument from autonomy thus attributes to trans inclusive practices an autonomy-promoting quality that its supposed beneficiaries will not themselves necessarily recognize. Moreover, by insisting that in respecting trans people's genders we are respecting their autonomy we ironically fail to respect the self-conceptions of those for whom their gender identity is not autonomous but an unchosen (and potentially painful) fact of life.¹¹

The argument could be complicated. Perhaps, the deflationist might insist, the point is *not* that people's genders should be respected because we should respect autonomous choice, but rather that respecting people's gendered self-identifications is to be preferred because this advances their autonomy globally. This line of argument, though, is also objectionable, on one of two counts. If the claim is that practices that respect a person's self-identification *generally* advance autonomy, then we have not yet explained why misgendering is wrong in a *directed* way. When I misgender you, I wrong you. The grounds of this wrong must surely involve some specific facts about *you*. But the proposed justification is wholly generic and as such insufficient. Alternatively, the claim may be that respecting a person's self-identification is part of a practice that *necessarily* advances *their very own* autonomy. But this claim looks false. True, being able to live authentically in a long-denied identity may well promote an individual's autonomy. Such a connection is not invariable though, trans people may give up social capital, standing, and status, along with real capital, relationships of all kinds, and their own health and wellbeing in pursuit of transition. It is implausible to claim that their autonomy is invariably advanced, unless there is some necessary connection between autonomy and gendered authenticity. If, though, there is such a connection it must be explicated, and the realist's strong suspicion is that any such explication will end up appealing to the real gender kinds that are unavailable to deflationists.

Finally, the appeal to autonomy here seems to be "one thought too many". The case is an interesting converse of the example by which Bernard Williams first introduced this idea. In his case, we are expected to extend special treatment to some without a further thought on the basis

¹¹ There is a comparable controversy over whether people "choose" to be gay. The most reasonable response here would seem to be one that respects the self-conception of gay people: both those who experience their sexuality as unchosen and those for whom it is a choice. An approach to queer theory that marginalized either perspective would be ipso facto objectionable. The same goes, *mutatis mutandis*, here.

of our special relationships (Williams 1981, pp. 17–18). Here, the idea is that if we owe an equal duty to all, then it is invidious to appeal to a special justification for our compliance with this duty in some particular case. Respecting a person’s gender is non-competitive in that according respect to one person in no way impedes our ability to extend an identical respect to others.¹² Thus the basis for the respect we extend to cis people in their gender should not differ from the basis for the respect we pay to trans people. Yet we do not respect the genders of cis people because of some fact about autonomy. We respect Barack Obama’s gender because he is a man. The same then should go for a trans person. When they discover that they are being respected out of regard for their autonomy, they are entitled to complain that this is one thought too many.

4 Interpretive gender kinds

At this juncture we face a dilemma: I have argued that we have reason to adopt something like MR. I have also suggested that the squeezing argument poses a real challenge for the advocates of that principle. We are pulled in two directions: if we are to ground what seem like intuitive moral claims we must posit objective gender kinds; according to the squeezing argument there are no such things. What gives? In this section, I will suggest a way out of the squeezing argument. My proposal is that gender kinds could be interpretive social kinds. I will claim that these avoid the challenges posed by deflationists.

4.1 Interpretivism

The notion of an interpretive kind derives from Ronald Dworkin’s jurisprudence (Dworkin 1986). I will start, then, by sketching the idea in its original jurisprudential context before considering how it might be extended to *gender* kinds. “Interpretive”, as Dworkin understands it, has a quasi-teleological ring to it: an interpretive practice is a practice to which we take the interpretive attitude, which is to say, a practice whose scope and content we do not take to be fixed and to which

¹² Some have argued against this idea: when we acknowledge trans women as women, the thought goes, we undermine their *sex*-based identity (e.g. Germaine Greer quoted in Serano 2016, p. 167, cf. Scheman 1997, p. 148). Arguably, though, the criticism here is *not* that their gender identity is not being respected, but rather that their beliefs about the *grounds* of this identity are not given due deference. This is a quite different issue.

we attribute a purpose in order to assess questions of scope and content. Interpretations are constructive in the sense that they aim to show the practice in its best light and in order to give such interpretations we must reason from the perspective of the practice itself rather than from some more personal point of view. Dworkin's theory of "law as integrity" develops this idea to give an interpretation of the law. Integrity is the virtue of principled consistency and moral coherency. According to Dworkin, this value is essential to the development of genuine associative community and helps to define the legal point of view. Law as integrity thus suggests that:

[P]ropositions of law are true if they figure in or follow from the principles of justice, fairness, and procedural due process that provide the best constructive interpretation of the community's legal practice. (Dworkin 1986, p. 225)

This interpretivist approach can, I will now argue, be extended to non-legal social practices, including our practices of gendering persons.

