

## Gender and Gender Terms

Philosophical theories of gender are typically understood as theories of what it is to be a woman, a man, a nonbinary person, and so on. In this paper, I argue that this is a mistake. There's good reason to suppose that our best philosophical theory of gender might not directly match up to or give the extensions of ordinary gender categories like 'woman'.<sup>1</sup>

To illustrate this, I offer a defense of *social position* theories of gender - theories which say that gender is a social system that privileges some and disadvantages others based on expectations about what social roles people should occupy in virtue of their perceived sex characteristics. A major objection to such theories of gender is that they don't give us an adequate account of who should count as women (or men, or genderqueer people, or etc). I agree that such theories *don't* give an adequate account of gender terms like 'woman', nor do they give us necessary and sufficient conditions for what it is to be a woman (or nonbinary, or a man, etc). But I argue that it's a mistake to expect them to. The project of developing a philosophical theory of gender can and should come apart from the project of giving definitions or truth conditions for sentences involving our gender terms. A theory of the social reality that explains gender might guide or shape how we use our gender terms without thereby being a theory of those terms, or a theory that straightforwardly gives us the extensions of those terms.

In (§1), I introduce the idea of the metaphysics of gender. In (§2) and (§3), I sketch the two main strands in contemporary philosophical accounts of gender - social position accounts and identity-based accounts - and then discuss a common worry for each: that they fail to correctly determine the class of people who should count as women (or men, or non-binary, etc). In (§4)-(§6), I articulate an alternative interpretation of Haslanger's social position theory of gender, according to which it's a mistake to think that such a social position theory is in the business of giving the extension for terms like 'woman'. I show how this interpretation of Haslanger's view avoids the objection that her view fails to correctly classify all women as women. Finally, in (§8) and (§9) I sketch a way in which we can understand our use of gender terms as deeply influenced by our theory of gender, without thinking that a correct theory of gender is thereby a theory of our ordinary gender terms or concepts.

---

### 1. A metaphysics for gender

---

<sup>1</sup> A note on terminology: there's some confusion in the relevant literature over whether the target of analysis should be ordinary language terms like 'woman', our concept of *woman*, or perhaps a social category *being a woman*. Indeed, sometimes the target seems to be all three. And typically where focus is on the latter two the idea is that this is the underlying social reality which our use of terms like 'woman' ought to track. (See especially Haslanger (2012)c.) I'm going to primarily focus on gender terms like 'woman', but my arguments about social position accounts and exclusion problems apply equally well if you instead think the target should be the concept of woman or the relevant social category or etc.

Those who attempt to give a metaphysics of gender face skepticism on two fronts. On one side, many people interested in metaphysics do not think that the interesting questions of metaphysics extend to gender. On the other side, many people interested in gender do not think that the interesting questions of gender extend to metaphysics.<sup>2</sup> And yet, such skepticism notwithstanding, a rich philosophical conversation on the metaphysics of gender has emerged and gained momentum in recent years.

Giving a metaphysics of gender is typically understood as the project of explaining what gender really is. The oft-repeated slogan is that gender is ‘the social meaning of sex’ - but this claim can be unpacked in a huge variety of ways. We might have lots of (probably inconsistent) folk beliefs about gender. We might use gendered language and gendered concepts in a variety of ways - some of them synonymous with sex, others clearly not. And so on. But when we’re engaged in the metaphysics of gender, we’re trying to theorize what it is in virtue of which people have genders, or in virtue of which members of a given gender can be said to have something in common with each other, or in virtue of which gendered norms and roles have the significance they do. That is, we’re trying to say what feature(s) of the world - if any - unify or explain gender.<sup>3</sup>

This project is often understood as at least in part a political, rather than purely descriptive. When explaining what gender really is, we can understand ourselves as attempting to engage in a ‘debunking project’ - an attempt to explain why commonsense beliefs about gender are false, confused, or misguided, often in ways that contribute to gender inequality. Similarly, when we are trying to figure out who is including in gender categories like *woman* or how we should use the term ‘woman’, we may in part be asking how we *should* use these terms and understand these categories (an ‘ameliorative’ analysis), rather than simply asking how such categories and terms are in fact used in their contemporary contexts (a ‘descriptive’ analysis).<sup>4</sup> We’ll return to this issue in (§6).

The options for metaphysical explanations of gender run the gamut from strongly deflationist or response-dependent all the way to naturalist<sup>5</sup> or essentialist.<sup>6</sup> What those attempting to give a metaphysics of gender have in common is not *how* they think gender should be explained, but rather the simple idea that gender is something that admits of metaphysical explanation. Skeptics about this project think that gender just isn’t the kind

---

<sup>2</sup> For a critical discussion and overview of some of this skepticism, see Alcoff (2005), ch. 6.

<sup>3</sup> Haslanger (2012)a characterizes this project as an attempt to ‘explain a variety of connected phenomena in terms of their relations to one that is theorized as the central or core phenomenon.’

<sup>4</sup> See Haslanger (2012)c.

<sup>5</sup> The most familiar way, historically, to defend the idea that gender is a natural kind is by denying a sex/gender distinction. But for a more nuanced defense of gender as a natural kind see Bach (2012), (2016).

<sup>6</sup> Gender essentialism has a philosophical history that tracks back at least as far as Aristotle, though it’s fair to say that most contemporary philosophical discussion of gender are strongly anti-essentialist. For an interesting argument that gender essentialism and social constructionism are compatible however, see Witt (2011).

of thing that admits of such explanation - perhaps what we call 'gender' is nothing more than a loose collection of words and beliefs<sup>7</sup>, or perhaps gender is not any one thing, but rather how we choose to 'perform' in response to sex-based stereotypes<sup>8</sup>, or perhaps it's just implausible that there's any sense of gender that is stable across differences in race, class, sexuality, and culture.<sup>9</sup>

A striking feature of the contemporary metaphysics of gender, though, is that it typically takes the task of explaining gender as the task of explaining what it is to be a woman (or a man, or genderqueer, or etc.) And thus attempts to give a metaphysics of gender often become attempts to give application conditions for gender terms such as 'woman'. Likewise, skepticism about the metaphysics of gender often arises from skepticism that there is any specific thing that it is to be a woman, or a to be man, or to be genderqueer, etc.

In what follows, I'm going to suggest that this is a place where philosophical analyses of gender have taken a bit of a wrong turn. And in doing so, they've gotten mired in debates that are ultimately red herrings. Giving a metaphysics of gender, I argue, should not be thought of as an attempt to give a metaphysics that directly accounts for the extension of the terms 'woman' and 'man'. And so giving a metaphysics of gender needn't be the project of giving a metaphysics that determines which people count as women, which people count as men, which people count as nonbinary, etc. Rather, giving a metaphysics of gender should be understood as the project of theorizing what it is - if anything - about the social world that ultimately explains gender. But that project might come apart from the project of defining or giving application conditions for our natural language gender terms like 'woman'.

---

## 2. Social position accounts

Contemporary gender metaphysics can be (roughly) divided into two main camps: social position accounts and identity-based accounts.<sup>10</sup> Broadly speaking, social position accounts explain gender by external factors. An individual's gender is a matter of how other people react to them, treat them, etc. More specifically, many social position accounts ex-

---

<sup>7</sup> This is, perhaps, the implicit consequence of some contemporary views in metaphysics which prioritize fundamentality (Barnes (2015)). But see also LaBrada (2016) for a discussion of gender eliminativism.

<sup>8</sup> As in, e.g., Butler (1990)

<sup>9</sup> See especially Spellman (1988)

<sup>10</sup> Note that while the dichotomy I'm presenting here distinguishes two main strands of discussion, it is not exhaustive. Stoljar (2011), for example, argues for a resemblance-nominalism about gender which includes both external social factors and internal sense of gender identity. And Briggs and George (manuscript) argue that gender categories are constructed from external, historical social factors, but that whether someone is a member of a particular gender is a question of whether they *should* be classed as a member of that gender. Briggs and George state that this moral question is often, though not always, determined at least in part by gender identity. Neither of these views fits easily on either side of the distinction I'm discussing here.

plain gender in terms of the material social (dis)advantage that is imposed on individuals based on collective norms and assumptions about sexed bodies. Explanation in terms of material (dis)advantage, however, isn't a necessary feature of such views - although it is a common one. For the purposes here, I'm treating the hallmark social position accounts of gender to simply be the claim that an individual's gender is determined by social factors external to that individual - how they are perceived, what roles they are expected to occupy, etc.<sup>11</sup>

There are many different ways of spelling out the details of this, some more inflationary than others. So, for example, on Ásta's (2011), (forthcoming) theory of gender, gender is a social property which is 'conferred' on an individual in a given context. To be a woman in a particular context is to occupy a certain social role - to have specific social 'constraints and enablements' in that context, and specific assumptions or expectations placed on you in that context. A specific gender property like 'woman' is conferred on an individual in a context by others in that context, and it's conferred based on assumptions about specific 'base properties' that being a woman is taken to track in that context. What those base properties are - as well as what the 'constraints and enablements' of gender properties are - can vary greatly depending on the context. In many situations, people confer gender based on assumptions about (perceived) biological sex or reproductive role. But in other contexts, people confer gender differently - they might confer it simply based on stated gender identity or pronoun preferences, they might confer it based on gender expression, or etc.

