

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Philosophy

2017

Categories We (Aim to) Live By

Dembroff, Robin

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp014q77fv007>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

CATEGORIES WE (AIM TO) LIVE BY

ROBIN A. DEMBROFF

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF

PHILOSOPHY

ADVISER: SHAMIK DASGUPTA

AND SARAH-JANE LESLIE

SEPTEMBER 2017

© Copyright by Robin A. Dembroff, 2017.

All rights reserved.

Abstract

We want to know who should be counted as belonging to various social categories. Who should be counted as black? ...as a woman? ...as gay? In both everyday and academic contexts, it is widely assumed that these questions turn on metaphysical analyses of what makes someone black, a woman, and so on. That is, it is assumed that we should count someone as (e.g.) a woman just in case they satisfy sufficient conditions for having the property *woman*.

My dissertation argues that this assumption is wrong: whether someone should count as a woman turns not on whether they satisfy the correct *metaphysical* analysis of what it is to be a woman, but on *political* and *ethical* considerations about how we ought to treat each other.

I begin by examining the case of gender, and argue that looking to metaphysical theories of gender to settle questions about who should be counted as a man or a woman misses the point. What matters for determining ethical gender ascriptions are normative questions about how we ought to treat others, and not facts about who is a man or a woman. From here, I develop an important theoretical upshot of this view: that there can be metaphysical truths the assertion of which tends to both implicate and reinforce social injustices. These truths highlight places where we can both critique and improve our social categories. In my final chapter, I apply this method to sexual orientation. There, I argue that common understandings of sexual orientation are metaphysically ambiguous and ethically problematic, and I propose a novel and revisionary account of sexual orientation.

Acknowledgments

I started graduate school at the University of Notre Dame in 2012. While there, I received generous advice from many faculty members, and especially Paddy Blanchette, Michael Rea, Meghan Sullivan, Peter van Inwagen, and Fritz Warfield. I remember with special fondness when, after announcing that I should read more about a philosophical problem – and I’d already read a lot! – Peter (affectionately) growled at me, “Don’t go read. Go think!” I remain grateful for this advice, and all the support and guidance I received at Notre Dame.

After transferring to Princeton in 2014, my research began to focus on feminist philosophy and social ontology. This research has hugely benefitted from countless interactions with my advisers, Shamik Dasgupta and Sarah-Jane Leslie. Many of the lessons I learned from them guided this project. I’ll mention just two: the confidence to pursue new questions (from Shamik), and a healthy skepticism about empirically detached metaphysics (from Sarah-Jane). This project comes out of combining these lessons with general excitement about metaphysics – modeled for me by Shamik, Boris Kment, and Gideon Rosen – as well as passion for feminist philosophy and social ontology.

My thinking about these latter topics would not be nearly as clear or informed as it is today if not for Elizabeth Barnes. Elizabeth has been unbelievably generous with her time, providing feedback on each part of this project, as well as more pieces of professional and personal advice than I care to count. In addition, her written work – along with that of Ásta, Sally Haslanger, Esa Diaz-Leon, Katharine Jenkins, Charlotte Witt, and many others – showed me how to do careful, socially engaged feminist ontology. I am incredibly grateful to count Elizabeth as a friend and as a mentor.

I want to also thank my philosopher-friends, many of whom gave me feedbacks on parts of this project, and all of whom have sharpened my thinking. In particular, I'd like to thank Christina Van Dyke, John Corvino, Amy Seymour, Cat Saint-Croix, Daniel Wodak, Sukaina Hirji, Brennan McDavid, Sam Preston, and Erin Islo. With deep love, I want to also express gratitude to Marilyn and Bob Adams. Pax et bonum, Marilyn – thank you for everything.

I have been grateful for copious feedback on this project, and especially for feedback from David Black, Justin Christy, Peter Finnochiaro, Vera Flocke, Rachel McKinnon, David Plunkett, Michael Rauschenbach, Father Raphael Mary Salzillo, Sami Sumpter, and Elizabeth Victor. I've also been fortunate to receive feedback from audiences at the Pacific, Central, and Eastern APA, Wayne State University, Western Washington University, Princeton, Notre Dame, Yale University, Tufts University, University of Michigan, UNC Asheville, UC Irvine's Perspectives on Gender Conference, New York SWIP, the Berkeley Social Ontology Group, PRPSEW, and the Feminist Utopias conference. As is the case with most works of philosophy, to call this project 'single-authored' does not do justice to the vast amounts of time and energy that others have generously donated to help me develop these ideas.

I'd like to also thank my non-philosophy-friends, and especially Ruthie Birger and Rebecca Neill, for their endless support and love, and for always having my back (even when I'm wrong).

Finally, I'd like to thank my favorite human and best friend: Erica Tibbetts. I won't try to capture in a few sentences all the happiness and perspective you bring into my life.

RAD

Princeton, April 2017

To Tibbs & Jack

My Home Away from the Bungalow of Distinctions

Contents

Abstract	iii
Acknowledgments.....	iv
1. Introduction	1
1.1 Oppressive Truths.....	4
1.2 Real Talk on the Metaphysics of Gender.....	8
1.3 What is Sexual Orientation?.....	12
2. Oppressive Truths	15
2.1 Introduction.....	15
2.2 Social Groups.....	18
2.3 Illustrating the Phenomenon.....	23
2.4 Oppressive Truths.....	25
2.4.1 Social Status and Injustice.....	28
2.4.2 Reinforcing Ontology via Ideology.....	35
2.5 Applications.....	41
2.5.1 Normative Implications for Ascriptions.....	41
2.5.2 Motivating the Ameliorative Project.....	45
2.5.3 Dissolving Identity Disputes.....	50

3.	Real Talk on the Metaphysics of Gender	56
3.1	Introduction.....	56
3.2	What Constrains Just Gender Ascriptions?	60
3.2.1	The Alethic View.....	60
3.2.2	The Normative View.....	68
3.3	Social Implications.....	70
3.4	Oppressive Truths.....	84
3.5	Methodological Implications.....	96
4.	What is Sexual Orientation?	101
4.1	Introduction.....	101
4.2	Methodology and Framework.....	105
4.2.1	The Engineering Project.....	105
4.2.2	Framework I: Purposes.....	108
4.2.3	Framework II: Target Concept.....	111
4.3	Bidimensional Dispositionalism.....	116
4.3.1	Preliminary Issues.....	117
4.3.2	Conceptualizing Bidimensional Dispositionalism.....	137
4.4	Conclusion and Implications.....	151
5.	Conclusion	153
	Bibliography.....	157

Chapter 1

Introduction

“Categories We (Aim to) Live By” explores the intersection of feminist philosophy and social ontology. As such, it is an exercise in using abstraction to understand and critique social experiences, and especially experiences directly related to the categorical frameworks that we live by. For this reason, I think it is important to contextualize my project by beginning with experience, and I will do so by recounting a memory – one that played a significant role in this dissertation’s development.

