

Some Internal Problems with Revisionary Gender Concepts

Abstract

Feminism has long grappled with its own demarcation problem – exactly what is it to be a woman? – and the rise of trans-inclusive feminism has made this problem more urgent. I will first consider Sally Haslanger’s “social and hierarchical” account of *woman*, resulting from “Ameliorative Inquiry”: she balances ordinary use of the term against the instrumental value of novel definitions in advancing the cause of feminism. Then, I will turn to Katharine Jenkins’ charge that Haslanger’s view suffers from an “Inclusion Problem”: it fails to class many trans women as women. Jenkins offers a novel norm-relevancy account of *woman* to avoid the Inclusion Problem. Unfortunately, Jenkins’ account has serious internal problems, i.e. problems by Jenkins’ own lights: it is unintelligible, or it suffers from an Inclusion Problem of its own. After that, I will develop novel arguments for the conclusion that the project of Ameliorative Inquiry is both incoherent and also impossible to complete – at least, impossible to complete in a trans-inclusive way. Trans-inclusive feminism, therefore, would do well to move beyond Ameliorative Inquiry. Insofar as that’s not possible, trans-inclusive feminism inherits the incoherence of Ameliorative Inquiry.

Introduction

Feminism has long grappled with its own demarcation problem – exactly what is it to be a woman? – and the rise of trans-inclusive feminism has made this problem more urgent. Once upon a time – and even now in some linguistic communities – our term “woman” picked out a biological class: *adult human biological females*. This is what Sally Haslanger (2012, 375) calls “the dominant manifest meaning,” what Toril Moi (1999, 8) calls “the ordinary understanding,” and what Talia Mae Bettcher (2009, 103) calls the “definitional account” of the term “woman.” I’ll call it “the traditional concept,” and it is reported by lexicographers in, for example, the Oxford English Dictionary.¹

¹ Though she goes on to reject it, Bettcher (2009, 105) says of this definition: “On the face of it, the definition ‘female, adult, human being’ really does seem right. Indeed, it seems as perfect a definition as one might have ever wanted.” To say that this is the traditional concept of womanhood is *not* to deny conceptual connections between this concept and other concepts of norms, identity, expression, or social status. It’s only to say that, when exercising this traditional concept of womanhood – when using the “dominant manifest meaning” of the term “woman” – to think of, for example, norms about women or acceptable modes of expression for women, is to think of norms about *adult human females*, and acceptable modes of expression for *adult human females*.

But, lately, some feminist philosophers have sought to revise our concept of womanhood, in order to advance the cause of social justice. Following Sally Haslanger, I'll call this deliberate conceptual engineering "Ameliorative Inquiry." The idea is to reflect on which concept we *should* associate with the term "woman," in order to best achieve the goals of feminism, e.g. calling out unjust discrimination and subordination of women, honoring the gender identifications of trans women, and the like.

In this paper, I will evaluate some of those recent attempts to intentionally revise our gender concepts. I'll begin with Sally Haslanger's social and hierarchical definition, and report Katherine Jenkins' (2016) recent objection that it fails to include many trans women. I will then turn to Jenkins' own revisionary concept of womanhood, defined in terms of taking enough of the right kind of gender norms to be relevant to oneself. I will argue that Jenkins' definition either excludes or marginalizes many trans women, or it is unintelligible.

I will also develop novel arguments for the conclusion that the project of Ameliorative Inquiry is both incoherent and also impossible to complete in a trans-inclusive way. It's incoherent because, while it endeavors to stay on the topic of gender, by intentionally introducing new concepts, it necessarily changes the subject. And Ameliorative Inquiry is impossible to complete in a trans-inclusive way because we're meant to evaluate proposed revisions of our gender concept in light of the goals of *feminism*. That requires already knowing what womanhood is, and so we can't use the revisionary concepts themselves to do the evaluation. But if we understand womanhood in the traditional, biological way, it's hard to see how the output of Ameliorative Inquiry could be trans-inclusive. Trans-inclusive feminism, therefore, would do well to move beyond Ameliorative Inquiry. To the degree that this is impossible, trans-inclusive feminism inherits the incoherence of Ameliorative Inquiry.

Ameliorative Inquiry and Sally Haslanger's Account of Gender

Sally Haslanger's influential work offers the following definition of womanhood, in social and hierarchical terms:

Someone is a woman if and only if she is systematically subordinated along some dimension (economic, political, legal, social, etc.), and she is "marked" as a target for this treatment by being, regularly and for the most part, observed or imagined to have bodily features presumed to be evidence of a female's biological role in reproduction.²

You won't be the first to notice that this definition of womanhood departs from the traditional conception of womanhood, introduced in the previous section. It seems perfectly possible, for example, for a man to meet Haslanger's definition

² See Haslanger 2000, p.39 and p. 42. She offers a similar definition of manhood, though in terms of privilege rather than subordination.

of womanhood. Suppose that a cis gender, heterosexual, biologically male transvestite occupies a subordinate position on the basis of observed *or imagined* bodily features linked to the female sex, due to his preferred style of dress. Even if he sincerely insists that he is a man, Haslanger's definition classifies him as a woman. And this is because Haslanger aims to disclose what she calls the "operative" concept of womanhood, "the more implicit, hidden, and yet practiced" concept (2012, 370), the "properties attributed to the things" when we distinguish members of the extension of the concept from non-members (2012, 105). You might have thought that the operative concept and the manifest concept coincide, and that the properties distinguishing women from non-women are just those represented by the manifest concept: being an adult, being human, and being biologically female. But, according to Haslanger, the operative and manifest concepts of womanhood do *not* coincide; our operative concept is masked, she says, by the manifest concept, "understood as defining *what women are by nature* in traditional terms" (2012, 93). The definition she offers above is what she takes to be the true operative concept: a *role*, or a *position* in society. She offers a functionalist, "black box" view of womanhood, a view indifferent to *how* the box functions, to what it is that plays the role. We can see this by placing a man in that role.³

Also, relying on that "dominant manifest meaning" of the term "woman," it seems perfectly possible for a woman to fail to meet Haslanger's definition of womanhood. Suppose, for example, that we achieve ideal gender justice: no group occupies subordinate positions at all; repressive hierarchies have been abolished. If so, Haslanger's definition says *there couldn't be any more women*. But you might think there could be, at least according to the traditional conception of womanhood. Also, as Mari Mikkola (2009, 565) points out, "the Queen of England is not oppressed on sex-marked grounds and so, would not count as a woman on Haslanger's definition."⁴ These issues emerge because

³ You might think it easy to sidestep this counterexample by erasing "or imagined" from Haslanger's definition. But this comes at a serious cost: womanhood would then require the actual observation of bodily features presumed to be evidence of a female's biological role in reproduction. And then, again, we could easily think of cases to show that this conception of womanhood departs from the non-revisionary conception, e.g. by imagining a woman who never undergoes the required inspection.

⁴ Perhaps you're of a mind that even the Queen is in fact systematically subordinated (economically, politically, legally, socially, etc.) on sex-marked grounds, maybe because she's still judged or constrained by gender norms. If so, just move to a nearby possibility where Her Majesty suffers no such subordination. Insofar as you judge it to be possible that she nevertheless remains a woman, Mikkola's point stands: oppression is not essential to the traditional conception of womanhood. But perhaps you doubt it's possible that even the Queen could escape subordination. As a last resort, then, attend to the intuitive possibility that a woman might be the last person on Earth, facing no systematic subordination, but remaining a woman nonetheless. Of course, Haslanger may bite the bullet here, and say that these cases are impossible. But the question is what you, the reader, judge to be possible, exercising the traditional conception of womanhood.

Haslanger's proposed definition makes subordination *essential* to the application of "woman."