On the more inflationary end of the spectrum, Charlotte Witt (2011) argues that gender is a 'uni-essential' property of social individuals. Witt thinks that human beings are comprised of three components - the human organism, the person, and the social individual. Social individuals are defined in terms of their relationships to others, and the social roles they occupy. Professor, father, sister, caretaker, friend - these are all properties of the social individual. Gender, according to Witt, is a social property which structures and unifies our other key social properties and social roles. You are not merely a sibling, you are a brother or a sister; you are not merely a parent, you are a mother or a father; and so on. And gender is, on Witt's view, a social role imposed on persons based on assumptions about reproductive role. The social system of gender divides individuals into two gender categories, man and woman - and structures so much of our social life around these categories - in order to organize and regulate reproduction and reproductive labor, often at the expense of women.

Ásta and Witt disagree quite markedly on the underlying metaphysics. For Ásta, genders are social properties conferred by others in a given context, which one can gain or lose as one moves from context to context. For Witt, gender is the social role which structures and unifies all other key social roles, and which is thus essential to the social individual (though not to the person or the human organism). But what both agree on is that a per-

---

<sup>11</sup> As Linda Alcoff (2005) explains the idea: 'The external situation determines the person's relative position, just as the position of a pawn on a chessboard is safe or dangerous, powerful or weak, according to its relation to the other chess pieces' (p. 148.)'

son's gender is determined by others - by the complex perceptions, judgements, and norms that other people impose on that individual.

Perhaps the most well-known and widely-discussed version of social position theory, however - and the one which I will focus on as a paradigm example of a social position accounts of gender - is Sally Haslanger's. According to Haslanger (2012)a, a person, S, is a woman iff:

- (i) S is regularly and for the most part observed or imagined to have certain bodily features presumed to be evidence of a female's biological role in reproduction;
- (ii) That S has these features marks S within the dominant ideology of S's society as someone who ought to occupy certain kinds of social position that are in fact subordinate (and so motivates and justifies S's occupying such a position); and
- (iii) The fact that S satisfies (i) and (ii) plays a role in S's systematic subordination, i.e., along some dimension, S's social position is oppressive, and S's satisfying (i) and (ii) plays a role in that dimension of subordination.

And conversely, S is a man iff:

- (i) S is regularly and for the most part observed or imagined to have certain bodily features presumed to be evidence of a male's biological role in reproduction;
- (ii) that S has these features marks S within the dominant ideology of S's society as someone who ought to occupy certain kinds of social position that are in fact privileged (and so motivates and justifies S's occupying such a position); and
- (iii) the fact that S satisfies (i) and (ii) plays a role in S's systematic privilege, i.e., along some dimension, S's social position is privileged, and S's satisfying (i) and (ii) plays a role in that dimension of privilege.

Let's unpack this a little. The central idea of Haslanger's dual definitions of woman and man is this: gender is a social structure that privileges some and disadvantages others based on assumptions about biological sex. There's no particular intrinsic property or feature that all women have in common, on this view. Rather, what women have in common is a structural feature: social norms dictate that they ought to occupy certain social roles, based on our assumptions about (real or imagined) bodily features we associate with a female's role in reproduction. And the roles that we assume they ought to occupy are in fact devalued. What the particular roles are can and does vary across culture, class, race, nationality, etc. But on Haslanger's view, there are always things which we associate as 'women's things', 'women's work', 'women's behavior', etc. These roles are assigned based on presumptions about biological sex (and especially reproductive role). And then whatever varying things we consider the 'women's things' to be are things we in fact socially devalue.

But Haslanger's view - and other, similar social position accounts - faces what Katharine Jenkins (2016) calls an 'exclusion problem'. The worry here is straightforward. Social posi-

tion accounts often fail to classify people we think should count as women as women.<sup>12</sup> This problem is especially pressing in the case of those whose experience of gender or gender presentation departs from our gender norms, such as trans women who do not easily ‘pass’ as cis women. Because social position accounts say (roughly) that whether you are a woman is determined by how other people respond and react to you, they often end up saying that trans women who aren’t easily or typically identified by others as women aren’t women.

Let’s continue using Haslanger’s view as an example. Haslanger’s account classifies only some trans women as women. But those trans women who aren’t regularly identified as having female-associated body parts will not meet condition (i) in Haslanger’s definition of ‘woman’ - they are not readily and for the most part identified as having features associated with a female’s role in biological reproduction. Note that it is not necessary for a person to be perceived as biologically female in order to be a woman, on Haslanger’s account. It’s sufficient that they be perceived as having features associated with femaleness (and especially a female’s role in reproduction). Nevertheless, many trans women - including many of the most economically disadvantaged and vulnerable trans women - don’t meet this criteria, and many more who currently do meet this criteria considered themselves women before they met it.

And it’s easy to see how this worry generalizes to social position accounts more broadly. Any view that explains what it is to be a woman, or a man, or genderqueer, or etc in terms of how other people treat, react, and respond to you will have difficulty explaining cases where a woman is not recognized by others as a woman, a genderqueer person is not recognized by others as genderqueer, and so on. Such exclusion problems are taken to suggest that social position accounts are metaphysically inadequate - they don’t fully explain the reality of gender, or what it is to be a woman. Because they define gender in external terms - via how others respond to you and perceive you - they can’t fully explain the phenomenon of *misgendering* (when others simply get your gender wrong), and in doing so they undervalue the role of internal aspects of gender, such as gender identity. Moreover, the same worries are also taken to suggest that such accounts are *politically* inadequate as well. By maintaining that, e.g., trans women aren’t women unless they are viewed by others in a specific way (because *what it is* to be a woman is to be viewed in a specific way), such views contribute to the deep harms of misgendering that trans people face.<sup>13</sup>

---

<sup>12</sup> See especially Jenkins (2016), McKittrick (2015)

<sup>13</sup> See especially Kapusta (forthcoming). Some defenders of social position accounts have objected here that part of their view is that systems of gender are unjust. So they are not saying that trans women shouldn’t be women - in a fairer system they would be - but merely that given the way the world is and the way that gender actually functions, you are not a woman unless you are socially positioned in a specific way. But as Watson (2015) points out, it’s hard to argue that there’s some specific way that all non-trans women are socially positioned which differentiates their experience from those of trans women.

---

### 3. Identity-based accounts

In contrast to social position accounts, according to which gender is explained by external features, internalist accounts explain gender, at least in part, via *internal* features. What gender you are is determined by how you feel about yourself, how you are inclined to behave, which groups you see yourself as belonging to, etc, and not - or not entirely - by how others respond to you. And so it is possible, on internalist accounts, that people treat you as being a gender other than the gender you really are.

Just as with social position views, internalist views can take many different forms. So, for example, Jennifer McKittrick (2015), argues that a person is a woman (or a man, or genderqueer, or etc) just in case they have a cluster of specific behavioral dispositions. You are a woman in a particular context, for example, just in case you have a cluster of behavioral dispositions coded as feminine in that context - including, e.g., self-identification as a woman, using women-specific spaces like bathrooms and locker rooms, wearing your hair long, wearing women's clothing, etc. None of these dispositions is, by themselves, sufficient to make you a woman; having some sufficient number of them together is.<sup>14</sup>

A striking feature of many internalist views is a strong emphasis on *gender identity*.<sup>15</sup> I'll call these views 'identity-based' theories of gender. Not all internalist views are identity-based views in the sense I'll be discussing here (e.g., McKittrick's is not<sup>16</sup>), but many of the most prominent are. For example, Talia Bettcher (2009), (2013), argues that a person's gender is determined by their sincere self-identification with that gender. So you are a woman, on Bettcher's view, just in case you sincerely self-identify as a woman.<sup>17</sup> But 'sincere self-identification' means more than simply a willingness to assert 'I am a woman'. Those who sincerely self-identify as women will 'live as women' - with the caveat that living 'as a woman' can mean very different things to different people, and the authenticity of gender expres-

---

<sup>14</sup> Note that for McKittrick these behavioral dispositions are internal (they are dispositions to behave rather than manifested behaviors) and they can be masked by other factors - e.g., you can be disposed to self-identify as a woman, but refrain from doing so because of fear of violence. But on McKittrick's view gender is still socially constructed - and the dispositions, though internal, are not intrinsic - because *which* behavioral dispositions we count as feminine or masculine, and so which dispositions matter to what your gender is, is determined by social norms. For a related view, see Julia Serano (2007).

<sup>15</sup> Note that for the purposes here I will use 'gender identity' simply to mean self-identification with a particular gender (or no gender), rather than anything more specific. 'Gender identity' is sometimes used in psychology to refer to a specific psychological feature that emerges in early childhood, for example - but as I am using the term a person's gender identity could evolve or change as they age.

<sup>16</sup> For McKittrick, a disposition to self-identify with a gender is just one of a cluster of behavioral dispositions that matter to gender classification - so it is important, but not especially or uniquely important in the way characteristic of the views I'm labelling 'identity-based'.

<sup>17</sup> Importantly, Bettcher (2013) notes that this is not what 'woman' means in everyday contexts, but that it is the resistant meaning of 'woman' which is common in trans-inclusive subcultures. She then argues that we should accept and employ the resistant meaning of 'woman', given the exclusionary harms that the dominant usage perpetuates.

sion can sometimes be impeded by basic needs of safety or economic security.<sup>18</sup> The point is simply that gender identity, on Bettcher's view, is not merely a fact about what propositions a person is inclined to assert; it is something that fundamentally structures a person's experience and behavior.