In December 2015, after an extended time away from my childhood home, I returned for the holidays. Sitting around the dinner table with my family one evening, I made an announcement. “I’m not a woman,” I said. “I’m genderqueer. So I want to ask you [my family] to use gender neutral language when you refer to me.” My politically and religiously conservative family did not warmly receive this announcement. One family member responded, “You can call yourself whatever you’d like – call yourself a tree toad if you’d like. But god made you a woman. You *are* a woman.”¹ I strongly voiced my disagreement. “Just because I have a certain sort of body doesn’t make me a woman.” This conversation, though it went on, didn’t improve, and everyone eventually stormed away to simmer in his, her, or their relative frustrations.

I’ve reflected a lot on this moment. In particular, I’ve wondered about what, exactly, was the source of our disagreement. For a while, I understood the disagreement to be about *what gender I*

¹ It’s important to emphasize that, for my family, I think that saying that I am not a woman but rather genderqueer sounded as absurd as saying that I am not human but rather a tree toad. While they may have some responsibility for being in this epistemic position, my focus here is not on the responsibility question.

have – that is, whether I *in fact* am a woman or not. I thought my family was getting the ontological facts *wrong* in saying that I am a woman, relying on a mistaken and outdated metaphysical picture of what makes someone a woman. I happily concluded that I was enlightened about the nature of gender and they were not. They ought to defer to *my* ontological expertise, I thought, particularly given that I had so much personally at stake in their ‘getting it right’.

I am not alone in interpreting confrontations such as this one as disputes over ontological facts. Just a cursory foray into discussions about how to classify persons into social categories – (e.g.) whether autism should count as a disability², whether Puerto Ricans should count as black³, whether transgender women should count as women⁴, or whether some heterosexual persons should count as queer⁵ – reveals that these discussions typically proceed *as if* they are fundamentally ontological disputes. That is, there appears to be a background assumption that, whatever the facts are concerning who falls within some social group, these facts constrain or even determine how we ought to classify persons. This assumption was particularly obvious in my dinner table dispute: My family assumed that I should be classified as a woman *because I am* a woman. I assumed that I should not be so classified *because I am not* a woman.

Importantly, this assumption appears not only during dinner table disagreements, but also within feminist ontology projects that explore the nature of social categorical frameworks, such as (e.g.) *gender*, *race*, *disability*, and *sexual orientation*. Generally speaking, this literature has focused on developing understandings of these frameworks that reveal them to be *socially constructed* rather than *natural*, where this typically is taken to mean that what accounts for

² See, e.g., Stevenson (2015) along with countless blogs discussing the merits of the phrase, ‘Autism is not a disability, it’s a different ability’.

³ See, e.g., Clemente (2011)

⁴ See, e.g., Burkett (2015).

⁵ See, e.g., Mortimer (2016).

persons' category membership are their social relationships, and not features that they possess by nature. (For example, on a social constructionist theory of gender, one might be a woman because one has been/is treated a particular way by others on the basis of having a female body, rather than simply because one has a female body.) These projects are immensely valuable for revealing the ways in which social forces form us, both individually and as members of a larger society, and do so in ways that are hidden under the guise of nature or essence.

Such projects typically are explicitly guided by political aims of changing commonly shared beliefs, practices and norms surrounding a social category. For this reason, they often deploy the following methodology: begin with a particular political aim for that social category (e.g., trans inclusivity in gender categories, or the elimination of racial categories) and then construct an analysis of the category that achieves this goal (e.g., a theory of *womanhood* that is trans-inclusive, or an anti-realist theory of race). While this methodology might appear free from the assumption that classification should follow ontology, the assumption become apparent whenever such projects are framed, not as theories of how these categories *ought* to be, but as they in fact *are*.⁶ Again, why think that social categories would turn out to align with our political aims? (If anything, shouldn't we expect them to misalign?) Yet here, too, we find the assumption that how we *ought* to classify persons aligns with the ontological facts about category membership.

I now question this assumption, and with it not only my initial interpretation of the dinner table dispute, but also a common project framing within feminist ontology. This dissertation

⁶ In addition to my puzzlement, these projects also tend to create suspicion and dismissal from readers not already sympathetic to the political aims. For example, imagine someone who has lived for many years with a trans-exclusive understanding of gender. Telling such a person that they've been wrong about gender their entire life, and gender really is (e.g.) determined by gender identity may lead to incredulity and possibly anger. Such 'telling' tends to come across as trying to strong-arm someone into particular political views by asserting they've been misperceiving reality. See, e.g., Nazworth (2015).

develops my criticism of this assumption, as well as the implications of rejecting it for both political and philosophical discourse. In the next chapter, I develop a framework for understanding situations in which the facts about how we should classify persons into social categories comes apart from the ontological facts about those persons' category memberships, as well as for understanding how we can reconstruct (for the better) these social categories. In the next two chapters, I apply this framework to particular gender and sexual orientation categories, drawing out the particular considerations in favor of revising these categories, and proposing a particular revision to our concept of sexual orientation. The following three sections of this introduction summarize these three chapters. In my final chapter, I discuss a common theme underlying these three chapters: the marriage of social critique and analytic metaphysics.

1.1 Oppressive Truths

Humans classify each other into an enormous number of social categories – to name just a few: *men, women, and genderqueer; black, white, Asian, and Native American; gay, queer, bisexual and straight*. We do this, as Ron Mallon nicely summarizes, for purposes including “explaining and predicting the behaviors of other individuals and groups; signaling to and coordinating with others; representing the world to ourselves; and stigmatizing, valorizing, and regulating the behavior of ourselves and others.”⁷ Given the centrality of these categories, much hangs on how persons are categorized – after all, how they are categorized guides how others interact with and evaluate them. It is unsurprising, then, that a number of ongoing controversies center on questions of social categorization, such as *who should be counted* -- that is, acknowledged and

⁷ Mallon (2016), 1.

treated – *as a woman?* ...*as disabled?* ...*as black?* ...*as a parent?* ...*as an American?*

A particular assumption is often employed on both sides of these controversies: Who we *should* count as members of some social category is constrained by who is *in fact* a member of that category. For example, those who think that autistic persons should not be counted as disabled typically claim that ‘autism is not a disability’; those who think that persons identifying as genderqueer should not be counted as women or as men typically claim that these persons ‘are not women or men’. On its surface, this rhetoric seems to assume that normative facts about how to count category membership align with ontological facts about category membership.