Finally, this proposed definition would, on occasion, reverse the order of explanation in an implausible way. For example, on the non-revisionary, "dominant manifest meaning" of *woman*, in sexism and misogyny, people are oppressed *because they are women*. Explanation ought to go in that direction: on these occasions, womanhood explains (without, of course, excusing) why some people were targeted and oppressed.⁵ But, on Haslanger's proposed definition, things are the other way around: people are women *because they are oppressed*.⁶ It's the oppression that explains why someone is a woman. To the degree that we find this implausible, we should admit that Haslanger's definition is revisionary, departing from the traditional conception of womanhood.⁷

Now, is all this a problem for Haslanger's project? The answer isn't obvious. In some places, it does sound as though she's attempting to give us the real definitions of our ordinary gender terms. For example, she gives a parallel account of race, one she explicitly models on her account of gender, and offers it "as one that captures our ordinary use of the term" (2008, 66). She says that her proposed definition of womanhood provides "a better explanation of how gender works" (2012, 135), and she sees herself "working within a semantic externalist model that looks to social theory to provide us with an account of our social terms, just as scientific essentialism looks to the physical sciences to provide an account of our naturalistic terms" (*ibid.*, 137).

So, Haslanger doesn't mean to change the subject, to leave our ordinary gender terms behind and define *new* homonyms "man" and "woman," and then speak, rather misleadingly, using those homonyms. Instead, she does indeed

⁵ I'd say that being female also explains the targeting and oppression. Of course, the non-revisionary sense of "woman" accommodates that. But the point here is that, on Haslanger's account, being a woman does not explain the targeting and oppression. Rather, only being female explains the targeting and oppression, and then it's this targeting and oppression which explains why the victim is a woman. If you find this infelicitous, you have here evidence that Haslanger's account departs from our ordinary use of the term "woman."

⁶ At least, oppressed in a particular sort of way. On this, Haslanger echoes Simone de Beauvoir's (1973, 301) famous dictum that "One is not born, but rather becomes, a woman." Less well known are de Beauvoir's problematic trans-exclusive dicta, for example (*ibid.*, 18) that "...there have always been women. They are women in virtue of their anatomy and physiology," and also (*ibid.*, 59) "...woman is of all mammalian females at once the one who is most profoundly alienated..., and the one who most violently resists this alienation."

⁷ Also, as we learned from Socrates' encounter with Euthyphro, to the degree that there is *any* order of explanation between Haslanger's definiens and her proposed definiendum, we should conclude that we don't have here a genuine analysis of the traditional concept of womanhood. An order of explanation shows that the definiens and the definiendum pick out *two* qualities, one of which explains the other, and so the analysis fails.

mean to give an account of *our* ordinary gender terms. It's just that – and this is crucial – she thinks commonsense judgments about paradigm cases are not the only standard by which we should measure analyses of our gender terms. What we thought were paradigm cases of women may not fall within the kind that turns out to be the best extension of our gender term “woman” (ibid., 135).⁸ So while apparent counterexamples like the ones above *matter*, they do not decisively derail Haslanger's project. For Haslanger, capturing our ordinary discourse is merely one factor in a process of conceptual engineering, and this process may, in the end, land on her proposed definition as the best available.⁹

What other factors are relevant in this process of conceptual engineering? According to Haslanger, we may come to accept definitions of our gender terms for the sake of accomplishing the goals of feminism – e.g. combating gender injustice – even if those definitions give the ‘wrong’ results for apparently paradigm cases. She says “I believe that races and genders are real categories to be defined in terms of social positions. I have come to this conclusion by considering what categories we *should* employ in the quest for social justice” (ibid., 365, emphasis added). She goes on to call such an inquiry “ameliorative,” since it looks toward what legitimate social or political purposes we might have in using gender terms, and develops concepts that would help us achieve those ends. She recommends each of her definitions “not because it provides an analysis of our ordinary discourse, but because it offers numerous political and theoretical advantages” (ibid., 366).¹⁰

So, Haslanger is engaged here in a broader project of *Ameliorative Inquiry*: when she evaluates a proposed definition of a gender term, she balances our ordinary use of gender terms against the instrumental value of the definition in achieving the goals of feminism.¹¹ And this is why apparent counterexamples to her definitions are not decisive refutations of her project. What would be more troubling, for Haslanger, would be evidence that her proposed definitions, in

⁸ Of the cases given above of apparently paradigmatic women who don't meet her proposed definition, Haslanger (ibid., 239) says: “I believe it is part of the project of feminism to bring about a day when there are no more women (though, of course, we should not aim to do away with females!).” Meditate on that statement – “part of the project of feminism is to bring about a day when there are no more women” – in order to appreciate how far Haslanger's definition departs from our ordinary conception.

⁹ Haslanger says (ibid., 239): “I'm happy to admit that there could be females who aren't women in the sense I've defined, but these individuals (or possible individuals) are not counterexamples to the analysis. The analysis is intended to capture a meaningful political category for critical feminist efforts, and non-oppressed females do not fall within that category...”

¹⁰ In a similar way, Bettcher (2009, 107) gives purely moral reasons in favor of revising our traditional gender concepts: she believes that the traditional gender concepts perpetuate a system of forced disclosure of genital status, which she says is invasive and sexually abusive.

¹¹ For a sketch of those goals and values, see *Resisting Reality*, 226-7.

addition to apparently misclassifying paradigm cases of men and women, actually *frustrate* the cause of feminism. That would be what we might call an “internal” problem for Haslanger’s project: a problem by her own lights, a problem on her own terms.¹² Katharine Jenkins has recently given an argument to this effect, and it’s there I now turn.

Katharine Jenkins’ “Inclusion Problem” for Haslanger’s Ameliorative Inquiry

Recall that, when we evaluate her proposed definition of gender terms, Haslanger asks us to weigh the costs of apparently misclassifying paradigmatic men and women against the benefits of advancing the project of feminism. By arguing that Haslanger’s definition stymies the cause of feminism, Katharine Jenkins (2016) raises an *internal* problem for Haslanger’s project of ameliorative inquiry, i.e. a problem by the project’s own lights. According to Jenkins, Haslanger’s definition is too narrow: it fails to classify some people as women whom it ought to, for the sake of social justice.

This is Jenkins’ “Inclusion Problem” for Haslanger, and it has two simple steps. First, Jenkins asserts (ibid., 396) as a foundational premise, that “trans gender identities are entirely valid – that trans women are women and trans men are men.” Further, she says, “[f]ailure to respect the gender identifications of trans people is a serious harm and is conceptually linked to forms of transphobic oppression and even violence.” That’s the force behind the objection: failure to respect trans gender identities will at least slow the cause of gender justice, and in fact may well promote injustice.¹³

Secondly, Jenkins says that Haslanger’s definition fails to respect trans gender identities, since it fails to classify some trans women as women. For example, she says (ibid., 399-400), trans women who do not publicly present as women may well not meet Haslanger’s definition: they won’t “function as women,” in Haslanger’s sense, since they won’t be observed or imagined to have female bodily features. Additionally, she says (ibid., 400), some trans women who *do* publicly present as women but are nevertheless *treated* as men wouldn’t meet Haslanger’s definition. And, finally, there may be trans women whose gender identifications are respected unconditionally, and *not* on the basis of observed or imagined bodily features associated with a female’s biological role in reproduction. Though such a person may well function as a woman, Jenkins says

¹² My Marxist and post-Marxist colleagues might prefer the term “immanent critique.” As Titus Stahl (ms.) defines it, we’re engaged in “a critique which derives the standards it employs from the object criticized... rather than approaching that [object] with independently justified standards.”

¹³ I do not mean to endorse or deny Jenkins’ foundational premise here. Since Jenkins takes it as foundational, I believe more progress could be made by raising an *internal* problem for Jenkins, a problem by her own lights, a problem that arises even if we grant this foundational premise.

(ibid., 400-1), this person would not count as a woman on Haslanger's definition, since she's not treated as a woman *on the basis* that Haslanger says is essential for womanhood.

So ends the Inclusion Problem. And, recall, the force of the objection is that failure to respect trans gender identities will frustrate the cause of feminism, by exacerbating gender injustice, promoting or facilitating transphobic violence, etc. As Jenkins puts it (ibid., 403), "in this feminist theory many trans women would only be counted as women in a secondary or peripheral sense – a paradigmatic case of marginalization."

Jenkins' objection to Haslanger is on point: Haslanger's definition does fail to respect the gender self-identification of some trans people. Insofar as Haslanger agrees that this cuts against the cause of feminism, Jenkins raises a powerful internal objection to Haslanger's project of Ameliorative Inquiry, a problem on Haslanger's own terms. However, Jenkins' proposed definition of "woman" has internal problems of its own. Next, I'll discuss Jenkins' view. After that, we'll see its internal problems.