Identity-based accounts can often sound, *prima facie*, like a version of essentialism. The claim that gender is (or is determined by) gender identity is easily interpreted as the claim that everyone has an innate sense of gender which is independent of how they behave, how they are socialized, how they are treated, etc. But identity-based accounts don't need to commit to anything like this.<sup>19</sup>

So, for example, Jenkins (2016) - drawing from Haslanger (2012)b - suggests that we should understand gender identity as a type of 'internal map': a pattern of responses, both conscious and unconscious, that guide social behavior. Jenkins characterizes gender identity as follows:

S has a gender identity of x iff S's internal 'map' is formed to guide someone classed as a member of X gender through the social or material realities that are, in that context, characteristic of Xs as a class.

And, more specifically:

S has a female gender identity iff S's internal 'map' is formed to guide someone classed as a woman through the social or material realities that are, in that context, characteristic of women as a class.

On this view, gender identity is not innate, essential, or independent of socially constructed norms about gendered behavior. Rather, we form gender identities in response to (contingent, socially constructed) gender norms, behaviors, and socialization. But one's sex doesn't determine how one forms gender identity. One can - in response to these contingent social features - form a gender identity that is more typically found in those with different reproductive organs, or form a gender identity that is uncommon, etc.

Identity-based accounts typically take gender identity to be a crucial part of what determines - or what *ought* to determine - the extension of our gender terms 'woman', 'man', 'genderqueer', 'agender', etc. This doesn't mean that, on such views, more public aspects of gender are unimportant. Bettcher, for example, emphasizes the importance of 'living

---

<sup>18</sup> See especially Bettcher (2007)

<sup>19</sup> To give perhaps the most anti-essentialist example, Appiah (1990) construes gender identity as nothing more than the direct response to gendered socialization. In contrast to Appiah, most contemporary accounts of gender identity will want to say that socialization doesn't directly determine gender identity, even if gender identity is in some important way determined by social factors, but there's a wide spectrum of views about how this works.



as' a woman<sup>20</sup> when saying that someone is a woman, and Jenkins emphasizes the importance of gender role, which she takes to be roughly what is described by externalist theories of gender like Haslanger's. But both Bettcher and Jenkins argue that we should use our term 'woman' to refer to all and only people who identify as women (or who have a female gender identity, to use Jenkins' terminology). And it is this characteristic feature of identity-based views - the idea that women are all and only the people who identify as women, and that we should use our term 'woman' to reflect this - that I want to focus on.

It's easy to see how views like this avoid the exclusion problems raised above. Regardless of whether they are perceived by others as women, trans women self-identify as women, and on identity-based accounts that's what matters. Because identity-based accounts view gender as something internally-determined, you can - contra social position accounts - be a woman and be the only one who realizes that, or be a woman despite the fact that others react to your gender presentation by telling you that you are lying or delusional.

Importantly, though, this approach also faces its own exclusion problems. It is doubtful, for example, that all cognitively disabled women have a sense of gender identity in the sense discussed.<sup>21</sup> But denying womanhood to cognitively disabled women seems like a gross injustice, especially given that cognitively disabled women are particularly vulnerable to gendered abuse. We would be denying that women without certain cognitive capacities deserve the label 'women', specifically because of how their cognitive capacities differ from what is typical. We would thus, in effect, be saying that, because of their disabilities, cognitively disabled women are not women, but are merely female.<sup>22</sup> This is not unlike how we attribute sex but not gender to non-human animals.

---

<sup>20</sup> Though it's important to note that 'living as' a woman is not the the same thing, for Bettcher, as occupying a public social position like the one described by Haslanger. There are, on her view, ways of living as a woman which are recognized in queer and trans-inclusive subcultures but which would not be recognized as ways of living as a woman in many other contexts.

<sup>21</sup> This point was, to my knowledge, first raised by Sally Haslanger in the PEA Soup discussion of Jenkins' paper: <http://peasoup.typepad.com/peasoup/2016/01/ethics-discussions-at-pea-soup-katharine-jenkins-amelioration-and-inclusion-gender-identity-and-the-.html>

Note that while her view places less emphasis on the role of gender identity and self-identification, the same exclusion problem arises for McKittrick's internalist account. On her view, a person must have 'some sufficient number' of 'sufficiently strong' behavioral dispositions to be a woman; this cluster typically includes dispositions toward self-identification, but it's at least *consistent* with McKittrick's view that if someone had sufficiently many other behavioral dispositions but did *not* identify as a woman, they might still be a woman. Nevertheless many cognitively disabled women will lack dispositions to behave in ways we typically code as female or feminine (given that cognitively disabled people are often disposed to behave in atypical ways), in addition to ostensibly lacking dispositions to self-identify with a specific gender, and so would seemingly not count as women on McKittrick's view.

<sup>22</sup> On Briggs and George (manuscript)'s view, gender should, ideally, be consensual, rather than imposed. But it is important to note that gender can never be fully consensual for everyone. If genders were only had consensually, then those who cannot consent to being gendered wouldn't have genders - but this withholding of gender would likewise be without their consent, and could potentially further marginalize them in our society.

Specific cases of exclusion aside, there are more general exclusion worries that arise for identity-based accounts. Suppose we treat a specific gender identity as a necessary — and perhaps sufficient<sup>23</sup> — condition for being a woman. Is there any sense of ‘gender identity’ such that all (and perhaps only) women have this same gender identity? Gender identity can’t merely be a disposition to state ‘I am a woman’ - you can be a woman with aphasia, or a woman with Alzheimer’s disease, or a woman with any number of conditions that would make it such that you lack such a disposition. Nor is the type of gender identity referred to in these conversations the same thing as the sense of gender identity that is often studied by psychologists - a type of sex-based identification that appears to form in very early childhood. Not all trans women, for example, form a female gender identity in this particular sense.<sup>24</sup>

So consider, instead, Jenkins’ proposal - that women are all and only those who share the same ‘internal map’. But here we encounter some worries. Do all the people we think of as women - trans women, intentionally gender non-conforming women, androgynous women, typically feminine women - have the same internal map? What does it mean for two people with very different experiences of their gender to have ‘the same’ internal map? At this point a proposal like Jenkins’ faces a dilemma: we can make the conditions for having a particular internal map specific and relatively strict, or we can make them fairly non-specific and vague. If we make the conditions for what it takes to have a specific type of internal map strict, we risk excluding some women. For example, we’d risk saying that especially gender non-conforming women aren’t women, or that neuro-atypical women aren’t women. Conversely, if we make the conditions for what it takes to have a specific type of internal map weaker, we risk incorrectly including people who are not women. So, for example, we risk saying that some genderqueer people are women, or even that some drag performers are women. The more specific we are about what it takes to have the right sort of internal map, the more we risk ruling some people out; but the less specific we are, the most we risk wrongly including people, which is itself a way of misgendering. The basic worry is simply that it’s very hard to characterize any internally felt sense of gender which all and only the women (or the men, or the genderqueer people, or etc) share.

---

<sup>23</sup> It’s a sufficient condition on Jenkins’ view - or at least it is a sufficient condition for being a member of the class which Jenkins’ thinks we should reserve the term ‘woman’ for. On Bettcher’s view I am less clear. Certainly sincere self-identification together with ‘living as’ a woman are jointly sufficient. And plausibly *sincere* self-identification (as opposed to simple self-ascription) is itself sufficient for living as a woman. Regardless, note that points similar to the over- and under-generalization worries that follow for gender identity can also be made for what should count as living as a woman. And the epistemic worries arise merely for treating a specific gender identity as a necessary condition for being a woman, regardless of whether it is sufficient.

<sup>24</sup> Indeed, there is significant resistance to the dominance of the ‘traditional trans narrative’, which insists that in order for a trans person’s gender identification to be legitimate they must make claims like ‘I’ve known my whole life that I was a boy’. See, for example, Stone (1992).

---

## 4. Metaphysics and language

Leaving aside debates over the metaphysics of gender for a moment, let's consider how the terrain looks in other areas of metaphysics. It's commonplace, within realist metaphysics more broadly, to maintain that one's metaphysics needn't line up neatly with the true sentences of natural language or the extensions of predicates. For 'There is a table' to be true, we don't obviously need a metaphysics that gives the real definition of tables, or the essence of tables, or etc. Our metaphysics can be one way, the true sentences of natural language another, and there needn't be any neat or direct mapping from one to the other.

Suppose we thought that the fundamental questions in metaphysics are just questions about microphysical entities; whatever these turn out to be, they won't include tables. Some people maintain that, if this is the case, 'There is a table' is false, though there's perhaps a paraphrase to something similar which is true, such as 'There are simples arranged table-wise'.<sup>25</sup> But it's increasingly common to deny that such paraphrase strategies are necessary, and likewise to deny that an absence of tables from our basic metaphysics means 'There is a table' is not a true sentence of the English language.

We can say instead that the truth of 'There is a table' is grounded in the existence and arrangement of microphysical entities, or that the existence and arrangement of microphysical entities are a truthmaker for 'There is a table', or 'There is a table' is true in virtue of the existence and arrangement of microphysical particles, or, etc.<sup>26</sup> There needn't, according to such views, be a direct equivalence between how we truly describe the world in natural language - e.g., as containing tables - and our metaphysics.