“Oppressive Truths” argues that there are conditions under which this assumption does not hold, and explores the argument’s implications. The argument centers around the concept of *oppressive truths*, or truths about social category membership⁸ such that *correctly* ascribing category membership – that is, counting someone as a category member according to the ontological facts about that category – leads to an unjust distribution of social status and reinforces those ontological facts. That is, the ontological facts about who belongs to a social category may be such that ascribing membership according to the ontological facts can both (i) unjustly distribute social status, and (ii) reinforce the ontological facts underlying this unjust distribution. The central idea, then, is that there can be cases where who in fact belongs to some category misaligns with who *ought* to belong to this category; moreover, in such cases, accurately ascribing membership in that category reinforces the unjust ontological boundaries around that category.

In order to motivate these claims, I break down the phenomenon of oppressive truths into two

⁸ In particular, this chapter focuses on oppressive truths as they relate to *social groups*, which I take to be social categories of persons such that a social status is associated with this classification, and this social status is conferred on someone when they are systemically ascribed this classification by others. However, I suspect the argument extends to social categories more broadly (e.g., *food, marriage, borders*).

parts, and independently motivate each part. First, how we count membership or non-membership in a social category can unjustly confer or deny the social status associated with that group. Moreover, it can do so even when our group ascriptions are true. And second, social category-term deployments (e.g., ‘woman’, ‘disabled’), including ascriptions of membership or non-membership in social categories, can reinforce the facts about which features determine membership in that category.

To start my analysis, I first describe examples of cases where the extension of persons who *should* be conferred a particular group’s social status – that is, the extension of persons who we ought to count as group members – comes apart from the extension of persons who are in fact counted as members of that group. I then consider why, in such a case, our unjust group ascriptions could be true. To defend this idea, I argue that on both extensional and intensional theories of reference, ascriptions that unjustly distribute social status might nevertheless be true. From here, I turn to the second part of the analysis. Drawing from Eric Swanson’s idea that deploying slurs can reinforce harmful background ideologies, I argue that, similarly, deploying social category terms in certain ways and contexts can reinforce harmful background ideologies. Moreover, since these ideologies establish and perpetuate social categories (and their boundaries), I argue that to reinforce these ideologies is to reinforce the ontological facts about the boundaries around social categories.

At first glance, this might seem like enough to unseat the assumption that the category ascriptions we should make are constrained by the ontological facts about category membership. But, in fact, I argue that reinforcing bad ideologies about our social categories is ubiquitous and often unavoidable. For example, we arguably would be better off without any social categories of race or of gender. But this hardly seems to suggest that we can eliminate these categories simply

by not ascribing membership in them – they are too deeply embedded in social practices.

Assuming that ought implies can, it seems something more is needed to establish that sometimes we ought to revise our ascriptions in order to revise our social categories. I argue that we ought to revise our ascription practices when there is an *available* alternative ideology about some category – that is, when there is a community with an alternative robust shared set of beliefs, norms, and practices surrounding that category – and when revised ascriptions would cue that alternative ideology.⁹

From here, I turn to the implications of my view for feminist ontology and for diagnosing ongoing identity disputes. With respect to the former, I first argue that the possibility of oppressive truths indicates that, when doing social ontology, we should not expect that an accurate metaphysical analysis of some category aligns with our normative intuitions concerning who *should* count as category members. Second, I argue that this leaves room for explicitly prescriptive ontological projects: projects that take seriously the idea that social categories in some significant sense are up to us, and propose revisions to the ideologies surrounding these categories in order to bring them into alignment with our normative intuitions.¹⁰

The framework for oppressive truths also has helpful insights for diagnosing disputes over how to count persons' social identities. Here, I focus on disputes over how to count the race or gender of persons who identify with a race or gender other than the one they were initially assigned at (or before) birth. One fairly popular position to take within these disputes is that

⁹ As I discuss in this chapter, gender is a good example here. Whereas the ideology surrounding, e.g., *womanhood* in many communities is trans-exclusive, there is a trans-inclusive alternative available ideology operative in trans-activist communities (discussed at length in Bettcher (2013)) which we can cue for the purpose of revising the ideology operative in trans-exclusive contexts.

¹⁰ This leaves open the question of how such alternative ideologies are developed in the first place. I take it the work of envisioning alternative ideologies is being done in feminist ontology, in grassroots movements, and in the social sciences. See, for example, social scientist Rogers Brubaker's (2006) *Ethnicity Without Groups*, which proposes a substantial shift in our ideology of ethnicity.

(e.g.) transgender women should be counted as women, but someone who has no African ancestry, but identifies as black (such as Rachel Dolezal¹¹) should *not* be counted as black. Often, this position is justified with an appeal to an ontological asymmetry between the cases: transgender women *are* women, but Dolezal is *not* black.¹² That is, a common approach to these cases is to argue that gender is mutable, while race is unchangeable. I suggest that my framework offers an alternative (though compatible) justification for the asymmetrical treatment of these identities: Even *if* transgender women did not belong to the social group *women*, this would be an oppressive truth that we should and can change by adopting available alternative gender ideologies. But this may not be true in the latter case: Even if someone like Dolezal is not in fact black, this may not be an oppressive truth, or one we currently could change.

Stepping back, recall the foundations of feminist ontology: to reveal the ways in which our social categories are not natural or immutable, but rather are in some important sense dependent on humans and their interactions. “Oppressive Truths” is a call to return to this foundation, which should lead us to suspect that many of our social categories not only exclude persons they should include and include persons they should exclude, but also, are our responsibility to improve in the pursuit of ameliorating oppression.

1.2 Real Talk on the Metaphysics of Gender

“Oppressive Truths” remains at a fairly high level of generality with respect to social categories. Rather than focus on any single category, it sets up a framework that applies to a variety of

¹¹ Dolezal is a woman with European American ancestry, who self-identifies as black and successfully passed as black for a number of years before being ‘outed’ by family members.

¹² See, e.g., Talusan (2015) “Dolezal identified as black, but I *am* a woman, and other trans people *are* the gender they feel themselves to be.” See also Allen (2017).

categories. In “Real Talk on the Metaphysics of Gender,” I turn my focus specifically to gender categories. Here, using the framework of oppressive truths, I provide two interrelated arguments against the claim that the gender ascriptions we should make are constrained by the facts about what genders persons in fact have.

What makes it the case that we should count (e.g.) transgender men as men, or that we should not count genderqueer persons as either men or as women? A standard response in both political and philosophical discourses is an ontological one: it is because (e.g.) transgender men *are* men, and genderqueer persons are *not* men or women. Call this the *alethic view* – the view that the gender ascriptions we *should* make are constrained by the metaphysical facts about what genders people *in fact have*. After establishing the prominence of the alethic view, both within progressive and conservative discourses concerning gender categorization, I claim that the alethic view is false, and provide two arguments in support of the *normative view*, or the view that the normative facts about gender ascriptions are *not* constrained by the metaphysical facts about what genders persons have.