Jenkins' Revisionary Definition of "Woman"

Jenkins rejects Haslanger's functional, social-role account of womanhood, for reasons just given. Instead, Jenkins proposes to ground womanhood in *psychology*, to reserve "woman" for people who *take* norms of femininity to be relevant to themselves.¹⁴ More carefully, she defines womanhood like so (ibid., 410ff):

Someone is a woman if and only if she has an internal 'map' that is formed to guide someone classed as a woman through social and material reality.

She adapts "classed as a woman" from Haslanger's definition: to be classed as a woman is to be, regularly and for the most part, marked as a target for subordination on the basis of observed or imagined bodily features presumed to be evidence of a female's biological role in reproduction. But, notice, it's not actually *occupying* any social role that grounds womanhood, for Jenkins. Rather, what grounds womanhood is the *having* of that "internal map."

Later, Jenkins says more about what it is to have this internal map (ibid., 411):

On my definition, having a female gender identity does not necessarily involve having internalized norms of femininity in the sense of accepting

¹⁴ True, Jenkins proposes two distinct senses of "gender": gender-as-class and gender-by-identity. But she's quite clear (2016, 417) that she supports "using the term ['woman'] to refer to people with a female gender identity and not, in general, using it to refer to people classed as women."

them on some level. Rather, what is important is that one takes those norms to be relevant to oneself; whether one feels at all moved to actually comply with the relevant norms is a distinct question.

To illustrate, Jenkins gives an example of a woman who knows that, according to her cultural context, she's supposed to shave her legs.¹⁵ But she chooses not to. Her knowledge of how her body should be (according to her society) is part of her gender identity, says Jenkins. This woman accepts that this dominant norm of feminine appearance applies to her; she has internalized it; it is therefore part of her "internal map." And – importantly – this is so even though she does not comply that norm. Jenkins' idea is that anyone who internalizes (without necessarily abiding by) enough of the right sort of gender norms will thereby qualify as a woman. And, again, to internalize a gender norm is to take it to apply to oneself, to take the norm to be *about* a type of person that one is.¹⁶

Jenkins offers her account as a solution to the Inclusion Problem. While Haslanger's definition fails to respect the gender self-identification of some trans women, those trans women will count as having a *female gender identity* on Jenkins' view, even if they're not *classed* as women by society (e.g. because they don't present as women). And, as I mentioned above, Jenkins (ibid., 417) thinks we should reserve the term "woman" for those who have a female gender identity. So, as long as they've internalized a sufficient number of the right kind of gender norms, trans women will count as women, on Jenkins' view. That

¹⁵ Of course, gender norms vary widely by cultural context. If you don't like Jenkins' example here, please replace it with one you prefer.

¹⁶ These points provide Jenkins with a handy and satisfying response to a recent objection from Matthew Salett Andler (2017). Andler argues that Jenkins' account will not be fully trans-inclusive, since "gender maps characteristic of people socially positioned as either women or men only allow a body to be experienced as unified if its features 'match' along the axis of a single sex," and some trans people experience bodily unification without such "matching." This is possible, Andler says, "[o]nly by adopting a radical gender map – distinct from the gender maps characteristic of people socially positioned as either women or men..." And therefore these trans persons wouldn't, on Jenkins' view, count as women or men. But Jenkins has two possible replies. First, even if Andler has correctly located some relevant gender norms (concerning the experience of bodily unification and "matching"), they are by no means the *only* gender norms, and trans people who don't take these norms to apply to them may still well take many other gender norms to apply to them, sufficiently many to meet Jenkins' definition. Also, Jenkins may take this occasion to remind us that *taking a gender norm to be relevant to oneself* does not require that one *abide* by that norm. (Recall Jenkins' example of the woman who knows that, according to society, she's "supposed" to shave her legs, and yet she doesn't.) And so, even if there is an experience-of-bodily-unification-requires-'matching' gender norm, and even if trans people don't conform to it, that's no barrier to their taking the norm to be relevant to them, and for that to count in favor of their being classified appropriately by Jenkins' lights.

much is right.¹⁷ However, as we'll see in the next section, Jenkins' view has an Inclusion Problem of its own.

An Internal Problem for Jenkins' Revisionary Definition of "Woman"

According to Jenkins, to be a woman (by identity) is to take a sufficient number of "norms of femininity" to be relevant to oneself, to apply to oneself, to be *about* or *for* oneself. Now, social norms are about (types of) people. And a problem emerges when we ask *who these norms of femininity are norms about*, who these norms are *for*. Norms of the medical profession are norms about doctors, nurses, etc.; that's who the norms are *for*. Academic norms are norms about professors, students, etc. Now, what about Jenkins' "norms of femininity"? Who are these norms *for*? Presumably, these norms that play a role in determining womanhood are norms *about women*, they are norms *for women*, they are statements of how women ought to behave. In that case, substituting talk of internal maps for equivalent talk of gender norms, Jenkins' definition of womanhood amounts to this:

Someone is a woman if and only if she takes (a sufficient number of) norms about women to be relevant to herself, i.e. to be *about* herself.

And now the problem becomes obvious: we've used "women" in a definition of womanhood. So, what is the meaning of "women" in that definiens? There are a few options, but each one has problems.

A first option – and the one that Jenkins takes – is to read "women" in Jenkins' proposed definiens as referring to Haslanger's social-role concept of womanhood: someone "classed" as a woman, someone subordinated on the basis of observed or imagined female biological sex. But this option inherits an Inclusion Problem of its own: the trans women who are excluded from Haslanger's definition can't sensibly take norms for women-in-Haslanger's-sense to be norms for *them*, and so they can't sensibly be women on Jenkins' view. They might falsely believe that they are among the women-in-Haslanger's-

¹⁷ This proposal, though, departs quite radically from our ordinary concept of womanhood. Jenkins is well aware of this, and is quite explicit that she's engaged in a revisionary project of Ameliorative Inquiry. But it's worth reflecting on the divergence between this proposal and our ordinary concept of womanhood. An anonymous referee helpfully points out intuitively possible cases like this: a woman, raised perhaps by radical feminists, who has not internalized enough of the right sort of gender norms to meet Jenkins' definition; a woman who, perhaps due to cognitive impairment, is unable to internalize enough of the right sort of norms to meet Jenkins' definition; a woman who, due to difficult relocation or immigration, has not internalized enough of the right sort of norms to count as a woman in her new culture. Jenkins' definition, then, seems to marginalize women with non-normative backgrounds, cognitive impairments, or difficult immigration experiences. (See also Elizabeth Barnes (forthcoming, 7-8) for similar points.) Insofar as you think this is a cost, you may want to balance it against the purported benefits of Jenkins' proposal in advancing the cause of feminism.

sense, as a result of a mistake or a delusion. Or they might irrationally have this set of beliefs: These norms are about all and only women-in-Haslanger's-sense; I am not a woman-in-Haslanger's-sense; Nevertheless, these norms are about me. So, on Jenkins' view, a trans woman who thinks this way can indeed be a woman, but only as the result of a mistake, or a delusion, or irrationality. And that's a kind of marginalization. But then Jenkins hasn't solved the Inclusion Problem that she herself leveled against Haslanger.¹⁸

A second option is that we might insert Jenkins' own revisionary definition of "woman" into the definiens. To be a woman, on this account, is to take norms about women—in this very same sense—to be relevant to oneself.

¹⁸ This problem remains even if we don't interpret Jenkins' proposed definition as requiring a *belief* that a sufficient number of the right kind of gender norms are directed at oneself. I am grateful to an anonymous referee for proposing to me a reading on which *taking these norms to be relevant to oneself* is to have a *sense* of the norms being things one needs to consider, respond to, and relative to which one experiences her behavior as either transgressive or not. And perhaps it's possible to have this experience even if one knows that the norm is not directed at oneself. However, I think the problem remains. To see how, consider that, as a man, there is some minimal degree to which I have a sense of gender norms about women as things I need to consider, respond to, and things that I knowingly transgress with my behavior. Yet this is not sufficient to make me a woman. So, we'll have to read this proposal in a thicker, more substantial way in order to exclude people like me from the category "woman." For a woman, these norms are things one must consider, respond to, and things that one transgresses in the right way *because these norms are about oneself*. But then my concern remains: this thicker, more substantial reading will marginalize trans women, by requiring a kind of delusion, or mistake, or irrationality of them in order to meet the definition of womanhood. For if these norms are in fact not directed at or *about* oneself, then it is irrational, or mistaken, or delusional to have the sense that these norms are things one must consider, respond to, and things that one transgresses in the right way *because these norms are about oneself*. Trans individuals would be victims of something like an illusion or a hallucination, on this option.