Those who endorse such views often argue that the truth conditions of natural language sentences are determined at least in large part by use. 'There is a table' is a true sentence of English simply because of our practices of communication and how we use our words. We reliably and successfully use our word 'table' in a way that renders 'There is a table' true. 'There is a table' would be a true sentence of English, on such a view, on pretty much any story about what fundamental metaphysics is like - it would be true even if tables were nothing more than collections of atoms in the void, or if tables were ultimately just bundles of tropes, or even if tables were ultimately just ideas in the mind of God.

Could we say something similar about the metaphysics of gender? For understandable reasons, this move has been resisted. Common usage of gender terms is currently in flux - much greater flux than terms like 'table' - so to begin with it's difficult to settle on what, if anything, 'the folk' or 'ordinary speakers' mean by their gender terms. Teenagers probably mean something quite a bit different from their grandparents; wealthy teenagers in Manhattan probably mean something quite a bit different from working class teenagers in

---

<sup>25</sup> See, for example, Van Inwagen (1990) and Merricks (2001)

<sup>26</sup> See, e.g., Bennett (2011), Cameron (2008), Dorr (2005), Schaffer (2009), Sider (2011)

Alabama. But suppose we narrow our focus to ordinary speakers in a specific context. Saying that, e.g., ‘x is a woman’ is true in a context just in case our ordinary usage of ‘woman’ in that context would include x doesn’t look appealing.

Again, the problem is one of exclusion. At least in many contexts - though this is beginning to change - our ordinary usage of gender terms seems to track perceptions about biological sex. So, for example, speakers will often reserve the use of gender terms like ‘woman’ for those they take to have certain biological features associated being female. People often refuse to grant that trans women are women, or that genderqueer people are genderqueer, etc. So the worry is this: saying that the correct usage of sentences involving gender terms is determined by use - the way that the correct usage of sentences involving ‘table’ is determined by how we use the term ‘table’ - leaves us unable to criticize the common use of gender terms the way we might want to. If I’m in a context where everyone is using the term ‘man’ in a way that only applies to people born with a penis and testicles, it looks troubling to grant that this usage is correct just because of how the term is employed in that context.

I’m sympathetic to the view that - barring skeptical hypotheses in which I’m just a brain in a vat - facts about tables aren’t the sort of thing people could be systematically wrong about, even if they could be systematically wrong about the metaphysical structure of the universe. With apologies to some metaphysicians, it doesn’t seem like we need to bring in a class of table experts to tell us the real truth about whether there are tables. But I do think people can be systematically wrong about gender. Moreover, I think that a theory of what gender really is ought to be able to influence how gender terms are used. It might not be surprising if our metaphysics of material objects doesn’t give us much guidance for how to use a word like ‘table’, but we want our metaphysics of gender to have at least *some* relevance to how we use words like ‘woman’.

But I think all of this is compatible with saying that social metaphysics doesn’t give us application conditions or definitions for gender terms. In standard analytic metaphysics, when we say that the truth and assertability of ‘There is a table’ is largely a separate matter from whatever metaphysics ultimately explains why there are tables, that’s typically because the metaphysical commitments we are talking about are pretty far removed from tables. We want an account of how we can say that reality is ultimately nothing more than atoms in the void, or states of affairs<sup>27</sup>, or basic qualitative ‘thisnesses’<sup>28</sup>, or Tractarian geometry<sup>29</sup>, or etc, in a way that explains the existence of a manifest image that includes tables, but which doesn’t require us to give illuminating truth conditions for ‘There is a table’ or provide a direct mapping from sentences involving tables to sentences involving our preferred esoteric piece of metaphysics.

---

<sup>27</sup> Armstrong (1978)

<sup>28</sup> Paul (2012)

<sup>29</sup> Turner (2016)

In contrast, when we are doing social metaphysics, the metaphysical commitments we're talking about - social structures, social identities, etc - are more closely related to natural language terms like 'man' and 'woman', even if they aren't the referents of those terms and even if they don't give us the application conditions for those terms. Unsurprisingly, the metaphysics of gender might be more directly relevant to how we speak about gender than a fundamental metaphysics of spacetime structure is relevant to how we speak about tables. But that's compatible with thinking that the metaphysics of gender doesn't directly give us specific or illuminating truth conditions for sentences involving natural language gender terms, any more than the metaphysics of fundamental spacetime structure gives us illuminating truth conditions for sentences about tables.<sup>30</sup>

In what follows, I'm going to argue that a social position account of gender can provide for a middle ground between two extremes - the view that says the correct metaphysics of gender straightforwardly gives us the application conditions for our gender terms, and the view that says that what's true about gender is determined simply by how people use natural language gender terms.<sup>31</sup> People can still be incorrect in their use of gender terms - incorrect because of the underlying metaphysical reality of gender - without the metaphysics of gender being seen as something which directly provides the definition of our gender terms.

---

## 5. Social position reconsidered

---

<sup>30</sup> Note that in saying this I'm not making any claim about whether or to what extent the interesting aspects of the metaphysics of gender are 'fundamental'. It's common, in wider debates in metaphysics, to hear claims like 'there is a table' is true, but tables aren't part of the fundamental structure of reality. Analogous claims for gender don't work very well when we're doing social metaphysics. Whatever the metaphysics of gender is, genders are unlikely to meet the criteria of 'fundamental' typically deployed in other areas of metaphysics - and we shouldn't expect them to. The claim here is not that there's a 'fundamental' structure to gender, but not to categories like 'woman'. Rather, the claim is simply that the most interesting or explanatory social categories don't map on neatly to our ordinary language terms and/or common-usage gender categories.

<sup>31</sup> It's worth noting that, as a result, some of the objections raised to views that try to separate metaphysical commitments from true sentences of natural language don't get off the ground for the type of view I'm defending here. For example, Amie Thomasson (2015), has argued against views which attempt to separate metaphysics from ordinary language via 'semantic ascent' - she claims that such views are committed to saying something equivalent to 'there are no tables, but 'there are tables' is true', and that a semantic argument can be given to show that this view is incoherent. Much can be said about this argument (I don't personally find it persuasive.) But the key thing to be said here is that the view I'm defending doesn't attempt to say anything like 'there are no women, but 'there are women' is true'. Of course there are women. And 'x is a woman' is true iff x is a woman. What I'm arguing is just that the metaphysical categories of interest - the bedrock social structures - are slightly different than 'woman', 'man', etc. Thomasson further argues for metaphysical deflationism, which certainly isn't in keeping with the approach I've taken here and which I don't have the space to discuss. But one thing to emphasize is that everything I'm saying is compatible with wild permissivism about social ontology: maybe infinitely many social categories exist, I'm just arguing that the ones we should consider as explanatorily most interesting are hierarchical social position categories.

Here is my central claim: on the best interpretation of social position accounts of gender, they shouldn't be thought of as giving us a metaphysical analysis of what it is to be a man or a woman, or of giving us straightforward application conditions for gender terms like 'man' and 'woman'. In what follows, I don't intend to give positive arguments for this view of gender. (That would be a far more expansive project.) Rather, I attempt to sketch of how it can be understood - what the basic idea is and how it differs from more traditional social position views - and then show how it can successfully avoid the most influential objection to social position accounts.

This spin on social position views is, in fact, in keeping with the approach to social position accounts that Haslanger herself undertakes. She maintains that her 'claim is not that my account [of gender] 'analyzes our concept', in the sense that it provides an interpretation of what people have in mind when they use the term, or that it is what determines the extension of gender/race language in a Fregean way, but that it captures the social reality that underlies our thinking and speaking, but is hidden from view.'<sup>32</sup> Similarly, she maintains that, having outlined the basic social categories she is interested in, 'we could simply bracket the terminological issues and just consider whether the groups in question are ones that are important to consider given the goals of our inquiry'.<sup>33</sup>

Indeed, within Haslanger's project 'woman' and 'man' are perhaps best understood as technical terms which aren't equivalent to the ordinary language terms. In much the same way, for example, 'part' is a technical term in discussions of mereology and doesn't correspond particularly well to the ordinary language term 'part'. Nevertheless, gender terms are politically loaded in a way that terms like 'part' are not, and there is a great deal of social significance attached to them. And using them (or perhaps, their homonyms) as technical terms is likewise politically loaded. People care deeply about the application of our gender terms and there's moral significance to their application. Haslanger originally argued that the political significance of our gender terms was a primary reason why we ought to use those (familiar) gender terms to refer to the (unfamiliar) social structures she describes. But I think that the direction that recent debates have taken - and specifically, how much they have focused on the use and meaning of words like 'woman' - is good reason to think that this move was misguided.

What I want to propose - using Haslanger's social position account as a template - is that we can decouple a social position metaphysics of gender from the definition or application conditions of our gender terms, but still maintain that such a metaphysics guides how we *ought* to use such terms.

Let's begin by a simple terminological replacement. Replace 'woman' in the Haslangerian definitions with 'feminized' and 'man' with 'masculinized', and we get the following results.