The first argument is from *social implications*. This argument proceeds in two steps: I first outline some of the usual social implications – and specifically, the social statuses – that are conferred by gender ascriptions. I then argue that, if the facts about how we should distribute gendered social statuses come apart from the metaphysical facts about gender categorization, the normative facts are what should matter when considering what gender we should ascribe to someone. That is, the social implications of counting someone as a woman or a man should guide our gender ascriptions *regardless* of the ascriptions’ truth-values.

Why think this? Again, we should think back to what background ideology of gender is operating in a particular context. In many contexts, the background gender ideology – that is, the

shared background beliefs, norms, and practices surrounding gender – dictates that, for example, in order to have access to women’s resources, to be evaluated according to feminine norms, or even to have one’s behaviors and desires be interpreted as those of a woman, one must be seen as a woman. If then, one lives in a context where, in addition, the ideology says that one is a woman only if one was born with female genitalia, the social category of *woman* has been constructed in such a way that persons who ought to have access to the associated material and interpretative resources are nevertheless excluded from this category. In such a situation, I argue, it is more important to give persons access to the resources and recognition that they ought to have then to ensure that gender ascriptions align with the operative social categories.

This argument sets up the way in which gender ascriptions can unjustly distribute gender status, and moreover, could do so even if these ascriptions were true. This gives us an initial reason to think that we ought to make gender ascriptions that justly distribute these statuses, regardless of whether these ascriptions are true. But this argument is subject to a substantial worry: namely, that we generally should try to assert truths, and moreover, that saying we should count (e.g.) trans men as men regardless of whether they are in fact men sounds like a condescending way of humoring the trans community.

My second argument draws out the ways in which this objection overlooks the possibility of oppressive truths, and how language can be used to revise gender categories. This argument says, in effect, that ontological facts about gender categories may be *oppressive* in the sense that true gender ascriptions not only can reinforce gender injustice, but also *reinforce* these oppressive facts. Moreover, if these facts are oppressive, then, I argue, we can shift the ontological facts about gender categories by revising our patterns of gender ascriptions. Central to this argument is the idea that gender categories are *constitutive* social constructions –they are constituted at least

in part by shared social practices of conferring particular gender statuses on persons perceived to have certain features (e.g., female bodies). In light of this, changing our practices surrounding gender ascriptions has the potential of affecting the boundaries of gender categories, even if new practices initially conflict with dominant practices. That is, the features that provided the basis for conferring gender status are (in a significant sense)¹³ up to us -- by revising our shared beliefs and practices surrounding which features should provided the basis for these ascriptions, we can in turn revise our gender categories. As Talia Bettcher argues at length, this has been achieved in trans activist sub-communities;¹⁴ I argue that such a shift should spread to other communities even if it requires making what are, in those contexts, false gender ascriptions.

If successful, these arguments carry substantial implications for the framing of feminist ontology projects concerning gender. Generally speaking, I think these projects tend to fall into two types. The first is descriptive projects, which aim to describe the features of our gender categories. Descriptive projects, such as Haslanger (2000), Ásta (2011), and Bach (2012), often highlight the ways in which our current categories are problematic, both with respect to the social statuses accompanying gender categorization as well as the extension of these categories. In contrast, prescriptive projects take on the challenge of finding ameliorative revisions to our current categories. Prescriptive projects, such as Jenkins (2015), McKittrick (2016), Diaz-Leon (2016), and Andler (forthcoming), use philosophical tools to articulate gender categories that ameliorate some of these problematic features. However, these projects have thus far been ambiguous in framing or even framed explicitly as descriptive projects. Without the alethic view of gender, though, we have little reason to think that our gender categories are non-oppressive, making a descriptive framing for these projects quite puzzling. The normative view provides

¹³ Of course, our options are constrained by physical and social environments.

¹⁴ See Bettcher (2013).

motivation for openly prescriptive projects. Given that ontological facts about gender categories may well be oppressive truths, it is not only important to describe these facts, but also to develop alternative gender ideologies, including new concepts of gender. The project of envisioning or clarifying these new concepts – for example, concepts on which gender categorization is determined by self-identification rather than genitalia – is a valuable ontological project; it illuminates gender categories that we *could* and perhaps *should* adopt, rather than only illuminating oppressive features of current gender categories.

1.3 What is Sexual Orientation?

“What is Sexual Orientation?” is an example of one such prescriptive project, and focuses on sexual orientation categories. In it, I begin by problematizing our current understanding of sexual orientation (and the categories thereof), and then develop a concept of sexual orientation that ameliorates problematic features of our current sexual orientation categories. On this proposed understanding – “Bidimensional Dispositionalism” – sexual orientation concerns what sex(es) and gender(s) of person one is disposed to sexually engage, and makes no reference to one’s own sex or gender.

In order to identify some of the problems with dominant current understandings of sexual orientation, I argue, one needs only to look at the myriad ways in which people talk about it. For example, it is described as an “enduring pattern of emotional, romantic, and/or sexual attractions,”¹⁵ as a “trait...predispos[ing] us to experience sexual attraction,”¹⁶ and as patterns of

¹⁵ American Psychological Association (2008).

¹⁶ LeVay (2011), 1.

“sexual desires and fantasies and sexual behaviors”¹⁷ we are disposed to have under certain conditions. It also is sometimes described in terms of attraction to “men, women, or both sexes,”¹⁸ to “the same and/or opposite gender,”¹⁹ or with no reference to sex or gender at all.²⁰

Clearly, there is no agreement about what sexual orientation is. Given this, it seems that the next question to ask is which (if any) of the various available understandings seems best to adopt. I argue that our current options are bleak. Almost without exception, they assume binary categories of sex or gender (i.e., male/female or men/women); they disagree on whether sexual orientation concerns gender-attraction (attraction to individuals with certain genders) or sex-attraction (attraction to individuals with certain sexes); and they typically leave us with no clear sense of when or under what conditions we can determine someone’s sexual orientation. These ambiguities and assumptions, I argue, are not only metaphysically problematic, but also normatively problematic. They maintain a historically policed majority/minority divide between heterosexuality and other sexual orientations, and they reinforce cisnormativity by blending together biological sex and gender, leaving no space for recognizing persons who are not cisgender²¹, or who are attracted to persons who are not cisgender.