Another anonymous referee helpfully points out that an alternative direction for Ameliorative Inquiry might be to say that a woman is someone who *desires* enough of the right kind of gender norms to be about her, or who *believes* that enough of the right kind of gender norms *should* apply to her, even if they don't in her current cultural context. This sort of desire-based account of womanhood would avoid the present objection, and seems to accord with the testimony of trans individuals. Julia Serano (2007), for example, describes her experience in terms of a "subconscious *desire* to be female." Genny Beemyn and Susan Rankin (2011, 54) recount the responses of two transgender women who participated in a survey of theirs, both of whom described their experiences in terms of *wanting* to be female. Arnold Grossman et al. (2005, 9) report that "Of the MTF youth, only 1 reported never wishing to be born of the sex other than her birth sex, while 26 always wished to be born of the sex other than their birth sex, with the other 4 using the sometimes points of the scale." More recently, trans writer Andrea Long Chu (2018) described her experience as follows: "I am trying to tell you something that few of us dare to talk about, especially in public, especially when we are trying to feel political: not the fact, boringly obvious to those of us living it, that many trans women wish they were cis women, but the darker, more difficult fact that many trans women *wish they were women, period*." Though a desire-based account along these lines will be subject to further objections that I lay out in this paper (and others that the reader may supply), it is, I believe, worth exploring for the Ameliorative Inquirer.

But then the definition is viciously circular. To feel the gravity of this defect, observe that, after two recursive iterations, the account amounts to:

Someone is a woman if and only if she takes norms about [*those who take norms about* (those who take norms about <...> to be relevant to them) *to be relevant to them*] to be relevant to herself.

The ellipses represent an irredeemable void in the middle of the definiens, and help us see why the definition cannot succeed in expressing a proposition. On this option, then, Jenkins' definition fails: it is literally unintelligible, due to vicious circularity.¹⁹

A third option is that we're meant to read "women" in the definiens in the non-revisionary sense, i.e. as picking out *adult human biological females*. To be a woman, on this account, is to take norms of femininity – i.e. norms about adult human biological females – to be relevant to oneself, to be about oneself. But, in that case, Jenkins' definition excludes trans women who don't think of themselves as biologically female, i.e. the many trans women who believe their gender does not "match" their (male) biological sex. They do not take themselves to be biological females – hence, in many cases, the desire to be female, and to transition physically – and so they do not take norms about biological females to be norms about them. But then Jenkins' definition says they are not women, in which case an Inclusion Problem reemerges for her own view. And this Inclusion Problem for Jenkins is plausibly worse than that which she raises for Haslanger, since it applies to all trans women who take their gender not to "match" their biological sex.²⁰

¹⁹ John Corvino (2000, 178) comes *close* to putting his finger on this difficulty, saying "The paradox of analysis is that an analysis is unhelpful if it exactly synonymous with the analysandum; unsuccessful if it is not. The analysis of gender in terms of self-conception exhibits the former difficulty..." Yet, contra Corvino, the problem here is not unhelpful synonymy, but rather unintelligibility due to vicious circularity. Bettcher (2009, 109) comes a bit closer, attributing to Jacob Hale the worry that if being a woman is merely a matter of believing that one is a woman, then "we seem to have some problem of circularity or regress. In practice this means that the criterion is virtually unintelligible." This passage suggests that Bettcher thinks the regress may not be vicious, and that the problem is merely one of practicality. The concern I'm expressing in this section is graver: the regress is indeed vicious, and in principle, not merely in practice.

²⁰ There's an emerging trend in some pockets of feminist philosophy to revise our ordinary notions of biological sex, and to say that biological sex itself is a matter of self-identification, or norm-relevancy, or the like. To be female is to identify as female, for example. (Observe how, for example, Jenkins slides between "woman" and "female gender identity" in her 2016 work, especially p. 409, note 33.) That proposal would avoid the problems of this paragraph. But problems similar to those besetting "the fourth option" below would emerge for this proposal. See this by asking yourself just what exactly those who identify as female are identifying as. The traditional conception of biological sex had an answer in terms of chromosomes, hormones, genitalia, etc., but what does this revisionary conception offer? It's hard to see how this revisionary proposal will avoid vicious circularity without plunging headfirst into unintelligibility.

Moreover, *if* medical technology cannot really change a person's biological sex,²¹ then on this alternative Jenkins is committed to saying that literally any trans woman would be *mistaken* to think of herself as a woman: to be women, biologically male trans women must (mis)take norms about biological females to be about them. For Jenkins, on this alternative, trans women can be women, but only via an error, and not sensibly or rationally. Paradoxically, on this lemma, they are women because they mistakenly think they are, i.e. in virtue of *mistaking* norms for women to be norms for them. And that is paradigmatic marginalization, which is a serious problem by Jenkins' own lights.

A fourth option is to say "women" in Jenkins' definiens is *not* Haslanger's sense, and also *not* the non-revisionary sense, and then to become quietists about its meaning. It means something else, something we can't say. But without any clear sense of "women" in the definiens, the definition becomes unintelligible, amounting merely to this:

Someone is a woman if and only if she takes norms about [...] to be relevant to herself.

Like the first option, there's a break in the right biconditional, and the definition fails to express any proposition. As far as definitions go, that's a vice.²²

²¹ Plausibly, one's biological sex is determined by one's hormone levels, one's primary and secondary sex characteristics, and one's genetics. And we likely weight these features according to how central or crucial we take them to be to membership in a biological sex category. Current medical technology can do quite a bit to modulate one's hormone levels, significantly less to transform one's primary and secondary sex characteristics, and nothing at all to alter one's underlying genome. Depending on how one weights these features, then, one may well conclude that current medical technology cannot really move a person from one biological sex category to another. If so, the Inclusion Problem for Jenkins' view is quite substantial indeed.

²² The concerns in this section apply also, *mutatis mutandis*, to Bettcher's (2013, 241) admittedly rough definition of "trans woman." She says: S is woman if and only if S is a trans woman or a non-trans woman. And S is a non-trans woman if and only if S is a woman and not a trans woman. And S is a trans woman if and only if S was assigned to the male sex at birth, currently lives as a woman, and self-identifies as a trans woman (or as a woman). As you can see, "woman" is eventually used on both sides of that initial, disjunctive definition of womanhood. (And "trans woman" is used in the definition of trans womanhood.) So the circularity concerns of this section apply to Bettcher's definition of "trans woman." The concerns of this section also apply, *mutatis mutandis*, to "self-identification" views of manhood and womanhood. It's clear how they apply to obviously circular definitions of the kind one finds in Bettcher (2017, 396): to be a woman is to sincerely self-identify as a woman. But these concerns also apply to more complicated analyses, for example that of Jennifer McKittrick (2015), who says that to be a woman (in a particular context) is to have a cluster of behavioral dispositions coded as feminine in that context. Just as a circularity threat arises for Jenkins with regard to "norms of femininity," for McKittrick the circularity threat arises when we wonder what it is for a behavioral disposition to be coded as feminine. Presumably that means something like: the disposition is widely considered normative for women. And in that case the circularity concern becomes salient.

A fifth option has been proposed to me a few times, by audience members and by anonymous referees, so it's worth mentioning here. On this proposal, we expunge from Jenkins' definition any mention of "feminine gender norms" or "norms about women," and instead just *list* these social norms, without referencing females or women: shave your legs; have sex only with men; etc. (choose your preferred examples). And then, the proposal continues, to be a woman is to take these norms to be about you, to be for you. And haven't we thereby escaped vicious circularity, and the inclusion problems mentioned above?