---

<sup>32</sup> Haslanger (2014), p. 29

<sup>33</sup> Haslanger (2012)a, p. 240

A person, S, is feminized in a context iff:

- (i) S is regularly and for the most part observed or imagined to have certain bodily features presumed to be evidence of a female's biological role in reproduction;
- (ii) That S has these features marks S within the dominant ideology of S's society as someone who ought to occupy certain kinds of social position that are in fact subordinate (and so motivates and justifies S's occupying such a position);
- (iii) The fact that S satisfies (i) and (ii) plays a role in S's systematic subordination, i.e., along some dimension, S's social position is oppressive, and S's satisfying (i) and (ii) plays a role in that dimension of subordination

In contrast, a person, S, is masculinized in a context iff:

- (i) S is regularly and for the most part observed or imagined to have certain bodily features presumed to be evidence of a male's biological role in reproduction;
- (ii) That S has these features marks S within the dominant ideology of S's society as someone who ought to occupy certain kinds of social position that are in fact privileged (and so motivates and justifies S's occupying such a position);
- (iii) The fact that S satisfies (i) and (ii) plays a role in S's systematic privilege, i.e., along some dimension, S's social position is privileged, and S's satisfying (i) and (ii) plays a role in that dimension of privilege

If we follow a Haslangerian interpretation of social structures, then being masculinized or feminized in a context will be the basic social reality that explains our complex social experience of gender. Our social world is structured in a way that codes bodies as male or female, assigns them (many and varied) roles as a result, and then systematically devalues the roles we think that those coded as female ought to occupy. For Haslanger, this is a basic fact about how the world is - as real and mind-independent as anything else. What's distinctive about this type of social reality, though, is that it's something made (and potentially something we could unmake) via our collective patterns of social interaction.

Saying that the social structure of masculinization and feminization is the ultimate metaphysical explanation of gender, however, needn't imply that it's the full story about gender, as Haslanger herself acknowledges.<sup>34</sup> Gender also encompasses gender identity, gender expression, and so on. And of course these things all matter greatly to our experience of gender, and saying that social position is the full or complete account of gender would be far too reductive.<sup>35</sup>

---

<sup>34</sup> 'Let me emphasize. . . that I do not want to argue that my proposals provide the *only* acceptable ways of to define. . . gender; in fact, the epistemological framework I employ is explicitly designed to allow for different definitions responding to different concerns.' Haslanger (2012)a, p. 221.

<sup>35</sup> It's also true, of course, that there are local subcultures - trans communities are a prime example - where gender functions very differently and non-hierarchically (see especially Bettcher (2013) and Ásta (forthcoming)). This is consistent with the Haslangerian picture of social structures - since the Haslangerian claim is about wider, systematic social norms - so long as these contexts are relatively localized and formed, at least in part, in response and resistance to hierarchical gender norms.

But a social position metaphysics allows us to say that these further components of gender can ultimately be explained in terms of the basic binary social structure that attributes social significance to perceived biological sex, and which privileges some and disadvantages others based on assumptions about what ought to follow from being perceived as male or female. Gender identity, for example, can be understood as an internally-felt sense of one's own relationship to social norms of gender formed in response to this basic social division. This is, for example, exactly the route that Jenkins takes in her account of gender identities as 'internal maps' formed in response to the norms and patterns of behavior that correlate to an underlying gendered social structure. But while the social division is binary, gender identity needn't be - we might form many and varied internal responses to being masculinized or feminized. Likewise, gender identity needn't correlate with one's social position. Most people who are feminized in a context identify as women, but some don't.

Gender expression can likewise be understood as a complex response to norms and assumptions about masculinization and feminization - you can consciously choose to embrace some, all, or none of the stereotypes associated with a social position you identify with or that others identify you with, for example. But what a social position account allows us to explain is why certain choices of expression have particular - and particularly gendered - significance. If someone perceived to have male sex characteristics wears makeup, it's socially significant (at least in many contemporary Western context.) Other socially significant choices of expression - having a lot of piercings or tattoos, e.g. - can be striking or unusual, but not in quite the same way. We have lots of behaviors we typically expect from - and impose on - people we perceive as male, and not wearing makeup is among them. We expect the people we view as male not to wear makeup partly *because* they're (perceived as) male. A social position theory can explain both why such gendered expectations are common and why, as a result, choices about gender expression are significant.

We might also want to identify additional social positions beyond masculinization and feminization. But we can do so while still understanding such further social positions as ultimately explained by - though distinct from - the basic binary positions. So, for example, we can explain further social positions in virtue of their relationship to unique ways one is masculinized or feminized in a context, or how one fails to be masculinized or feminized in a context.

Here are two examples of what we might consider further social positions, distinct from masculinization and feminization.

A person, S, is a *gender outlier* in a context C iff either:

- (i) S is perceived in C as having physical characteristics, real or imagined, associated with male's role in biological reproduction;
- (ii) S is perceived as attempting to occupy a feminized role in C
- (iii) That S satisfies (i) and (ii) plays a role in S's systematic subordination

or:



- (iv) S is perceived in C as having physical characteristics, real or imagined, associated with female's role in biological reproduction;
- (v) S is perceived as attempting to occupy a masculinized role in C
- (vi) That S satisfies (iv) and (v) plays a role in S's systematic subordination

A person, S, is a *gender confounder* in a context C iff

- (i) S cannot be reliably identified in C as having the physical characters we associate with either a male or female's role in biological reproduction, or S is reliably identified as having physical characteristics we associate with both a male and female's role in biological reproduction;
- (ii) Because S satisfies (i), S is neither masculinized nor feminized in C
- (iii) That S satisfies (i) and (ii) plays a role in S's systematic subordination

Basically, as I'm using the terms, you are a gender outlier if you're perceived as trying to occupy a specific gendered social position, but are also perceived as having sex-characteristics that we typically associate with a different gendered social position. Trans people who don't 'pass' as cis will be gender outliers, as will some cis people with very nonconformist gender expression. In contrast, you're a gender confounder if other people routinely can't figure out whether to respond to you as someone who is masculinized or someone who is feminized. Many genderqueer people are gender confounders in this sense, as again are some gender nonconforming people, regardless of their identification.<sup>36</sup>

I don't mean to defend these categories so much as to offer them as illustrations of the kinds of social positions - explained by but distinct from a basic binary - that might be worth exploring further. One of the advantages of a social position account of gender like Haslanger's is that it offers an explanation of why gender policing is so common and gender nonconformity is so heavily penalized. Our basic gender categories have a 'stay in your lane' requirement built into them: we assume that there are two ways that bodies can be (based on assumptions about reproductive role), and then mandate that there are ways you ought to behave, things you ought to identify with, ways you ought to express yourself based on being sorted into one of those two categories. If we think you're trying to occupy the 'wrong' social position, or we can't figure out what social position we think you ought to occupy, the social penalties are often severe.

The basic point I'm pushing is this. A social position account can say that the various aspects of gender are ultimately explained by a social structure that imposes norms and expectations (and which privileges some and disadvantages others) based on perceived biological sex and biological reproductive capacity. But a social position account isn't thereby committed to saying that such a social structure is everything there is to gender, or straightforwardly yields the extensions of our terms 'man' and 'woman', and so on. On

---

<sup>36</sup> See especially Watson (2016) for a discussion of how issues of misgendering and passing can commonly arise for gender nonconforming people who aren't trans.

this view, there is a bedrock social structure<sup>37</sup> that gives rise to the complicated, multi-faceted social experience of gender. When doing the metaphysics of gender, this basic social structure is something it makes sense to focus on. But it would be overly reductive to say that such a social structure is what gender is, or what gives us the extension of our gender terms. Gender is many, complicated things - but many, complicated things which are ultimately explained by a hierarchical social structure. The underlying structure of gender is binary, but that needn't mean that there are only two genders.

---

## 6. Exclusion redux

Perhaps the most influential objection to Haslanger's view of gender - and social position views of gender more broadly - has been the exclusion problem. But on my interpretation, discussions of the exclusion problem for Haslanger's view - including whether trans women are women, whether the Queen of England is a woman<sup>38</sup>, etc, can be sidestepped entirely. These type of worries have been raised repeatedly against Haslanger's account, and they focus on whether Haslanger's picture gives us an adequate explanation of who should fall under the extension of our term 'woman'. Trans women who don't pass as cis women aren't regularly and for the most part perceived as having features associated with a female's role in biological reproduction, but they're still women. So are gender-noconforming women who are regularly misgendered as men.<sup>39</sup> And it seems at least possible that some women don't occupy a social position that is in fact subordinate, while still being women - maybe the Queen of England isn't subordinated along any dimension, maybe stories about Amazons are still stories about women, even though they're stories that take place in a matriarchy, etc.

And so the exclusion worry is straightforward - but in my view not a deep problem with social position accounts. It's bad to say that trans women aren't women; it's weird for it to be debatable whether the Queen of England is a woman. But there's no problem with saying that some women (both cis and trans) aren't feminized in some contexts, especially since it's a part of Haslanger's view that this is a matter of injustice. And there's similarly no problem with saying the Queen of England or the Amazons might not be feminized in some contexts. On this reinterpretation of Haslanger, some women - both cis and trans - won't be feminized in some contexts, and some of the people who are feminized in some contexts - including some genderqueer people, some trans men, and some feminine cis men - won't be women.

---

<sup>37</sup> One that, I'd argue, exists in order to control and regulate sexual activity and reproduction.

<sup>38</sup> Mikkola (2009)

<sup>39</sup> Again, see Watson (2016) for discussion of the misgendering that can arise for gender-conforming people who are not trans.