I suggest that this conceptual confusion leaves us with room for a new prescriptive ontological project: one that aims to develop a new concept of sexual orientation that is better suited for our normative aims. In taking on this project, I first set out some of those normative aims. I argue that we would be better off with a concept that clarifies the criteria for ascribing sexual orientation; is consistent with a distinction between biological sex and gender; reduces or eliminates the presumption that being cisgender and heterosexual is normatively standard and

¹⁷ Stein (1999), 45.

¹⁸ American Psychological Association (2008).

¹⁹ Human Rights Campaign (2014).

²⁰ Stein (1999), 45.

²¹ That is, persons who identify with a gender other than the gender they were assigned at birth.

statistically normal; and is conducive for establishing legal and social protections for those who are not cisgender or heterosexual.

The concept of sexual orientation I then develop distinguishes between biological sex and gender, allows gender- and sex-attraction to vary independently, and does not reference one's own sex or gender. In doing so, it absorbs the salience of gender and sex in our current understanding of sexual orientation, but does not taxonomically privilege cisgender or heterosexual persons. In addition, it suggests that the conditions under which we should assess someone's sexual orientation by their sexual behaviors are the 'ordinary conditions' – that is, the conditions under which sexual orientation terms typically are deployed. This is meant to acknowledge that sexual orientation categorization is socially contingent, rather than based on immutable biological traits, and also to allow flexibility in our future understanding of which behaviors are reliable indicators of sexual orientation.

The upshot of this account is a socially situated, novel understanding of sexual orientation that furthers the goal of reducing gender- and sexual orientation-based oppression. It reveals some of the ways in which our current understanding of sexual orientation is problematic, and aims to motivate philosophers and non-philosophers alike to continue re-imagining sexual orientation categories.

Chapter 2

Oppressive Truths

2.1 Introduction

Who should be counted -- that is, acknowledged and treated -- as belonging to social groups such as *women, disabled persons, black persons, or parents*? Such questions incite arguments that typically aim to establish and police particular boundaries around these categories.¹ For example, debates over who should have access to women's bathrooms have focused on questions like *what makes someone a woman?* or *are trans women really women?* Participants in these (increasingly common)² debates assume that someone should be counted as a woman – and so have access to women-only spaces such as women's bathrooms – only if that person in fact is a woman. So too, debates over the policies and applications of the Americans with Disabilities Act have focused on questions such as *who is disabled?*³ Here we find a similar assumption: someone should be counted as disabled and so be provided with the protections and accommodations required by the

¹ I here use 'social group' to mean a group of persons such that persons viewed as members of this category are treated as having a particular social role. For example, *politicians, Americans, hipsters, mothers* are social groups, and *people with a quarter in their left pocket* is not. See section 2.2 for more on this point.

² See Brubaker (2016), 416: "Recent decades have seen a massive destabilization of prevailing categorical frameworks for the organization of cultural and somatic difference. ... The landscape of identity categories has become much more complex, fluid, and fragmented. Uncertainties and ambiguities in identifying oneself or categorizing others have been thematized and highlighted, and prevailing practices of counting, categorizing, and classifying have been challenged."

³ See, e.g., debates over the 2008 ADA Amendments Act, which broadened the number of physical or mental impairments that count as a disability under the ADA. [ADA National Network (2013).]

ADA only if that person in fact is disabled.⁴ Both debates, then, share a general background assumption: someone should be counted as a member of a particular social group, where this means acknowledging and treating that person as a group member, only if they in fact are a member of that group.

We can formulate these and similar background assumptions in the following way:

Assumption: The facts about who should be counted as a member (or non-member) of a social group X are constrained by the facts about who in fact belongs to social group X.⁵

Importantly, this assumption does not only arise in public discourses such as the ones mentioned above. It also arises in philosophical literature regarding the ontological nature of social groups. For example, in philosophical discussions regarding the metaphysics of gender, numerous authors have reasoned from the idea that transgender women ought to be acknowledged and treated as women to the idea that the correct metaphysical theory of gender (if there is one) will ensure that transgender women fall within the category *women*.⁶ This reasoning again employs the above assumption, because it begins with facts about which persons should be acknowledged and treated as belonging to a social group, and then uses these facts as a basis for determining which persons in fact belong to that group. Such reasoning is successful only if we can read off who belongs to a social group by looking to who should be counted as group members.

⁴ For another example, consider the debate over whether persons who smoke or sell marijuana are criminals. These debates appear to have a similar assumption: someone should be incarcerated or treated as a criminal just in case that person in fact is a criminal.

⁵ Or, in other words, someone should be acknowledged and treated as belonging to X *only if* that person is an X.

⁶ See, e.g., Barnes (2016), Jenkin (2015), McKittrick (2015), and Mikkola (2011). See Chapter 3 for further critique of this assumption.

Although this assumption is widespread, I think it is false. The facts about who belongs (or does not belong) to a social group can come apart from the facts about who should be counted as (that is, acknowledged and treated as) members of that group. This is for two reasons. First, the facts about who belongs (or does not belong) to a social group may be *oppressive truths*. Second, under certain conditions, making *false* group ascriptions (i.e., incorrectly counting group membership) can combat the corresponding oppression.

My argument will be broken down into two parts. First, I motivate and develop a theory of oppressive truths and their relationship to group ascriptions. Oppressive truths, here, will be understood as truths about social group membership such that *correctly* ascribing social group membership – that is, counting someone as a group member according to the ontological facts about that group – leads to an unjust distribution of social status and reinforces the relevant oppressive truths.⁷ That is, the ontological facts about who belongs to a social group may be such that ascribing membership according to the ontological facts can both (i) unjustly distribute social status, and (ii) reinforce the ontological facts underlying this unjust distribution.⁸

Next, I turn to the normative upshots of my view. I first argue that, under certain conditions, false group ascriptions can combat the unjust distribution of social status by working to revise oppressive truths. Under such conditions, then, the target assumption does not hold: who should be counted as a group member comes apart from who is a group member. This is because we

⁷ Here, I follow Ann Cudd in broadly understanding oppression as “a harm through which groups of persons are systematically and unfairly or unjustly constrained, burdened, or reduced by any of several forces... [that is] social *injustice*.” [Cudd (2006), p. 23.] Injustice, then, includes situations in which groups of people are systemically and unfairly ascribed (or denied) a particular group status.

⁸ I take Kapusta (2016), 514 to be hinting at the possibility of oppressive truths with respect to gender categories, as Kapusta states that she “has not touched upon the metaphysical question” of whether transgender women are women, or what makes them women, in order to avoid the danger of “excluding or marginalizing at least some people.” Ontological facts about gender categories such that, when widely employed, result in excluding and marginalizing people (usually transgender people) would count as oppressive truths, on my view. See Chapter 3 for more on the application of oppressive truths to gender.

should make ascriptions that combat unjust distribution of social status, and under certain conditions, false group membership ascriptions can do this.