I think not, since this fifth proposal has two problems. The first is that it's left mysterious why we chose *that* list of norms to put in the definition, rather than any of the other innumerable many combinations of norms that we might have chosen. Why is it thinking that one is supposed to shave her legs, etc. that makes one a woman, rather than, to pick a norm at random, thinking that one is supposed to serve arrest warrants, and read Miranda Rights to those served? The answer, of course, is obvious: we chose *that* list of norms for our definition of womanhood because *those are norms about women*. We can see this by imagining the shape of this fifth proposal if gender norms were to change. Currently, in many cultural contexts, there remains a fairly strong feminine gender norm about shaving legs. But it's eroding. Imagine that, in fifty years, that norm no longer exists. At the present time, the proponents of this fifth proposal say that "Shave your legs!" is among the norms relevant to womanhood. But they wouldn't say that in fifty years, for a cultural context in which that feminine gender norm is absent. And this suggests that, *really*, proponents of this fifth proposal are operating with a concept of womanhood that makes reference to norms about *women*. In that case, our concerns about circularity remain, and have been merely hidden backstage, as it were.

Secondly, even if this proposal offered a satisfying, non-circular definition of womanhood, it plausibly would not escape the inclusion problems mentioned above in the discussion of the second and third options. And that's because, in order for trans women to meet this definition of womanhood *sensibly*, it would have to be the case that these listed norms are in fact norms about them, norms for them. For example, I can't rationally take the norm *serve arrest warrants and read Miranda rights to those served* to be about me, since it's about law enforcement officers, and I know I'm not a law enforcement officer. Similarly, norms like *shave your legs* are about a class of people, they are for a particular group of people. And one can sensibly take those norms to be about oneself only if one is in that group.²³ So, in order to determine whether trans women can sensibly meet this proposed definition of womanhood, we'll have to first uncover whether these listed norms are norms for them. Since these are norms for women, this will

²³ Of course, one might have a justified false belief that one is a member of a group, e.g. as a result of a mistake or a delusion. But presumably Jenkins and others don't want to say that trans women can be *bona fide* women only as a result of a mistake or a delusion. That's a kind of marginalization.

require us to reflect on the meaning of “woman,” in which case this fifth option ultimately collapses into one of the previous four proposals.²⁴

The upshot of all this is that either Jenkins’ proposed definition of womanhood is unintelligible, or it excludes a substantial number of trans women. Either way, by her own lights, Jenkins’ proposed definition should not be the one we land on if we’re engaged in Ameliorative Inquiry.

Setting those difficulties aside, one may still wonder whether Jenkins’ proposal is so revisionary that we’re no longer using our traditional gender concepts anymore, but instead introducing new homonyms. And similarly with Haslanger’s proposed definition. More on this in the next section, along with reasons to think Ameliorative Inquiry is incoherent as well as impossible to complete (in a satisfactory, trans-inclusive way). We’ll see that trans-inclusive feminists have good reason to part ways with Ameliorative Inquiry, if they can.

Ameliorative Inquiry is Incoherent and Impossible

In this section, we’ll lay out reasons to think that (i) considered as a set of doctrines, there’s a contradiction within Ameliorative Inquiry, rendering it incoherent, and (ii) considered as a methodology, Ameliorative Inquiry is either impossible to complete, due to circular definitions of crucial terms, or else it cannot be completed satisfactorily, i.e. in a way that results in trans-inclusive definitions.

Let’s first explore the charge that Ameliorative Inquiry is incoherent. To begin, let’s clarify some terminology involved in the debate. *Concepts* are mental representations of items in the world, something like mental file folders, labeled with words from natural language and stuffed with information about their referents. A *conception* is that information we associate with some concept, the scraps we put into our mental file folders.

Now, the tension I want to point out has two parts. On the one hand, practitioners of Ameliorative Inquiry (e.g. Haslanger, Jenkins, et al.) don’t mean to change the subject. We’ve been talking about men and women for millennia,

²⁴ Perhaps one might successfully excise norms from their cultural context, and give them a new target by going through a little ceremony like so: “I hereby declare that the following norms are about me!” And then one lists the norms, which were about a group one doesn’t belong to, but which perhaps are now about oneself. However, if someone declares that, according to her, the norm “Do X!” applies, then her performance implies that she thinks she ought to do X. If she undertakes this performance with no intention of doing X, then she has acted in bad faith. And yet, as Jenkins herself points out, many women (and men) flout gender norms. They know the gender norms apply to them, and yet have no intention of abiding by them. And surely many trans women are the same. But then, this proposal we’re now considering would entail that these trans women are acting in bad faith when they identify as women. Surely this is a bad result, by the proposal’s own standards.

and they mean to continue the conversation without changing the subject matter. Haslanger wants to “provide us with an account of our social terms.”²⁵ And Jenkins says, “Over the last few decades, feminist philosophers have sought to develop an analysis of gender concepts, and of the concept *woman* in particular.” To those who say “Caitlyn Jenner is not a woman,” e.g. Jenkins replies, “No, you’re wrong.” But “Jenner is a woman” contradicts “Jenner is not a woman” only if the same gender concept features in both. Otherwise, it’s a merely verbal dispute. And that would be a bad result, from the perspective of Haslanger, Jenkins, and others.

But, on the other hand, we have the second half of this tension: these theorists provide surprising new *conceptions* that, they say, are to be associated with our familiar gender *concepts*. These innovative new conceptions are the outputs of Ameliorative Inquiry, of conceptual engineering. When these theorists inquire into the meaning of our gender terms in this ameliorative way, they’re not guided only by how the word is actually used. They also consider how it *should* be used. For example, here’s how Jenkins (*ibid.*, 395) puts it: “Ameliorative inquiries thus make use of normative inputs...Ameliorative analysis is not bound to comply with our ordinary understanding or use of a concept: the target concept may be revisionary, provided that it furthers the goals guiding the analysis.”²⁶ And the conceptions they offer are intentionally revisionary. They *intentionally* depart from the way the concepts are ordinarily used.

The problem is that this intentional revisionism is apt to originate new concepts, rather than modify the conception of our existing concepts. In other words, we can’t use gender concepts in an intentionally revisionary way that is also non-originating. Ameliorative Inquiry necessarily changes the subject, resulting in merely verbal disputes. And this is in tension with its commitment *not* to change the subject, but to remain on the subject of gender, of womanhood.

For an example of an originating use of a concept, suppose I name my daughter “Anscombe.” In doing so, I’ve borrowed the name *word*—in the thin sense of “word,” i.e. the sound-type and inscription-type—the label on my Anscombe mental file folder. But, by intentionally departing from the ordinary use of that name, I’ve originated a new concept, a new mental file folder, which I fill up with information about my daughter, and *not* information about the famous philosopher. And we can prove this: you can say “Anscombe is a famous philosopher” and I can say “Anscombe is not a famous philosopher” and we needn’t thereby contradict each other, even if we’re using every concept other than *Anscombe* univocally. It’s a merely verbal dispute.

²⁵ It’s clear that by “terms” here she doesn’t mean words in the thin sense, i.e. those sound- and inscription-types. She’s not interested in the syntax or etymology of our gender *words*. She’s interested, rather, in gender itself, and our corresponding concepts.

²⁶ And Haslanger characterizes Ameliorative Inquiry like so (2012, 376): “Ameliorative analyses elucidate ‘our’ legitimate purposes and what concept of F-ness (if any) would serve them best (the target concept). Normative input is needed.”

A nice explanation of this phenomenon comes from the theory of concept individuation developed by Michael Tye and Mark Sainsbury (2012, 42). On their view, concepts – like words – are individuated by their origins. And whether a concept is used in a non-originating way, or a new concept is originated, is determined by the user’s degree of deference to previous uses:

Non-originating uses [of concepts] are marked by both of the following features: 1. The use involves deference to other uses, by the same subject or other subjects. 2. The use makes possible the accumulation of information from other uses, by the same subject or other subjects... One characteristic of originating uses is the absence of any such conformist requirement.