In addition to exclusion worries, there have been other terminology-related objections to Haslanger's social position account. For example, there are general concerns that 'get rid of women' shouldn't be the goal of feminism, and that defining 'woman' in solely negative terms is unhelpful.<sup>40</sup> But again, there doesn't seem to be any problem with defining feminization in a context in purely negative terms, or with saying that we want to get rid of the social structure of feminization (where 'feminization' is a technical term).<sup>41</sup>

If I'm right, these problems aren't actually worries about the substance of social position accounts like Haslanger's. The problems are, instead, primarily ones of labelling. The issue isn't inadequacies of social position accounts per se, but rather with their inability to serve as proxies for or definitions of our gender terms. There's perhaps also a problem of not paying enough attention to other important aspects of gender - though these aspects needn't be part of the bedrock metaphysics of gender in order to be given due importance. Undervaluing the importance of various aspects of gender - perhaps especially gender identity - could be part of why a social position theorist like Haslanger was tempted to say that we should use the term 'woman' to refer to the social structure she describes because doing so will be politically effective. I don't think it has been politically effective, but that doesn't take away from the merits of a social position account itself - it just means we might want to rethink the terminology we use for such an account.

Haslanger's use of the terms 'woman' and 'man' is guided by her deeply influential notion of an *ameliorative project*. When we are engaged in an ameliorative project for x, according to Haslanger, we are not simply asking 'what is x?', we are asking 'what do we want x to be?' This can sound like an odd question for metaphysics. If I'm offering a theory of reality, the way I want it to be includes unicorns, an absence of paper cuts, and dogs that live forever. Sadly, wishing doesn't make it so. But 'what do we want x to be?' for social metaphysics shouldn't be interpreted as choose-your-own-adventure theorizing. Instead, the goal is to describe systems of injustice, with the hope that describing them can help us understand how to alter them. The 'what do we want x to be?' here is 'how do we want x to change?'

Importantly, though, we might want to separate the conceptual and semantic components of an ameliorative project from the metaphysical ones. Ostensibly, we want the social reality of gender to change and we want our gendered concepts and terminology to change. Haslanger's assumption was that the politically effective thing to do was combine these two projects: have gender terms refer to the underlying social structure she argues ulti-

---

<sup>40</sup> See especially Saul (2006)

<sup>41</sup> NB: saying that we want to get rid of the social structure of feminization - one which assumes there is a way a person ought to be or a role they ought to occupy based on perceptions of biological sex - doesn't mean that we want to get rid of all the roles or features we in fact think of as 'feminine'. Plenty of things we stereotype as feminine - empathy, care, nurture, etc - are valuable. The thought is simply that it would be better (for everyone, not just for women) if we didn't impose empathetic or care-giving roles - and then socially devalue such roles - based on perceptions of biological sex.

mately explains gender. But this isn't the only way to approach an ameliorative project for gender and gendered terminology.

---

## 8. Amelioration for gender terms: permissivism about self-ascription

I've argued that we should separate a social position metaphysics of gender from the application conditions of our ordinary language gender terms. But that leaves us with the gaping question of how to think about our gender terms. The advantage of tying the definition or application conditions of ordinary language gender terms to social metaphysics is that it allows for a type of externalism about the meaning of those terms.<sup>42</sup> It lets us say that ordinary speakers are just wrong when, e.g., they say that you have to have a vagina to be a woman. If we let go of the idea that the underlying social reality of gender - whatever it may be - yields straightforward application conditions for gender terms, then how should we think about the meaning of such terms? In what follows, I sketch an account of how an ameliorative approach to gender terms could be guided by a metaphysics of gender - in this case, a social position metaphysics - without that metaphysics being a theory of what it is to be a woman, a man, a genderfluid person, a genderqueer person, etc. I begin with an account of the self-ascription of gender terms, and then move on to discuss the use of gender terms more broadly.

I'm going to start from the assumption that there is probably no one thing that terms like 'woman', 'man', 'genderqueer' or 'genderfluid' mean. This isn't a unique feature of gender terms, of course. Many of our ordinary language terms are flexible, malleable, and can mean different things in different contexts. Many, maybe most, speakers probably take themselves to be using terms like 'woman' and 'man' as sex terms<sup>43</sup>, but I'm going to assume a basic level of externalism - speaker intention doesn't determine speaker mean-

---

<sup>42</sup> Haslanger (2012)e

<sup>43</sup> Bettcher (2013), argues that ordinary speakers - rather than using gender terms as sex terms - use gender terms specifically as terms to denote genital status. So 'man', in the dominant usage, means 'person with a penis and testicles' and 'woman' means 'person with a vulva and vagina'. But I don't think this is quite right. Ordinary speakers seem happy, for example, to say that the character Varys in *Game of Thrones* is a man and use male pronouns to refer to him, even though Varys is a eunuch (and has been since early childhood.)

ing, and the two can come apart.<sup>44</sup> The prevalence of intersex conditions seems to be enough to show that our gender terms are not simple synonyms for biological sex terms - even if ordinary speakers often take them to be. Research increasingly shows a spectrum of sex variation between the male and female binaries. But ordinary speakers seem happy to attribute terms like 'man' or 'woman' to people with various intersex conditions, so long as their gender expression and presentation is binary and has been consistent throughout their life.

The starting point which seems most plausible to me is that our gender terms are complex, messy, and often refer to a gerrymandered cluster of features - including sex, perceived sex, gender identity, gender expression, etc. They can and do mean different things in different contexts. I don't think there's any one uniquely correct definition of terms like 'woman', and definitional projects which seek to give necessary and sufficient conditions for the application of terms like 'woman' seem doomed to counterexamples.<sup>45</sup>

This view of gender terms is broadly similar to (and inspired by) the contextualist account of gender terms outlined - though not endorsed - in Saul (2012). On the view Saul describes, gender terms are contextually variable, and there is no unique class of people they refer to or feature they pick out. They can and do mean different things in different contexts. I'm sympathetic to the view Saul describes, though for the purposes here I'm

---

<sup>44</sup> There's a worry, though, about assuming this type of externalism. If Haslangerian structures exist, might they serve as something like a reference magnet for the use of our gender terms? (That is, might a realist social metaphysics end up forcing us to say that gender terms - despite appearances - refer to whatever the underlying social reality is?) I'm not too concerned about this point. A typical story about reference magnets says that the way the world is can *sometimes* trump use, or be a tiebreaker when use is ambiguous. But use still matters. And our use of gender terms is so strikingly at odds with with Haslangerian structures that I don't think there should be much concern that the mere existence of Haslangerian structures might trump that usage. As already discussed, speakers of English wouldn't say that in trying to achieve gender justice we are trying to achieve a society without any women, or that stories about Amazons are not stories about women, or etc. But these are all consequences of Haslanger's technical use of 'woman'.

<sup>45</sup> To be clear, I don't think this is unique to gender terms. Attempts to give necessary and sufficient conditions for the correct application of most *any* ordinary language term is unlikely to be a productive project, in my view. (See especially Sider (2011) for compelling discussion on this point.) Where I think gender terms like 'woman' might be interestingly different from terms like 'table' is simply that the underlying social reality of gender might more directly constrain what we think is the correct usage of 'woman' than our views about the underlying reality of physical objects constrains our usage of 'table'.

not wedded to the specifics of a contextualist account.<sup>46</sup> All I really need is for gender terms to be somewhat variable or flexible in which class of people they pick out.

People who violate our gender-normative expectations in some way - either by identifying as a binary gender different than the one assigned to them at birth, or by identifying as a non-binary gender, or etc - are often told that they aren't *really* the gender that they say they are, where this is meant to be some deep claim about reality. But on the view I'm defending, such claims don't make sense. There's something that it is to really be feminized in a context. That underlying social structure exists and there are facts about our relationship to it. There may well be other social positions - being a gender outlier, being a gender confounder, etc - grounded in this basic structure, and if so there are also facts about individuals' relationships to those social positions. But none of these give us application conditions for terms like 'man', 'woman', 'genderqueer', etc. There aren't any context- or language-independent facts about what it really is to be a woman, a man, a non-binary person, etc. There are just the multiple, imprecise, inconsistent, muddled ways we used those words in various contexts.

On this understanding of gender terms, there will doubtless be some contexts or ways of using the word 'woman' in which - strictly speaking - it's true that, for example, someone is a woman if and only if they were born with a vagina and ovaries. But there are also many other contexts and many other ways of interpreting the term - interpretations according to which identifying as a woman is sufficient to be a woman, contexts in which presenting as a woman is sufficient to be a woman, etc. Likewise, there are contexts in which identifying as neither a man or a woman is sufficient to make it the case that you are neither, interpretations according to which identifying as an x, for whatever gender term x, is sufficient to make it the case that you are an x, and so on.

If I'm right, there aren't any deep, language-independent facts about which people are women, which people are genderqueer, etc. But these terms - even if they are not metaphysically significant - are deeply politically and personally significant. Misgendering people (i.e., refusing to use a person's preferred pronouns and gender terms) is a major source of harm for trans and non-binary people. By misgendering someone, we refuse to acknowledge an intensely felt aspect of their identity. In the process, we also reinforce various harmful and false beliefs about the gender binary - that people are somehow supposed to accept the gender they were assigned at birth, that there is some robust fact about whether someone is really a man or woman, that gender always corresponds to sex, that there are only two ways to experience gender identity (as a man or as a woman), etc.

---

<sup>46</sup> Though see Diaz Leon (2016) for a detailed discussion of Saul's view, and a defense of contextualist accounts of 'woman'. Diaz Leon argues that 'woman' is a politically significant term, and that part of what determines what 'x is a woman' means in a context include 'instrumental, moral, and political considerations having to do with how X should be treated'. I'm skeptical, however, that there is always a fact of the matter, in a context, about who *should* count as a member of a particular gender (sometimes different moral and political goods conflict), and I'm also reluctant to say that political considerations are part of what determines truth value for politically significant terms. I think politically effective 'noble lies' should be possible.