Several considerations favor treating gender as an interpretive kind. First such a construal makes sense of disagreement about the extension of gender kinds. It is striking that this parallels the argument that supports Dworkin's interpretivist account of law: he is concerned to explain how there could be pervasive and sophisticated disagreement about what the law is.¹³ Semantic theories explain the appearance of disagreement by claiming that the various parties are using the relevant terms in different ways. (In debates about gender, this position is most clearly exemplified by relativists, who hold that disagreement arises because in different idiolects gender terms can mean different things.) The problem, though, is that this approach cannot explain genuine disagreement: for *either* the disagreeing parties are talking past one another *or* they are using the same term in which case deep and persisting disagreement seems inexplicable (cf. Dworkin 1986, pp. 45–6; McGrath 2021, pp. 31–7). Disagreement is also an entrenched feature of contemporary discourse about gender: people disagree reflectively about gender terms, this disagreement persists even when the empirical facts are settled, and the various parties want to maintain that their answers are right. People who deny that Hunter Schafer is a woman aren't merely using words incorrectly, they are substantively in error about what is required for womanhood. While semantic

¹³ It also captures a number of the touted advantages of expressivism about gender, McGrath 2021, p. 30.

theories cannot capture this datum, in the interpretivist picture, it is easy to make sense of disagreement. If an informed interlocutor sincerely disagrees with me about Hunter's gender, then I must conclude that they are offering a rival interpretation of the practice. Moreover, it is easy to see why such disagreements might arise. In interpreting womanhood, for instance, we naturally face conflicting pressures to emphasize both the differences and the similarities between women. (Highlighting similarities helps to analyze oppression and develop resistance strategies; highlighting differences safeguards against essentialism.) The result is an ongoing process of contestation, in which new interpretations supersede older views and the various competing pressures pull our conceptions of womanhood in different directions.¹⁴

A second advantage of the interpretivist model is that it makes space for social facts to matter, while allowing us to advance moral claims about the scope of the relevant practice. In Dworkin's picture, we are interpreting the enormous body of legal precedent, statute, and the like when we make claims about the content of the law. Likewise gender does not float free of social context: upon our arrival into the world we confront both the facts of anatomy and the millennia of social practice that have been built thereon. In both cases, we are interpreting some pre-existing practice whose shape constrains the kinds of interpretations that can be constructed.¹⁵ According to interpretivists, though, while social facts matter in fixing the scope of an interpretive kind, they do not wholly determine it. Our accounts need not be prisoners of the status quo where that is unjust; while our interpretations are not unconstrained, there *is* scope to argue for more inclusive accounts of gender, for instance.

Finally, the interpretivist account seems to fit well with the way gender *actually* functions in societies in at least two senses. First interpretivism does not just allow us to explain good faith disagreement about gender, but it also *predicts* that gender will be contested given the open-ended nature of the interpretive process. It is thus some evidence in support of interpretivism that gender

¹⁴ Frye 1996, p. 1002 calls for the construction of a "concrete and historically real positive category of women—deliberately, creatively elaborating and articulating the differences among women in, by, and as a means to constructing a sociality, a symbolic order, a web of meanings of and among women". On the face of it, the interpretivist model looks like a strong candidate for what an implementation of this ideal of womanhood not as an "other" but as a site for contestation might look like.