In using “Anscombe” to baptize my daughter, I do *not* defer to any other uses, either by others or to my own previous uses. And, in my new name-using practice, information about the name’s referent is *not* absorbed into the common use of “Anscombe.” Rather, information about the referent of my new “Anscombe”-using practice – that is, information about my daughter – is compiled separately from the information we compile in our ordinary “Anscombe”-using practice – that is, information about the philosopher. Hence, it is an originating use; the result is a new concept. Deference makes the difference.²⁷

Similarly, suppose, following Haslanger and Jenkins, we use “woman” in an intentionally revisionary way, knowingly departing from the ordinary usage, consciously not deferring to the standard use. As with the “Anscombe” case just

²⁷ Notice the crucial role of deference in cases that are meant to pump semantic externalist intuitions. For example, in Tyler Burge’s famous case, we consider a man who mistakenly thinks he has arthritis in his thigh. This is impossible, since arthritis is an ailment of the joints. But why do we suppose that he *falsely* believes he has arthritis in his thigh, rather than that he *truly* believes he has tharthritus in his thigh, where tharthritus is an idiosyncratic concept of an ailment that manifests like arthritis, but which may also be in the thigh? The answer is deference. When Burge imagines the patient is informed by his doctor that arthritis is an inflammation of the joints, “[t]he patient is surprised, but relinquishes his view and goes on to ask what might be wrong with his thigh.” This response manifests his disposition to defer, and fuels our externalist intuitions in Burge’s case, i.e. that he’s not using an idiosyncratic concept (defined by the conception inside his skull), but rather that he has an incomplete grasp of a concept shared by the larger communal linguistic practice to which he defers. Ameliorative Inquiry lacks this disposition to defer to the larger communal linguistic practice, and is explicit about its revisionary intentions.

An anonymous referee helpfully pointed out that deference comes in degrees, and so perhaps it is hasty to think that ameliorative inquirists exhibit no deference to the larger communal linguistic practice. I agree that deference comes in degrees. However, if one intentionally departs from the communal use of a word, and if one persists in this use knowing full well that it is revisionary, then, no matter what else is going on, one does not defer *enough* to the communal linguistic practice to join it. And this is how things are in the “Anscombe” case above, how things would have been with Burge’s Arthritis Man had he resisted correction, and how things are with Ameliorative Inquiry.

mentioned, though we borrowed the *word*, we've originated a new concept. And, this being the case, "Jenner is a woman" and "Jenner is not a woman" need not contradict, and we'd lose any ground to criticize the users of the original gender concepts as saying something false. It's a merely verbal dispute, and, even if this revisionary use catches on and produces valuable social and political ends, it does so at the cost of changing the subject away from womanhood.²⁸

Haslanger is sensitive to this concern, and she admits (2000, 34) that "it isn't entirely clear when a project crosses over from being explicative to revisionary, or when it is no longer even revisionary but simply changes the subject." She goes on to propose a semantic condition on the appropriation of existing terminology (ibid., 35): "the proposed shift in meaning of the term would seem semantically warranted if central functions of the term remain the same, for example, if it helps organize or explain a core set of phenomena that the ordinary terms are used to identify or describe." She makes it clear, in a footnote on that sentence, that what she's proposing here is not a criterion for sameness of meaning. Indeed, she's considering when a term like "woman" should *shift* its meaning, i.e. come to be associated with a new, revisionary concept.²⁹

²⁸ This is so even if we take Haslanger and Jenkins to be engaged in what has become known, following Burgess and Plunkett (2013) as "conceptual ethics." Even Burgess and Plunkett take Haslanger to be introducing new *concepts* and then attaching old *words* to them (ibid., 1094): "[Haslanger] introduces concepts to pick out the relevant properties and proposes we lexicalize them with extant terms: 'race' and 'gender'. Her project therefore involves conceptual innovation, which is often a crucial part of conceptual ethics." This origination of new concepts produces ambiguity. And while that conceptual innovation, if it catches on, may well accomplish valuable social and political ends, the point here is simply that it does so at the cost of changing the subject. Consider also David Plunkett's (2015, 841) recent work on "metalinguistic negotiation," which he characterizes thusly: "If we think of the meanings of terms as the concepts that they express, then we can put this thought as follows: each speaker has a preferred concept for using in the context in hand, and they are advocating for that view." Just so. The Ameliorative Inquist appropriates existing terms ("man," and "woman") to express her favored concepts. Insofar as these concepts are intentionally revisionary, they are *new* concepts, and she has introduced ambiguity, and changed the subject. Plunkett (ibid., 864) compares Haslanger's project of ameliorative inquiry to a compatibilist's project of replacing the obviously incompatibilist folk concept of free will with another, new concept that is compatible with determinism, and admits that this process "will involve a change of subject from the original pre-theoretical one that we were talking about when we originally used the term 'free will'." I'm arguing for the analogous point here, about gender terms.

²⁹ For a related, more recent discussion, see Haslanger (2005), especially where she says, in note 11: "It might be useful to see this by analogy with other terminological developments in science. Although our understanding of and even our definition of 'atom' has changed over time, it is plausible that there is something worthwhile we have been and continue to be talking about." I believe that she intends to apply the same **Staying on Topic** criterion in this case as well. Also see Joshua Glasgow's (2009, especially 126ff) discussions of what he calls "revisionism" versus "substitutionism" in the philosophy of race.

And so Haslanger admits that Ameliorative Inquiry produces new meanings. But Haslanger thinks that, so long as we're warranted in appropriating the existing *term* "woman," we haven't changed the subject, even if we come to use the term with a new meaning. And we're warranted in appropriating existing terminology if the new use of the term will continue to help organize or explain a core set of phenomena that it was previously used to identify or describe. However, there's a serious problem with Haslanger's proposed condition here: it's either trivial – in which case it can't help us settle whether her revisionary gender concepts change the subject – or, if it's not trivial, then it's false, and in any event Haslanger's revisionary gender concepts don't meet her proposed condition for staying on topic. It's hard to settle the triviality question, because of the vagueness of her proposed condition – due to phrases like "core set," "help organize and explain," etc. Let's put the condition a little more carefully:

(Staying on Topic) For any concept C1 expressed by some term T, and any concept C2, C1 and C2 are about the same topic, and so we're warranted in using T to also express C2, **IF:** C2 helps organize or explain a core set of phenomena that C1 is used to identify or describe.

From one angle, this condition looks trivial. In order to meet the condition for staying on topic, it must be that the new concept and the old concept are about the same "core set" of phenomena. But that just sounds like another way of saying that they're about the same subject. So, in that case, in order for us to know whether we've stayed on topic according to **Stay on Topic**, we must already know whether we've stayed on topic. And that's not helpful.

To avoid the problem of triviality, we might loosen our understanding of being about the same "core set" of phenomena, so that it's possible for two distinct concepts, with two distinct meanings, on two distinct subjects, to nevertheless be "about the same core set of phenomena." But, in this case, Haslanger's proposed **Staying on Topic** condition is vulnerable to refutation by counterexample. To see this, notice that the history of science is littered with examples of theoretical terms that were eliminated – and that we would not have appropriated without a change of subject – despite the fact that the new terms we introduced plausibly still organized or explained a core set of phenomena that the eliminated term was used to identify or describe. Take "miasma" and "germ," for example. The miasma theory of disease posited the existence of miasma – a noxious "bad air" – as the cause of diseases like cholera, malaria, and the plague. The miasma theory was eventually supplanted by our modern theory of germs, where our concept of a germ also "helps organize and explain" that same "core set" of phenomena that miasma was meant to organize and explain – cholera, malaria, etc. Yet the corresponding conceptions were so different that appropriation of "miasma" to refer to germs was not semantically warranted. Despite both theories being about the same "core set" of phenomena, not even

the appropriation of “miasma” could have prevented a change of subject.³⁰ So, if **Staying on Topic** is non-trivial, the condition is not actually sufficient; there are counterexamples. And so **Staying on Topic** can’t help Haslanger argue that her revisionary gender concepts don’t change the subject.

But even if **Staying on Topic** were non-trivial and also not false, it still couldn’t help Haslanger explain why her revisionary gender concepts don’t change the subject. And that’s because her revisionary gender concepts plausibly don’t meet the sufficient condition laid out in **Staying on Topic**. After all, the traditional use of “woman” – the “manifest concept” that Haslanger aims to debunk and replace – picks out adult human biological females. The phenomena associated with that conjunction of properties – being an adult, being a human, being biologically female – are the “core set of phenomena” that “woman” is traditionally used to identify and describe. Revisionary gender concepts (such as Haslanger’s) pick out different conjunctions of properties, and therefore they plausibly “organize and explain” a different “core set of phenomena” than the traditional concept of womanhood. That’s the upshot of the counterintuitiveness worry of revisionary gender concepts; that’s what makes them revisionary.