I then argue that my framework of oppressive truths is a useful tool for motivating a revisionary approach within feminist ontology (sometimes called the *ameliorative project*), as well as dissolving puzzles over how to ascribe gender or race to persons who identify with a gender or race other than the one assigned to them at birth.

2.2 Social Groups

Before diving into my framework of oppressive truths, I want to clarify what I mean by ‘social groups’. I will use this term to refer to a classification (or, if you prefer, a kind) of persons such that (a) a social status is associated with this classification, and (b) this social status is conferred on someone when they are systemically ascribed this classification.⁹ (Paradigm examples of social groups include *women*, *disabled persons*, *black persons*, *gay persons*, *Americans*, and so on.)

While much can and has been said about social statuses, and their relationship to social groups, I will limit myself here to describing the intuitive idea behind (a).¹⁰ This is the idea that humans take certain features (e.g., *having female reproductive features* or *having a Southern accent*) to be socially significant, and use these features to sort themselves and others into various classifications for purposes such as understanding, evaluating, and predicting the behavior of themselves and others. Social groups, then, differ from classifications such as *left-*

⁹ Ásta (2013) refers to such groups as ‘human kinds’. However, I prefer the term ‘social groups’ in order to be clear that these kinds have social significance.

¹⁰ For more on this topic, see Haslanger (2000), (2012) “But Mom...”, Young (1990), Darwall (ms), among many others.

handed persons or *persons between 5'3" and 5'5"*, which do not have an associated social status, (though such features may impact how someone navigates the world).

We can now turn to unpacking (b). Drawing from Ásta's work, we can describe the property *having X-status* (for some social group X) as a *conferred* property – that is, “a property that something has in virtue of some attitude, action, or state of subjects, or group of subjects.”¹¹

According to Ásta, conferred properties can be broken down into five aspects:

Conferred property: What property is conferred

Who: Who the subject is

What: What attitude, state, or action of the subjects matter (for conferral)

When: Under what conditions the conferral takes place

Grounding property: What the subject is attempting to track (consciously or not), if anything.¹²

To illustrate this framework, we can consider Ásta's example of a strike in baseball. On a conferralist account of the property *being a strike*, a pitch has the property of *being a strike* just in case an umpire judges the pitch to be a strike in the context of a baseball game. The umpire will make this judgment according to whether or not they perceive the pitch to have a certain grounding property – namely the property of *having crossed the strike zone* – but a pitch is a strike regardless of whether or not the pitch in fact has the grounding property. The pitch is a strike in virtue of the umpire's call. However, in making their call, the umpire attempts to track

¹¹ Ásta (2013), 719

¹² Ásta (2013), 719

whether or not the pitch has the grounding property of *having crossed the strike zone*.¹³

Return to the case of social statuses. Here, I suggest, one has the *conferred property* of *having X-status* in virtue of facts about whether one is systemically acknowledged and treated as having X-membership. That is, someone has X-status in virtue of facts about whether that person is *systematically ascribed X-membership*. As in the case of a strike, these ascriptions will be based on whether someone is perceived as having particular grounding properties – in particular, whether they are perceived as having the properties that determine X-membership.¹⁴ For example, consider a context where *having female reproductive features* is taken to determine whether someone belongs to a group (*women*) that has an associated social status. Here, someone will have be conferred the social status associated with *women* so long as they are systematically counted as a woman, and they will be so counted just in case they are perceived as having female reproductive features. Of course, this opens the possibility that, in some cases, someone who *does* have these features – and so, has group membership – nevertheless lacks the associated status because others do not perceive these features. (And, similarly, in some cases, someone who does *not* have these features is nevertheless conferred the associated status.)¹⁵

If this picture remains fuzzy, consider a toy example: In Dr. Seuss's children's book *The*

¹³ Darwall (ms) provides a similar conferralist account of various types of recognition respect, such as honor.

¹⁴ Of course, this leaves open the possibility that, in some cases, X-membership and X-status cannot come apart. This does seem true for *some* groups – in particular, groups where X-status is a necessary condition for X-membership. For example, consider a group like *celebrities*. Someone belongs to this group only if they have the associated social status – to have the status of a celebrity is necessary for being a celebrity. Arguably, a particular social status is also necessary (if not necessary and sufficient) for belonging to groups such as *hipsters* and *political leaders* – that is, groups such that social recognition as belonging to that group is required for group membership. In such cases, the properties *belonging to X* and *having X-status* are identical, and are conferred properties. I diverge from Ásta in dividing (for at least certain social groups, such as race and gender groups) *belonging to that group* from *having the social status of a group member*, and in understanding the second but not the first as a conferred property. On Ásta's conferralist account, membership in social group always is a conferred property – that is, always requires social recognition. I take this division to make better sense of cases of passing, as well as our intuition that someone can belong to a social group (e.g., *queers*, *women*) without others recognizing this membership.

¹⁵ For clear examples of either sort, see cases of 'passing' as belonging to another social group – for example, the case of Jack Williams (aka Frances Clalin), a woman who lived as a man for three years in order to fight in the Civil War. For more on passing see, Mallon (2004) and Piper (1992).

Sneeches, he describes a society of Sneeches, wherein some Sneeches have star-shaped symbols on their stomachs, and others do not. In this society, the Sneeches perceived as having stars on their bellies are socially differentiated from those perceived as not having stars on their bellies, and their social status is based on this differentiation. Those that are perceived as being star-bellied have social power and control resources as well as the movements and freedoms of plain-bellied Sneeches. That is, it is not simply the case that some Sneeches are star-bellied and other are plain-bellied – Sneeches are divided into two social groups: *star-bellied Sneeches* and the *plain-bellied Sneeches*. Moreover, what social status a Sneechee has is determined by how they are systemically ascribed group membership by others. (One can imagine two Sneeches with identical but ambiguous belly markings, wherein only one has *star-bellied Sneechee* social status, because only one is systemically ascribed membership in this group.) That is: being ascribed membership in one of these groups determines a Sneechee's social status.

With this framework, we can return to the notion of oppressive truths. Recall that oppressive truths are facts about social group membership such that accurate social group ascriptions lead to the unjust conferral or denial of a social status. We now can put this a bit more precisely:

Oppressive truths: Facts about which features determine membership in a social group X such that to systemically confer the property *having X-social status* on the basis of these features unjustly distributes this property and reinforces these facts.