¹⁵ For this reason our account does not face problems in the "twin-earth" case discussed in George and Briggs 2019; incorporating a historical dimension at the level of *categories* draws the sting of this case.

has been realized in different ways in different times and places. In historical perspective, gender turns out to be continually in flux, repeatedly redefined in response to the vagaries of historical context (Scott 1986, p. 1074).¹⁶ Second the interpretivist account gives us one way to make sense of the idea that gender is socially constructed. Although it remains controversial what exactly “social construction” amounts to (see Ásta 2018; Hacking 1999; Haslanger 2012c), the interpretivist picture does realize at least one of its likely hallmarks. According to Ian Hacking social kinds are characteristically “interactive” in that the act of classification and the nature of the classified things are mutually and reciprocally interdependent (Hacking 1999). This is clearly true of interpretive kinds: changes to the shape of some social practice affect the self-conceptions and thus the interpretive thinking of the people whose lives are enmeshed in said practice.

These three points give us some reason to, at the very least, try the interpretivist account of gender on for size. In and of itself this does not immediately translate into a first-order account of gender and it is not intended to do so. Interpretivist kinds can be interpreted in many ways. Nonetheless, since part of what we are looking for is a realist account that avoids oppressiveness, interpretivism needs to be friendly to non-exclusionary accounts of gender. In §4.2, I show that interpretivism can meet this demand, by sketching an inclusive interpretation that has some similarities to the view presented in George and Briggs 2019.

Before we can consider the question of inclusiveness, though, we should say something about how the interpretivist picture avoids the trap set by the squeezing argument. Interpretive kinds are *not* natural kinds. The difference between legal and non-legal systems, for instance, does not seem to inhere in the intrinsic or essential properties of the relevant systems, doesn’t obviously support inductive reasoning, and doesn’t seem to carve nature at its joints. At the same, time, however, interpretive kinds allow us to insist that the *real* facts about gender in our society diverge from the operative gender kinds. Consider an analogy: a government body *g* adopts what looks like a discriminatory practice *p*. While there exists laws against discrimination, the policymakers in *g* maintain that *p* does not contravene these laws. The case comes to trial in court *c*. It is

¹⁶ There is a slight tension here with the minimal realist’s emphasis on the “universality” of gender kinds. I do not think this is fatal. Against the relativist, the minimal realist rejects intra-societal relativism about gender, but they acknowledge that there may be temporal and cultural variation in the scope of gender kinds. I envisage this to imply that there are, for example, universal gender kinds within the Western world.

decided that p is legal. But this judgment is appealed: c , the lawyers argue, got the judgment *wrong*. Their appeal is successful, and the initial verdict is overturned. p was in fact illegal all along and c was wrong about the law. The plaintiffs are compensated for a violation of their civil rights that occurred long before any legal judgment was made. Notice that in this analogy both societal practice and the initial judgment in c were inconsistent with the law. Even though this fact was not respected by the various operative mechanisms, p was illegal from the very beginning. Analogously, a person could really belong to a gender kind even if the operative mechanisms that prevail in their society fail to recognize this.

Of course, gender kinds are very different from the law. There typically exist legal mechanisms that allow the system to arbitrate rival interpretations and to officially settle questions about the law's content and scope. No such mechanisms exist in the case of gender. Interpretation is more open-ended and fluid. Nonetheless, just as it is possible to argue that someone has *misinterpreted* a play, for instance, we can affirm that our preferred interpretations of gender kinds are the right ones and give arguments in defense of this conclusion. In particular, it is both consistent and not improbable that a). our gender kinds could have a particular extension, but b). operative social practices could massively fail to accord with the relevant extension. Indeed, this is just what happens when ontological oppression takes the form of failing to recognize gender kinds that ought to secure social uptake. We get, then, to say that Hunter Schafer is a woman and was always a woman no matter whether or not society recognized this. In doing so, we uphold the spirit and the letter of MR without falling into the trap set by the squeezing argument.

4.2 Abolitionism, inclusion, and the value of gender

In this section I aim to show how it might be possible to develop an inclusive interpretation of gender. To give a full account, we would first need to consider what function a system of gender could serve that might justify its existence. I think there are answers that could be given here, but discussion (which invites us to consider the merits of gender abolitionism) takes us far beyond the parameters of this paper. For that reason in what follows I have simply *assumed* that gender structures could have a positive role in our society.

This helps, to explain how an inclusive interpretation of gender might be possible. We start in a culture in which meanings are attached to sex in such a way as to define gender kinds. In virtue of this, people who share a gender kind have a common set of experiences and have a shared interest in making sense of these experiences. To do so, they interpret and reinterpret the gender kinds to which they belong, sustaining and rewriting the cultural meanings inscribed by the gendered social structure as a whole.¹⁷ These acts of interpretation are guided by the shared goal of making sense of their experiences, but they also seek to show this practice in the best possible light: *ceteris paribus* we want to make the best of our lives.