To take just one example, imagine that gender utopia in the future, where there is no more subordination on sex-marked grounds, even though there are still males and females. All the “core phenomena” that the traditional concept of womanhood identifies and describes are still present: there are still adult human biological females. Yet, according to Haslanger’s revisionary definition of “woman,” there are no women, since there are no people subordinated on sex-marked grounds. And so the “core phenomena” of sex-marked subordination that Haslanger’s definition attempts to “organize and explain” are absent, in this situation. This shows that the “core phenomena” associated with Haslanger’s proposed concept is distinct from the “core phenomena” associated with the traditional concept. So, **Staying on Topic** will *not* rule that Haslanger has stayed on topic and is warranted in appropriating “women.” So, even if **Staying on Topic** were non-trivially true – and we’ve already seen reason to believe it’s either trivial or false – it still will not help Haslanger avoid the charge that she’s changed the subject. And likewise for other revisionary products of Ameliorative Inquiry, e.g. Jenkins’ norm-relevancy account.

Summing up this first concern, these two commitments of Ameliorative Inquiry stand in tension:

³⁰ The account of concept individuation from Tye and Sainsbury above gives a nice explanation of why there are two concepts here, of why our use of “germ” originated a new concept, rather than modifying the conception associated with the concept of miasma: *lack of deference* to the prior “miasma”-using practice. The germ theorist contradicts, and does not defer to, the miasma theorist. This is the same lack of deference to previous linguistic practice that characterizes revisionary gender concepts.

- A. Ameliorative Inquiry gives us an account of our existing gender concepts. It does *not* originate new concepts, thereby introducing ambiguity and merely verbal disputes.
- B. Ameliorative Inquiry gives us accounts that are intentionally revisionary, knowingly departing from the way our gender terms are ordinarily used, with an eye toward advancing the cause of social justice.

Therefore, Ameliorative Inquiry is incoherent. These two commitments of Ameliorative Inquiry cut against each other.

The second problem I'd like to point out is that Ameliorative Inquiry, understood as a *method*, is impossible to perform. Or, at the least, it's extremely difficult to see how the output of Ameliorative Inquiry could solve Jenkins' Inclusion Problem, i.e. to see how Ameliorative Inquiry could ever be completed in a satisfactory, trans-inclusive way.

The problem has to do with Ameliorative Inquiry's "normative inputs." As Haslanger says (*ibid.*, 366), Ameliorative Inquiry "seeks to identify what legitimate purposes we might have (if any) in categorizing people on the basis of... gender, and to develop concepts that would help us achieve these ends." What sort of goals and purposes might these be? She says elsewhere (*ibid.*, 240) "For the purposes of a critical feminist inquiry, oppression is a significant fact around which we should organize our theoretical categories," and she then alludes to "the feminist... values guiding our project..." She says (*ibid.* 226-7): "At the most general level, the task is to develop accounts of gender... that will be effective tools in the fight against injustice." She then goes on to say this broad project is guided by four concerns, including the need for an account that takes seriously "the agency of women," and which will "aid feminist... efforts to empower critical social agents." She also hopes (*ibid.* 240) that the analysis of gender can help us "identify and critique broad patterns of... sexual oppression."

Here's the problem. What makes our inquiry an ameliorative inquiry is that we set aside – we bracket off, we ignore – the project of analyzing the *ordinary* use of the term "woman," that manifest concept, in order to reflect upon certain "normative inputs," to consider what our concepts *should* be. And, to understand these normative inputs, we reflect on the project of *feminism*, its goals and values. Yet surely the values and goals of feminism will be defined partly in terms of womanhood, as we saw in some of the quotations of the last paragraph. In Ameliorative Inquiry, Haslanger says, we're looking to develop concepts that take seriously the agency of *women*, and which aid *feminist* efforts.

The problem, of course, is that the procedure of Ameliorative Inquiry asks us to set aside the ordinary understanding of womanhood in order to reflect on certain normative inputs, which inputs are given to us partly in terms of womanhood. And so we're at a sticking point in the process, a gap in the method. How shall we understand feminism and womanhood while doing

Ameliorative Inquiry, *before* we've developed revisionary gender concepts in order to advance the cause of feminism and the condition of women? If we leave these terms *undefined*, then Ameliorative Inquiry, considered as a method, becomes unintelligible, and therefore impossible to complete. If we say instead that we're meant to understand "woman" using the target concept developed as a result of Ameliorative Inquiry, this will be no help, since we cannot complete the process and develop that concept until we understand the values and goals of feminism, which requires understanding "woman." We're commanded to pull ourselves up by our own bootstraps.

To appreciate the impossibility of this task – and to assist you in the bracketing-off process that Ameliorative Inquiry recommends – try this project instead: I have a new term I'd like to define, "blarg." I'd like to define it in a way that will best advance the cause blargism. By that I mean I'd like to combat all unjust discrimination and oppression of the blargs. Now, please tell me how to best define "blarg" in order to do this. Seems impossible, right? How could you define "blarg" under these conditions without first knowing what blargism is, i.e. without already knowing what the blargs are?

Therefore, if during Ameliorative Inquiry we really are meant to (i) set aside ordinary understandings of womanhood while also (ii) reflecting upon the values and goals of feminism (defined in terms of womanhood), the project is impossible to complete, just like the "blarg" project. But perhaps both of those assumptions can be disputed. Perhaps, when considering these normative inputs to Ameliorative Inquiry, we're meant to understand "woman" in the ordinary *adult human biological female* way. When Haslanger lays out the four concerns guiding the broad project of Ameliorative Inquiry, for example, she does mention *females* (ibid. 266-7).³¹

But, on this alternative, it's doubtful that the result of Ameliorative Inquiry will be trans-inclusive. For consider that, on this alternative, we're meant to ask ourselves what definition of "woman" will best promote justice for *biological females*. Focusing our attention in this way requires us to ignore the oppression and violence directed toward biologically male trans women. Those legitimate concerns simply aren't on the table, on this way of doing Ameliorative Inquiry. But then by what path of reasoning would we arrive at a definition of woman that includes those trans women? It's doubtful that we would, and obscure how we even could.

A second proposed fix might be this: while we do bracket off our ordinary understanding of "woman" during Ameliorative Inquiry, we also purge the normative inputs of our Inquiry of any mention of women or feminism. Instead, in doing Ameliorative Inquiry, we ask ourselves only what definition of "woman" would best serve our *very general* project of social justice. But to

³¹ And this is how Jennifer Saul (2006, 126) reads Haslanger, concluding "The (legitimate) point of gender concepts, then, is to help us to fight oppression *on the basis of sex*" (emphasis mine).

appreciate how difficult this task would be – and to avoid inadvertently falling back into framing the inquiry in terms of actual womanhood – try a new task similar to the one above: I have a novel term I’d like to define, “blorg.” I’d like to define it in a way that will best advance the cause of social justice *generally*. By that I mean I’d like to combat all unjust discrimination on the basis of religion, sex, gender, race, ethnicity, culture, class, sexual orientation, age, disability, etc., etc. *All of it*. Would you please tell me how to best define “blorg” in order to do this?

Now, what are the odds that the definition you’ll arrive at for “blorg” is anything in the neighborhood of definitions of “woman” that we’ve considered in this paper? It’s hard to see how you’d get from here to there, on this alternative protocol for Ameliorative Inquiry. Worse, our previous problem reemerges, since even this very general normative input still requires that we understand womanhood, in order to understand the need to combat unjust discrimination on the basis of *gender*. (If we leave out that value of combating unjust gender discrimination, then the concerns of women are completely left out of this normative input, and the odds that we’ll arrive at a sensible and useful definition of “woman” wither toward zero.)³²

I present all this not as a knockdown argument against Ameliorative Inquiry, but rather as a request for clarification. Perhaps, though I can’t see it, there is a way to solve this problem. I’d be grateful to learn it. But, until then, trans-inclusive feminists have cause to be concerned about saddling themselves with the incoherence of Ameliorative Inquiry, and therefore reason to look beyond Ameliorative Inquiry toward a more promising project, if they can.