As a result, I think the ameliorative project - when focused on gendered language - gives us an argument for permissivism about the self-ascription of gender terms. When a person says 'I am an x' for some gender term x, we should interpret the term 'x' in a way that makes the speaker's claim true. If someone says 'I am a woman', we have a choice about how we interpret her. We can hear her as making a claim about her own self-affirmation or internally felt sense of gender, about her gender expression, about her gender performance, etc. And there will inevitably be a way of interpreting her that makes her claim true - literally true, just as true as anyone else's statement of 'I am a woman' - regardless of what sex characteristics she has. Similarly, if someone says 'I am genderqueer - I'm not a man or a woman', we can interpret that as a claim about their experience of gender identity or gender self-expression in a way that makes it true (literally true, just as true as anyone else's self-ascription of a gender term). We don't need to first develop a theory of what it is to be non-binary - and of whether there is a difference between genderqueer and being agender, for example - to say that their self-ascription of gender is true.<sup>47</sup>

Simply put, these words are flexible and can mean many different things. Their meaning is politically and personally important in many cases - being able to have 'I am an x' recognized as true is not only part of having your own identity validated, it's also a part of fighting against oppressive norms about how people must or should experience gender (e.g. you must have a gender that's binary, you must have a gender that corresponds to your sex, etc.) We thus ought to, whenever possible, be permissive about our interpretation of self-ascribed gender terms, and likewise do our best to create contexts in which such permissivism is the norm.

Although Saul (2012) describes a broadly similar picture of gendered language, she refrains from endorsing the model she describes, in part because she is worried about its political ramifications. A view like this allows us to say that it's true that trans women are women, true that genderqueer people are nonbinary, etc. So far so good. But it also has the result that, in different contexts or on other interpretations of these terms, it's true - just as true - that trans women are not women and that genderqueer people are not non-binary, etc. The very flexibility that allows us to easily say 'Trans women are women' is true also allows us to easily say, in a different context or according to a different interpretation, 'Trans women aren't women' is true. Saul worries that this is unacceptable - not only does it make trans women's claims to being women (or genderqueer people's claims to being nonbinary, or etc) no better than their denials, it also undermines the importance

---

<sup>47</sup> Dembroff (in progress) develops a somewhat different argument for a similar type of permissivism about gender self-ascription. On their view - which they develop through the idea of 'unethical truths' - whether a sentence like 'x is genderqueer' is *literally* true is irrelevant to whether we should *treat* the statement as true. Dembroff's claim is that what matters to gender self-ascription are normative facts, not semantic or metaphysical ones. Dembroff and I agree about a lot. Our major point of disagreement is this: Dembroff thinks that the metaphysics of gender is ultimately irrelevant to how we should use gendered language, whereas I think the metaphysics of gender can inform our use of gendered language without thereby being an account of gender terms. (More specifically, I'm skeptical that we can adequately assess what the normative facts about gender *are* without a metaphysics of gender.)

of claiming a gender. If we grant that trans women are women only because the term ‘woman’ has multiple meanings or is contextually shifty, we don’t seem to be recognizing the moral, political, and personal importance of saying that someone is a woman.

But this worry can be avoided on the picture I’m defending here. It’s a mistake, I think, to focus too much on the mere truth of sentences like ‘x is a woman’. Truth is relatively easy to come by for natural language sentences (especially if we don’t endorse a robust correspondence-style theory of truth). Communication, on the other hand, is hard. There are probably contexts in which sentences like ‘Trans women aren’t women’ or ‘There’s no such thing as being nonbinary’ are literally true. But it doesn’t follow that such sentences are ever assertable, appropriate things to say. There are many things, over and above the basic content, which are communicated by a typical utterance of a sentence like ‘Trans women aren’t women’. This often includes things like ‘Gender is determined by biology’, ‘There is a correct way to express and experience gender’, ‘There’s something wrong or defective about people whose gender identity is different from the gender they were assigned at birth’, and so on.

On the view I’m defending, these are false *in any context*, simply because they misdescribe the basic social reality of gender. Whether it’s true that someone is genderqueer might be a contextually flexible matter that’s determined by how we use language, but whether it’s true that biology determines all the complex behavioral features we associate with gender is not. Similarly, whether it’s true there are facts about how people *should* or *ought* to experience gender is not. These are language-independent questions of what the social world is like. Thus even if one can, strictly speaking, truly say in a context ‘Trans women aren’t really women’ or ‘There’s no such thing as being genderqueer’, much of what one typically communicates by such an assertion will be false, making it an incorrect and inappropriate thing to say.<sup>48</sup>

More generally, while the truth of a statement like ‘x is genderqueer’ hinges on the flexibility and mutability of natural language terms, the facts about the underlying social and normative issues - which are what I think explain the political importance of our gender terms - do not. The reason why we should be permissivist about the self-ascription of gender terms is that our binaristic gendered social structures are oppressive and should be challenged. Biology doesn’t determine gender identity, gender identity is an important part of people’s experience of gender in society, and so on. The political importance of gender ascriptions, I suggest, is less about the application conditions for particular natural language terms, and more about treating people as having first-person authority about their own gender identity and expression.<sup>49</sup>

---

<sup>48</sup> Note also that this isn’t an issue confined to trans and non-binary individuals. There are doubtless contexts in which, e.g., ‘real women’ denotes a certain type of stereotypical femininity, such that ‘Childless women aren’t real women’ or ‘Butch lesbians aren’t real women’ are, strictly speaking, true. See especially Leslie (2015). But again, on my view these sentences always *communicate* things which are false. For an argument that we should treat pernicious generics in general as false, see Haslanger (2012)f.

<sup>49</sup> See especially Bettcher (2009)



---

## 9. Amelioration and the flexible use of gender terms

Contra some, however, I don't think we should reserve the use of gender terms for all and only those who would describe their own gender identity using such gender terms.<sup>50</sup> As argued above, I think that we should, whenever possible, treat the sincere self-ascription of gender terms as true. But it doesn't follow that gender terms should always or only refer to gender identity.

As I'm understanding the semantic component of the ameliorative project, our use of these flexible terms should be guided by our political and social goals. And while self-affirmation and self-identity matter to gender, they aren't the only things that matter. Gender is a many-splendored thing. Plausibly, the best way to capture this is by allowing for flexibility and mutability in the way we use gender terms (just as we want to allow for flexibility and mutability in the way people experience and express gender.) I think it's a mistake to argue that gender terms should exclusively refer to gender identities - even if they often refer to gender identities - simply because it's a mistake to think that there is any one thing that terms like 'woman' mean.

A major motivation for this, as discussed previously, is that treating gender terms as synonyms for gender identities creates exclusion problems in just the same way that treating gender terms as synonyms for social positions does. We want to be able to say that cognitively disabled are women, regardless of their gender identity or self-ascription. But more generally, there are also contexts where it seems perfectly legitimate to focus on whether people are socially classed as women, regardless of their gender identity. So, for example, if we are talking about the wage gap between women and men, what seems to matter most is whether someone is perceived by their employers and co-workers as a woman. If someone is a trans man but is not out in his workplace, for example, (and so presents as a woman and as female in his workplace, and is assumed to be a woman by all his co-workers and employers), his salary should also arguably be part of this conversation.<sup>51</sup>

Similarly, there are contexts in which it seems to matter primarily whether someone has the sex-related features we typically associate with women, regardless of their gender identity. So, for example, if we are talking about how doctor's tend to under-treat and dismiss women's health problems like endometriosis, the experiences of a genderqueer person with a uterus are relevant to that conversation, regardless of whether they would

---

<sup>50</sup> See Jenkins (2016) for an argument that we should reserve the terms in this way.

<sup>51</sup> In many cases in which we use gender terms in ways that contradict gender identity, we do so in ways that communicate a lot of false information, but crucially I don't think this is always the case. If we treat being a 'woman in the workplace' as primarily a matter of social position - primarily a matter of how others treat, react, and respond to you - in order to investigate how that affects things like salary, we aren't thereby communicating that one must have certain biological features in order to be a woman. Some uses of 'woman' are contextually very specific. And so, perhaps unsurprisingly, whether it's correct to classify a person as a woman might depend on the specific goals we have and the specific questions we're asking. See especially Anderson (1995).

self-identify as a woman. Likewise, if we are talking about how doctors tend to downplay women's reports of pain, the experiences of a genderqueer person with female sex characteristics should be a part of that conversation, even if they don't self-identify as a woman in most contexts.

Insisting that there is one thing that our gender terms like 'woman', 'man', 'genderqueer', etc really mean is unhelpful. Not only does it limit our ability to describe the complex ways in which people experience gender - including gender role, gender identity, gender expression, etc - but it also reinforces the idea that there's something that it *really* is to be a woman, be a man, be androgyne, etc. If I'm right, there are facts about gender - there is a social reality to gender that is independent our how we talk about and think about gender. And I've argued that this reality is best understood via a social position account of gender. But it doesn't follow that there are mind- and language- independent facts about who the men, women, nonbinary folk, genderfluid folk, pangender folk, etc are. And being flexible about our use of gender terms - and respecting people's ability to truly describe their own gender using their own preferred terms - is one part of the process of combatting a system of gendered oppression that is, at its bedrock, binary and based on social role.