There are many ways, verbal and non-verbal, to count someone as a member of a social group.¹⁶ I will here focus primarily on cases of group ascriptions occurring through speech acts, in which one is explicitly said to fall under the label associated with that group. (E.g. “Alice is a woman.”) Importantly, I do not assume any particular view about the reference of social group terms – they may refer only to those who possess the properties grounding group membership, they may refer to those who possess grounding properties *and* the corresponding status, or they may refer to those who have the corresponding status. For example, suppose one holds that the property of being female has social significance, such that those perceived as having this property are grouped under the label ‘women’ and are conferred a particular status. One might think that ‘woman’ refers to females, that it refers to females who have been conferred the associated status, or that it refers to those who have been conferred this status (even if they are only passing as female). On my terminology, ‘woman’ is here a social group term, regardless of which semantic view is correct, because to count someone as falling under this label confers the associated status. Nothing that follows hinges on which of these views one favors regarding the referent of a given group term.¹⁷

From here, we can turn to illustrating the phenomenon of oppressive truths. After this, I will analyze this phenomenon, arguing that the possibility of oppressive truths reveals that *true* group ascriptions can unjustly confer social status, as well as reinforce oppressive truths – i.e., reinforce facts about which features determine membership in that group.

¹⁶ Thanks to Zoe King for raising this point.

¹⁷ That said, I should make one slight clarification for those who think that social group terms might refer to natural properties underlying group classifications (e.g. that ‘woman’ picks out persons with XX chromosomes). In that case, when I talk about revising group term deployment, I am advocating a reference shift for those terms, and not simply ontological revision in the groups associated with them. But, e.g., if ‘woman’ picks out those with the conferred social status, then revising use of this term will change who is a woman without changing the meaning of ‘woman’. Thanks to Ásta for pressing me on this point.

2.3 Illustrating the phenomenon

Let's begin with an illustration characterizing the phenomenon of oppressive truths. Consider the social group *parents*. This group is associated with a far-reaching as well as practically and symbolically important social status. Although the details of this status change across contexts, being conferred the social status of a *parent* likely includes (among other things) being expected to feel love for and commitment to one's child, being obligated to care for the child's physical and emotional needs, and being taken to have certain rights, such as the right to have the child in one's care, and to make medical and educational decisions on the child's behalf. Having this status likely also means that others will expect one to self-identify as a parent, and will be disposed to interpret one's behaviors as those of a parent -- (e.g., interpreting attendance at parent/teacher conferences as unremarkable, or lack of affection toward one's child as shocking and blameworthy).

Now consider the gymnast Simone Biles. At the age of three, Simone came under the care of her grandparents, Ron and Nellie. They have remained her primary caregivers since, and Ron, Nellie, and Simone all identify Ron and Nellie as Simone's parents. Nevertheless, when Simone competed in the 2016 summer Olympics, NBC commentator Al Trautwig said on-air that Simone "was raised by her grandfather and his wife, and she calls them 'mom' and 'dad'."¹⁸ This comment, perhaps unsurprisingly, was met with an onslaught of criticism, as it suggested that Ron and Nellie are merely *called* Simone's mom and dad, and are not in fact her mom and dad. In response to this criticism, Trautwig dug in his heels and tweeted, "[Her grandparents] may be

¹⁸ See US Magazine, "Simone Biles Responds to NBC Sportscaster's Comment About Her Parents". August 10, 2016.

mom and dad but they are NOT her parents.”¹⁹ In tweeting this, Trautwig clearly insisted that, with respect to Simone, he would *not* ascribe membership in the group *parents* to Ron and Nellie – that is, he would *not* confer the status of a parent on Ron and Nellie. Again, his claim was met with wide criticism.

Clearly, many people find Trautwig’s claim that Ron and Nellie are not Simone’s parents unsettling, even reprehensible. But why? That is, what’s *wrong* with Trautwig’s claim that Ron and Nellie are not Simone’s parents – what explains people’s negative evaluations of this claim (and similar claims)? One explanation is that Trautwig’s claim is problematic because it is false; he misidentifies the features grounding membership in the group *parents*. For example, perhaps features grounding membership in the group *parents* are legally-based, such that present and former legal guardians of a minor (like Ron and Nellie) *are* parents. Trautwig’s claim, we then might think, is problematic because it is false. Ron and Nellie *are* Simone’s parents, given the facts about what features make one a parent.

This explanation is insufficient. Imagine a situation in which the social status of *parents* remains the same, but Ron and Nellie were unable to become Simone’s legal guardians (though their relationship was otherwise exactly the same).²⁰ In this situation, on the hypothesis that legal guardianship grounds membership in the group *parents*, Trautwig’s claim would be *true*. That is, we can easily imagine a scenario in which legal guardianship of a child was the feature determining whether someone is a parent, and in which Trautwig would be correct to say that Ron and Nellie are not Simone’s parents. Nevertheless, even here, Trautwig’s claim could be (rightly) criticized – even if Ron and Nellie in fact were not Simone’s parents, it seems like there would be something *wrong* with saying so. This suggests that our evaluation of Trautwig’s

¹⁹ @AlTrautwig, “They may be Mom and Dad, but they are NOT her parents.” August 7, 2016.

²⁰ Such situations occur frequently – for example, when the caregivers are politically disenfranchised (e.g., undocumented immigrants), or the child already has two legal guardians, etc.

failure to count Ron and Nellie as parents – and importantly, our criticism of this failure – does not depend on the truth or falsity of his claim, and instead is based on a *normative* evaluation.

Trautwig's claim illustrates a broader, more systematic phenomenon. To better understand this phenomenon, I now turn to describing and motivating a framework of oppressive truths. On this framework, whether or not Trautwig intended to, group ascriptions similar to and including his unjustly deny the social status of *parents* to Ron and Nellie. And so, *even if* these ascriptions are true – that is, even if the boundaries around the group *parents* currently are such that persons like Ron and Nellie are not parents – these ascriptions systematically create an unjust distribution of the accompanying social status. Moreover, if true, they additionally *reinforce* the ontological boundaries around the group *parents*, and with them, the oppressive truth that persons who ought to have the status of parents are not parents.

2.4 Oppressive Truths

My framework for understanding oppressive truths can be broken down into two parts:

First: Sometimes, how we ascribe membership or non-membership in a social group unjustly confers or denies to persons the social status associated with that group. (Even when the ascriptions are true.)

Second: Social group term deployments, including ascriptions of membership or non-membership in a social group, can reinforce the facts about which features determine membership in that group.

In what follows, I unpack each part of this framework. But putting these pieces together, we arrive again at the idea that oppressive truths are facts about which features determine membership in a group X such that conferring X-status on the basis of these features (via group ascriptions) both unjustly confers (or denies) a social status to certain persons, and also reinforces the facts in question.