A concrete example helps to make the point. Imagine we are concerned to interpret “womanhood”: where possible, we prefer inclusive conceptions of “woman” since those show the kind as solidaristic and compassionate. At the same time, we should avoid interpretations that would make womanhood oppressive. We therefore favor an interpretation that accords people substantial agency in affiliating or disaffiliating themselves from gender kinds and deprecate interpretations that trap people in gender kinds with which they don’t identify; in other words, we interpret gender in such a way as to recognize individuals’ first-person authority over their gender classification (cf. Bettcher 2014). This same permissiveness favors allowing for a rich variety of genderqueer identities, beyond those normally encompassed in the gender binary. These considerations also recommend allowing people to opt out of a gender-system when it is not meaningful for them and thus support interpretations that include an agender kind.¹⁸

At this point it might be suggested that what we have described just is gender abolitionism, our intentions notwithstanding (Earp 2021, p. 39). But this is a mistake: while individuals are indeed accorded maximal freedom to negotiate the relevant social space in this picture, they nonetheless enter into a world in which sex comes with certain social meanings. Moreover, since interpretivist kinds are historically extended practices “anchored” to their origins in sex kinds, this vision can count as a gendered world, vast differences to our world notwithstanding (cf. George and Briggs 2019, pp. 21–3).

¹⁷ Under conditions of patriarchy the interpretations of the oppressed groups are substantially coerced, but we can imagine that they could be otherwise.

¹⁸ Cf. George and Briggs 2019, p. 29 especially.

5 Conclusion

I have argued that there *are* considerations favoring deflationism: the demands of normativity and objectivity are hard to reconcile. But I have also suggested that construing gender kinds as interpretive kinds can give us the requisite reconciliation and offers a better account of the wrong of misgendering. I suggest, then, that we try interpretivist versions of minimal realism on for size.

References

- Ásta (2013). "The Social Construction of Human Kinds". *Hypatia* 28.4, pp. 716–732.
- (2018). *Categories we Live By: The construction of sex, gender, race, and other social categories*. Oxford: UP.
- Bach, Theodore (2012). "Gender is a Natural Kind with a Historical Essence". *Ethics* 122.2, pp. 231–272.
- (2016). "Social Categories are Natural Kinds, not Objective Types (and Why it Matters Politically)". *Journal of Social Ontology* 2.2, pp. 177–201.
- Barnes, Elizabeth (2014). "Going Beyond the Fundamental: Feminism in Contemporary Metaphysics". *Proceedings of the Aristotelian Society* 114.3, pp. 335–351.
- (2020). "Gender and Gender Terms". *Noûs* 54.3, pp. 704–730.
- (2022). "Gender without Gender Identity: Identity-Based Views of Gender and the Problem of Cognitive Disability". *Mind* 131.523, pp. 836–862.
- Bettcher, Talia Mae (2009). "Trans Identities and First-Person Authority". In Laurie J. Shrage (ed.), *"You've Changed": Sex Reassignment and Personal Identity*. Oxford: UP.
- (2013). "Trans Women and the Meaning of "Woman"". In Alan Soble, Nicholas P. Power, and Raja Halwani (eds.), *Philosophy of Sex: Contemporary Readings*. 6th edition. Lanham, MD: Rowman & Littlefield.
- (2014). "Trapped in the Wrong Theory: Rethinking Trans Oppression and Resistance". *Signs* 39.2, pp. 383–406.
- Burgess, John P. (2008). "Being Explained Away". In *Mathematics, Models, and Modality: Selected Philosophical Essays*. Cambridge: UP, pp. 85–103.
- Dembroff, Robin (2018). "Real Talk on the Metaphysics of Gender". *Philosophical Topics* 46.2, pp. 21–50.
- (2020). "Beyond Binary: Genderqueer as Critical Gender Kind". *Philosophers' Imprint* 20.9, pp. 1–23.
- Dembroff, Robin and Wodak, Daniel (2018). "He/She/They/Ze". *Ergo* 5.14, pp. 371–406.
- Diaz-Leon, Esa (2016). "Woman as a Politically Significant Term: A Solution to the Puzzle". *Hypatia* 30.1, pp. 245–258.
- Dworkin, Ronald (1986). *Law's Empire*. Cambridge, MA.: Belknap Press.
- (1996). "Objectivity and Truth: You'd Better Believe it". *Philosophy & Public Affairs* 25.2, pp. 87–139.
- Earp, Brian D. (2021). "Abolishing Gender". In David Edmonds (ed.), *Future Morality*. Oxford: UP.