---

## 10. Summing up: a middle ground for social position theories

Debates in the metaphysics of gender have typically been construed as debates over how we should explain what it is to be a woman and what feature(s) of the world the term 'woman' should refer to. Particular theories focus on attempting to give a definition or application conditions for gender terms like 'woman', and on explaining what it is that all and only women have in common with each other. Unsurprisingly, though, it has proven very hard to specify anything that all and only the women (or the men, or the genderqueer) have in common with each other, and extant theories are plagued by exclusion problems. Rather than thinking that this should lead us to skepticism about the metaphysics of gender, however, I've argued that this should instead lead us to slightly rethink the goals of the project. The task of giving an explanation of the social reality of gender can and should come apart from the task of saying what it is to be a woman.

I've argued that many of the problems associated with social position theories of gender are problems simply with thinking that a social position account is the kind of thing that can give us application conditions for terms like 'woman'. A social position theory of the kind I'm defending here doesn't include any objective, language-independent facts about which individuals count as women. On this view, there are real and objective facts about gender, but the social reality of gender doesn't neatly map on to our ordinary-language gender categories. It doesn't follow, though, that the issue of who the women are - or our use of the term 'woman' - is politically or philosophically insignificant. Which people fall under the extension of the term 'woman' can vary depending on the context, but plausibly anyone who is truly described as a woman - whether because they identify as a woman, because they have female sex characteristics, because they are perceived as having female sex characteristics, etc - is a recipient of gender injustice along some dimension. A social position account of gender can help us to explain why this is the case, with-

out thereby explaining what it is to be a woman. Similarly, a social position account can help us explain why the social reality of gender can in general be harmful to people, whether or not those people are described by certain natural language terms in a given context. A social position metaphysics of gender says that our social reality is structured in a way that places norms and expectations on people based on our perceptions of their sex. These structures constrain us - they tell us there's a way that we ought to behave, speak, dress, socialize, work, etc based on the ways in which other people react to our bodies. And those constraints can be harmful across the board - for women, men, genderqueer, gender-fluid, gender-anything. The substantial work of the metaphysics of gender, I've argued, should be in explaining the nature of these constraints, rather than in explaining the application conditions for natural language gender terms.

### Works Cited

- Alcoff, Linda (2005): *Visible Identities: Race, Gender, and the Self*. New York: Oxford University Press.
- Appiah, Kwame Anthony (1990). But would that still be me? Notes on Gender, 'Race,' Ethnicity as Sources of Identity. *Journal of Philosophy* 87 (10):75-81.
- Armstrong, David Malet (1978). *A Theory of Universals. Universals and Scientific Realism Volume II*. Cambridge University Press.
- Ásta (2013). 'The Social Construction of Human Kinds.' *Hypatia* 28 (4):716-732.
- Ásta (forthcoming). *Categories We Live By*. OUP.
- Bach, Theodore (2016). Social Categories are Natural Kinds, not Objective Types (and Why it Matters Politically). *Journal of Social Ontology* 2 (2):177-201.
- Bach, Theodore (2012). Gender Is a Natural Kind with a Historical Essence. *Ethics* 122 (2): 231-272.
- Bennett, Karen (2011). Construction area (no hard hat required). *Philosophical Studies* 154 (1): 79-104.
- Bettcher, Talia Mae (2013). "Trans Women and the Meaning of 'Woman'". In A. Soble, N. Power & R. Halwani (eds.), *Philosophy of Sex: Contemporary Readings, Sixth Edition*. Rowan & Littlefield 233-250.
- Bettcher, Talia Mae (2009). Trans Identities and First-Person Authority. In Laurie Shrage (ed.), *You've Changed: Sex Reassignment and Personal Identity*. Oxford University Press
- Bettcher, Talia Mae (2007). Evil deceivers and make-believers: On transphobic violence and the politics of illusion. *Hypatia* 22 (3):43-65.
- Briggs, R.A. and B.R. George (in progress). Science Fiction Double Feature: Trans Liberation on Twin Earth.

- Butler, Judith (1990). *Gender Trouble*. New York: Routledge.
- Cameron, Ross P. (2008). Truthmakers and ontological commitment: or how to deal with complex objects and mathematical ontology without getting into trouble. *Philosophical Studies* 140 (1):1 - 18.
- Dembroff, Robin (in progress). Real Talk About the Metaphysics of Gender.
- Diaz Leon, Esa (2016). ‘Woman’ as a Politically Significant Term: A Solution to the Puzzle. *Hypatia* 31(2), pp. 245-58.
- Dorr, Cian (2005). What we disagree about when we disagree about ontology. In Mark Eli Kalderon (ed.), *Fictionalism in Metaphysics*. Oxford University Press 234--86.
- Epstein, Brian (2015). *The Ant Trap: Rebuilding the Foundations of the Social Sciences*. Oxford.
- Jenkins, Katharine (2016). Amelioration and Inclusion: Gender Identity and the Concept of Woman. *Ethics* 126 (2):394-421.
- Haslanger, Sally (2012): *Resisting Reality*. Oxford: OUP
- Haslanger, Sally. (2012)a: ‘Gender and Race: (What) Are they? (What) Do we want them to be?’ in *Resisting Reality*. Oxford OUP, p. 221-247.
- Haslanger, Sally (2012)b: ‘You Mixed? Racial Identity without Racial Biology.’ in *Resisting Reality*. Oxford: OUP, p. 273-297.
- Haslanger, Sally. (2012)c: ‘Social Construction: The ‘Debunking’ Project’ in *Resisting Reality*. Oxford: OUP, p. 113-138.
- Haslanger, Sally. (2012)d: ‘What are We Talking About: the semantics and politics of social kinds’ in *Resisting Reality*. Oxford OUP, p. 365-380.
- Haslanger, Sally. (2012)e: ‘What Good Are Our Intuitions?: Philosophical Analysis and Social Kinds’ in *Resisting Reality*. Oxford OUP, p. 381-405.
- Haslanger, Sally. (2012)f: ‘Ideology, Generics, and the Common Ground’ in *Resisting Reality*. Oxford OUP, p. .
- Haslanger, Sally. (2014): ‘Individualism, Interpretation, and Injustice: A reply to Stahl, Betti, and Mikkola’. *Krisis* 1, p. 24-38.
- Kapusta, Stephanie. (Forthcoming). “Misgendering and Its Moral Contestability,” *Hypatia*.
- LaBrada, Eloy (2016). ‘Categories We Die For: Ameliorating Gender in Analytic Feminist Philosophy.’ *PMLA* 131 (2):449-459.
- Leslie, Sarah-Jane (2015). Hilary Clinton is the Only Man in the Obama Administration. *Analytic Philosophy* 56(2), p. 111-141.

- Merricks, Trenton (2001). *Objects and Persons*. Oxford University Press.
- Mikkola, Mari (2009). Gender concepts and intuitions. *Canadian Journal of Philosophy* 39 (4):pp. 559-583.
- Mikkola, Mari (2008). Feminist perspectives on sex and gender. Stanford Encyclopedia of Philosophy.
- Mikkola, Mari (2006). Elizabeth Spelman, Gender Realism, and Women. *Hypatia* 21 (4):77-96.
- McKittrick, Jennifer (2015). A dispositional account of gender. *Philosophical Studies* 172 (10): 2575-2589.
- Paul, L. A. (2012). Building the world from its fundamental constituents. *Philosophical Studies* 158 (2):221-256.
- Saul, Jennifer (2012). Politically Significant Terms and the Philosophy of Language: Methodological Issues.” Eds. S. L. Crasnow and A. M. Superson. *Out from the Shadows: Analytical Feminist Contributions to Traditional Philosophy*. Oxford: Oxford University Press.
- Saul, Jennifer (2006). Philosophical Analysis and Social Kinds. *Proceedings of the Aristotelian Society, Supplementary Volumes*, vol. 80, 2006, pp. 89–143.
- Schaffer, Jonathan (2009). On what grounds what. In David Manley, David J. Chalmers & Ryan Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press 347-383.
- Serrano, Julia (2007): *Whipping Girl*. Emeryville, CA: Seal Press.
- Sider, Theodore (2011). *Writing the Book of the World*. Oxford University Press.
- Spellman, Elizabeth (1988). *The Inessential Woman: Problems of Exclusion in Feminist Thought*. Boston: Beacon Press.
- Stoljar, Natalie (2011). Different Women. Gender and the Realism-Nominalism Debate. In Charlotte Witt (ed.), *Feminist Metaphysics*. Springer Verlag 27–46.
- Stone, Sandy (1992): ‘The ‘Empire Strikes Back: A Post Transsexual Manifesto’ in Kristina Straub and Julia Epstein, eds. *Body Guards: The Cultural Politics of Gender Ambiguity*. New York: Routledge.
- Turner, Jason (2016). *The Facts in Logical Space: A Tractarian Ontology*. Oxford: OUP.
- van Inwagen, Peter (1990). *Material Beings*. Cornell University Press.
- Watson, Lori (2016). “The Woman Question,” *Transgender Studies Quarterly* 3:1-2: 248-255.
- Wesselius, Janet Catherina (1998). Gender Identity Without Gender Prescriptions. *Symposium: The Canadian Journal of Continental Philosophy* 2 (2):223-235.

Witt, Charlotte (2011). *The Metaphysics of Gender*. OUP Usa.

Witt, Charlotte (2011). What Is Gender Essentialism? In *Feminist Metaphysics*. Springer Verlag 11--25.