Conclusion

³² Recently, Robin Dembroff (forthcoming) proposes this method of Ameliorative Inquiry: “On this different approach – the ‘imitation’ approach – classification practices can be revised based upon those that already exist within other communities. Given that operative gender kinds vary across contexts, looking at other communities will reveal alternative gender kinds and corresponding classification practices. From here, those in one context can attempt to revise their operative gender kinds by mirroring or otherwise imitating the structures and practices that already exist in other contexts.” There are two problems with this “imitation approach.” First, if I’m really meant to set aside or bracket off my concept of “woman” that I wish to revise, then I won’t be in a position to know which communities are using that concept and which aren’t, in which case I won’t know which communities are candidates for imitation. (Of course, I could easily tell which communities are using the *word* “woman,” but that’s not enough to know whether they’re using the selfsame concept that I wish to revise.) Second, even if I could identify communities that are using the same concept as I am, if I really do bracket off my concept of womanhood that I wish to revise, then, when it comes time to evaluate how well the uses of that concept in these communities accomplish the goals of feminism (cashed out in terms of womanhood), I will face the same problems I just discussed in the paper. So even Dembroff’s “imitation approach” to Ameliorative Inquiry will be impossible to complete, at least in a trans-inclusive way.

Ameliorative Inquiry, with regard to “woman,” balances our ordinary use of gender terms against the instrumental value of proposed definitions in achieving the goals of feminism. Sally Haslanger’s proposed analysis of “woman” departs substantially from our ordinary concept, and also excludes many trans people. Katharine Jenkins’ proposed analysis of “woman” is an even further departure from our ordinary use, and, as I’ve argued, either is unintelligible or it too excludes many trans people.

We then saw reasons to believe that the entire enterprise of Ameliorative Inquiry, understood as a set of doctrines, is incoherent: we cannot *both* continue using our old gender concepts *and also* use them in an intentionally revisionary, non-conforming way. Ameliorative Inquiry inevitably introduces ambiguity – new homonyms – and thereby results in merely verbal disputes, a change of subject. And we gained reason to believe that, with regard to “woman,” Ameliorative Inquiry understood as a method or procedure that requires grasping the cause of feminism – of *women* – will either be impossible to complete due to unintelligibility (if, in the normative inputs to Ameliorative Inquiry, “woman” is left undefined, or is defined in a viciously circular way), or will be impossible to complete, at least in a way that includes all or even most trans women (if, in those normative inputs, “woman” is defined as adult human biological female). These are the internal problems with revisionary gender concepts that I spoke of in my title, and they ought to propel trans-inclusive feminism in a new direction, away from the project of Ameliorative Inquiry.

There is urgency here, since, unless it can do without Ameliorative Inquiry, the project of trans-inclusive feminism is incoherent.³³

References

Andler, Matthew Salett (2017). "Gender Identity and Exclusion: A Reply to Jenkins," *Ethics* 127(4): 883-895

Barnes, Elizabeth (forthcoming). "Gender and Gender Terms," *Nous*, 1-27

³³ Does the project of trans-inclusive feminism require Ameliorative Inquiry? Well, that depends on whether the project of trans-inclusive feminism requires that "woman" (or "female") apply in an unqualified way to trans women and "man" (or "male") apply in an unqualified way to trans men. Given that this usage intentionally departs from the traditional definitions of those terms, and requires conceptual engineering motivated by normative inputs, the project of trans-inclusive feminism *would* require Ameliorative Inquiry. It would therefore be incoherent. But if trans-inclusive feminism can find a way to advance its cause without conceptual revision, it may avoid the incoherence noted in this paper.

Exploring that possibility is beyond the scope of this paper, though the reader may be interested in a recent project from Barnes (forthcoming) and another from Robin Dembroff (forthcoming). Both of these projects will be attractive only to those who are willing to reject what Dembroff calls "the Real Gender assumption": that gender classifications should track the gender kind membership facts. According to Dembroff, gender kinds are social kinds, and *not* features of mind- or language-independent reality. Barnes agrees, saying (ibid., 17) "There aren't any context or language-independent facts about what it really is to be a woman, a man, a nonbinary person, etc." So, if you, like me, think that some people *really are* men, or women, or neither, independent of our minds and language, then these projects aren't for you. But despite throwing overboard the reality of gender, these projects don't escape the problems discussed in this paper. And that's because both projects require or propose *some* understanding of gender terms. And this understanding leads to a familiar dilemma. For example, Barnes says (ibid., 3) "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'." Even if this is true, giving a metaphysics of gender cannot "come apart" from *understanding* what "woman" means, since it would be impossible to determine what explains womanhood if we have no idea what womanhood even is. (Could we determine what explains blarghood, having no idea what it is to be a blarg? I can't see how.) But, there's a dilemma: this understanding of womanhood will be either traditional or revisionary. The former is trans exclusive, and the latter has the internal problems we've seen in this paper. Similarly, when Dembroff argues that we should modify the membership conditions of social kinds for the sake of social justice, the social kind concepts advocated for will be either traditional or revisionary. And, so, the same dilemma that confronts Barnes also arises for Dembroff: either trans exclusion, or some internal problems with revisionary gender concepts.

de Beauvoir, Simone (1973). *The Second Sex* (New York: Vintage Books)

Beeman, Brett and Susan Rankin (2011). *The Lives of Transgender People* (Columbia University Press)

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), 98-120

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 (2017). "Through the Looking Glass," in *The Routledge Companion to Feminist Philosophy*, eds. Ann Garry, Serene J. Khader, Alison Stone (Routledge), 393-404

Burge, Tyler (1979). "Individualism and the Mental." In *Midwest Studies in Philosophy, IV*, edited by Peter French, Theodore Uehling, and Howard Wettstein (Minneapolis: University of Minnesota Press), 73-121.

Burgess, Alexis and David Plunkett (2013). "Conceptual Ethics I," *Philosophy Compass* 8(12): 1091-1101

Butler, Judith and Liz Kotz (1992). "The body you want: Liz Kotz interviews Judith Butler," *Artforum International* 31(3): 82-89

Chu, Andrea Long (2018). "On Liking Women," *n+1 Magazine* 30

Corvino, John (2000). "Analyzing Gender," *Southwest Philosophy Review* 17(1): 173-180

Dembroff, Robin (forthcoming). "Real Talk on the Metaphysics of Gender," *Philosophical Topics*

Glasgow, Joshua (2009). *A Theory of Race* (Routledge)

Grossman, Arnold et al. (2005). "Parents' Reactions to Transgender Youths' Gender Nonconforming Expression and Identity," *Journal of Gay & Lesbian Social Services* 18(1): 3-16

Haslanger, Sally (2000). "Gender and Race: (What) are They? (What) do We Want Them to be?" *Noûs* 34(1): 31-55

Haslanger, Sally (2005). "What Are We Talking About? The Semantics and Politics of Social Kinds." *Hypatia* 20(4): 10-26

Haslanger, Sally (2008). "A Social Constructionist Analysis of Race" in B. Koenig, S. Lee, and S. Richardson, eds., *Revisiting Race in the Genomic Age* (New Brunswick, NJ: Rutgers University Press), 56-69

Haslanger, Sally (2012). *Resisting Reality* (New York: Oxford University Press)

Jenkins, Katharine (2016). "Amelioration and Inclusion: Gender Identity and the Concept of *Woman*," *Ethics* 126: 394-421

McKittrick, Jennifer (2015). "A dispositional account of gender," *Philosophical Studies* 172 (10): 2575-2589

Mikkola, Mari (2009). "Gender Concepts and Intuitions," *Canadian Journal of Philosophy* 39(4): 559-583

Moi, Toril (1999). *What is a Woman?* (Oxford University Press)

Plunkett, David (2015). "Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology of Philosophy," *Inquiry* 58: 828-874

Sainsbury, R.M. and Michael Tye (2012). *Seven Puzzles of Thought and How to Solve Them: An Originalist Theory of Concepts* (Oxford University Press)

Saul, Jennifer (2006). "Gender and Race," *Proceedings of the Aristotelian Society*, Supplementary Volume 80: 119-143

Serano, Julia (2007). *Whipping Girl: A Transsexual Woman on Sexism and the Scapegoating of Femininity* (Berkeley, CA: Seal Press)

Stahl, Titus (manuscript). "What is Immanent Critique?" Available at: <https://philarchive.org/archive/STAWII-3>