- Foucault, Michel (1978). "About the Concept of the "Dangerous Individual" in 19th Century Legal Psychiatry". *International Journal of Law and Psychiatry* 1.1. trans. Alain Baudot and Jane Couchman, pp. 1–18.
- Fricker, Miranda (2007). *Epistemic Injustice: Power and Ethics of Knowing*. Oxford: Oxford University Press.
- Frye, Marilyn (1996). "The Necessity of Differences: Constructing a Positive Category of Women". *Signs* 21.4, pp. 991–1010.
- George, B. R. and Briggs, R. A. (2019). "Science Fiction Double Feature: Trans Liberation on Twin Earth". MS. URL: <https://philpapers.org/archive/GEOSFD.pdf> (visited on 04/15/2022).
- Gibbard, Allan (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA.: Harvard UP.
- Hacking, Ian (1999). *Social Construction of What?* Cambridge, MA: Harvard UP.
- Haslanger, Sally (2012a). "Gender and Race: (What) Are they? (What) Do we want them to be?" In *Resisting Reality: Social Construction and Social Critique*. Oxford: UP, pp. 221–247.
- (2012b). "Ontology and Social Construction". In *Resisting Reality: Social Construction and Social Critique*. Oxford: UP, pp. 83–112.
- (2012c). "Social Construction: The 'Debunking' Project". In *Resisting Reality: Social Construction and Social Critique*. Oxford: UP, pp. 113–138.
- Jenkins, Katharine (2016). "Amelioration and Inclusion: Gender identity and the concept of Woman". *Ethics* 126.2, pp. 394–421.
- (2023). *Ontology and Oppression: Race, Gender, and Social Reality*. Oxford: UP.
- Kripke, Saul A. (1980). *Naming and Necessity*. Cambridge, MA.
- McGrath, Sarah (2021). "The Metaethics of Gender". In Russ Shaffer-Landau (ed.), *Oxford Studies in Metaethics*. Oxford: UP.
- Mikkola, Mari (2006). "Elizabeth Spelman, Gender Realism, and Women". *Hypatia* 21.4, pp. 77–96.
- (2016). *The Wrong of Injustice: The Role of Dehumanization in Feminist Philosophy*. Oxford: UP.
- Rawls, John (1971). *A Theory of Justice*. Cambridge, MA.: The Belknap Press.
- Rosen, Gideon (2017). "Metaphysical Relations in Metaethics". In Tristram McPherson and David Plunkett (eds.), *The Routledge Handbook of Metaethics*. New York: Routledge.
- Saul, Jennifer (2012). "Politically Significant Terms and the Philosophy of Language: Methodological Issues". In S. L. Crasnow and A. M. Superson (eds.), *Out from the Shadows: Analytical Feminist Contributions to Traditional Philosophy*. Oxford: UP.
- Scheman, Naomi (1997). "Queering the Center by Centering the Queer: Reflections on Transsexuals and Secular Jews". In Diana Tietjens Meyer (ed.), *Feminists Rethink the Self*. Boulder, CO.: Westview Press, pp. 124–162.
- Scott, Joan W. (1986). "Gender: A Useful Category of Historical Analysis". *The American Historical Review* 91.5, pp. 1053–1075.
- Serano, Julia (2016). *Whipping Girl: A transsexual woman on sexism and the scapegoating of femininity*. second edn. Berkeley, CA.
- Stoljar, Natalie (2011). "Different Women. Gender and the Realism-Nominalism Debate". In Charlotte Witt (ed.), *Feminist Metaphysics: Essays in the Ontology of Sex, Gender and the Self*. Dordrecht: Springer.
- Williams, Bernard (1981). "Persons, character and morality". In *Moral Luck: Philosophical Papers 1973–1980*. Cambridge: UP.