Before proceeding, I should note that my discussion is framed as if a single, discrete social status is associated with a given social group. This is a practically motivated simplification of a messy reality. What status is associated with a particular group may vary widely across contexts and typically does not have a precise delineation – for example, the status typically associated with *parent* among lower-class white families in the South may differ in significant ways from the status typically associated with parenthood among upper-class Chinese families in the Pacific Northwest, and it might be difficult to precisely articulate the features of either status. My present focus is on group ascriptions that confer the status associated with that group *in the ascriptions' context*.²¹

At this point, you may be wondering why this paper brings in the notion of truth – why not focus on the independently interesting argument that social group ascriptions can unjustly confer or deny particular social statuses, and leave it at that?²² Why also argue that *true* ascriptions can do this, and – when they do – that they reinforce their own truth by reinforcing the facts about who does and does not belong to a social group?

²¹ I make this point because, not only can the status associated with a group vary across contexts, but also, so can the features determining group membership. For example, there is a clear sense in which the social group *US citizens* has persisted across hundreds of years, even though the features determining membership in this group as well as the associated status has varied.

²² I recognize that the notion of ‘truth’ is typically applied to verbal assertions, and not to non-verbal ways of communicating. Group ascriptions, of course, can occur verbally or non-verbally. While I here focus on the case of verbal ascriptions, the notion of ‘accuracy’ or ‘correctness’ can be substituted for ‘truth’ in the case of non-verbal group ascriptions.

Consider the fact that there are many different ways of evaluating claims. One way of evaluating a claim is based on its truth or falsity. If I tell a friend who is deathly allergic to peanuts that the smoothie I handed to them does not contain peanuts, the most important mode of evaluation for this claim is its truth or falsity. But there are other ways to evaluate claims. We also could evaluate them for their informativeness. Suppose I've borrowed your car, and I can't figure out how open the gas tank door. I call and ask you where I can find the gas door lever, and you respond, "It's in the car." Your claim is true, but nevertheless I would be annoyed because your statement is extremely uninformative. In this paper, I want to highlight yet another type of evaluation for a claim: evaluations based on whether ascriptions reinforce oppressive truths about group membership (and along with these truths, unjust practices of conferring or denying social statuses). I hope to make clear that not only can we evaluate claims based on such reinforcement, but also that under certain conditions, this evaluation is more important than evaluations based on truth and falsity. Talking about truth is necessary for achieving this goal. In particular, it's important for seeing that, even if these two modes of evaluation come into conflict, such that certain group ascriptions are true but reinforce oppressive truths, we sometimes can rightfully criticize these ascriptions. Or, to reiterate the central focus of this paper: sometimes, which group ascriptions we *ought* to make are not constrained by which ascriptions are true.

We can now turn to examining each part of this framework.

2.4.1 Social Status and Injustice

First: Sometimes, how we ascribe membership or non-membership in a social group unjustly confers or denies to persons the social status associated with that group. (Even when the ascriptions are true.)

In some cases, how we confer or deny a particular social status through our group ascriptions misaligns with how we ought to confer or deny that status. This means something fairly broad: it means that, sometimes, the extension of persons who *should* be conferred a particular group's status – which is the extension of persons who ought to be systematically ascribed membership in that group – comes apart from the extension of persons who are in fact conferred that status. This is because the properties that should determine whether someone has that status are different than, an expansion on, or a reduction of the grounding properties that are used to confer that social status.

Return to Trautwig's claim that Ron and Nellie are not Simone's parents. Claims like his, which place a firm biological boundary around the group *parents*, are not unusual: Adopted children are frequently asked questions like 'who are your *real* parents?', gay parents are not acknowledged as their child's parents, and in general, persons raising children within non-traditional family structures often are not recognized as parents. There seems to be something normatively *wrong* with such group ascriptions, and I suggest it is this: these ascriptions unjustly deny *parent*-status to people should be treated as parents. That is, the features taken to determine membership in the group *parents* (i.e., the grounding property for the conferred property *having*

parent-status) in some way differs from the features that, given the associated status, *ought* to determine membership in this group.

Consider a different example: the social group *Americans*. In the United States, the status of an American is the status of an insider – to be seen and treated as an American is to be seen and treated as ‘belonging’ in the US. In contrast, to lack this status is to be seen and treated as an outsider. Sometimes the differences between persons with and those without this status are subtle – for example, the differences might manifest in the form of others making assumptions about where someone is from, what interests or preferences they have, and whether they share certain cultural experiences or exposures. At other times, the differences are obvious. For example, the fact that Native Americans’ interests often are treated as different than or even in conflict with American interests, the fact that the employment of immigrants and naturalized citizens are described as threatening to American workers, and the fact that persons seen as non-American often are only tolerated as residents, rather than seen as having equal right to political and economic engagement all indicate that these persons do not have the status of an American. In short, although the social group *Americans* is largely symbolic, the boundaries around this group have very real, and often material consequences.²³

Perhaps unsurprisingly, the grounding properties for conferring this status often are not properties merely concerning whether someone is a permanent resident in the US, or even a US citizen. According to a study by Jeremy Straughn, more than 89% of US citizens reported that in order to be “truly American,” it is important that one speak and write English.²⁴ An additional 71% reported that it is important that one is born on US soil, and 54% reported that it is important that one identify as a Christian. All this points to a state of affairs in which, in many

²³ Symbolic boundaries like these are “an essential medium through which people acquire status and monopolize resources.” Lamont & Mointer (2002), 168.

²⁴ Cox, et al. (2015).

contexts, persons who are naturalized citizens, who are not fluent in English, or who are clearly adherent to Islam, Buddhism, or another non Judeo-Christian religion are not conferred the status associated with being an American. In these contexts, the features that determine whether they are conferred this status come apart from the features that *should* determine whether they are conferred this status. Such a state of affairs is unjust because, in an important sense, it is one in which persons are not given their due: persons who deserve to be acknowledged and treated as Americans are denied this recognition and status.

Moreover, group ascriptions that result in unjust state of affairs such as this may be *true*. Consider that fact that, prior to 1920, women were not able to vote in the United States. Prior to 1920, then, group ascriptions such as ‘women are not voters’ or ‘Susan is not a voter’ were true. But, of course, women *should* have had the social status of voters – it was a grave injustice to deny them this status. Moreover, there are clear instances where ascribing to women non-membership in the group *voters* established (or at the very least, perpetuated) this injustice, even though the ascriptions were true: In the 1915 North Carolina Supreme Court case *Beckett v. Knight*, the court was deciding whether women should be allowed to receive fees for notary services. The court ruled that they should *not*, citing the reason that, “Women are not voters, and...notary fees should be reserved exclusively for voters.”²⁵ Here, the court’s claim that “women are not voters” was obviously true. But through this ascription (along with many other similar ascriptions), women were denied *voter*-status, including (in this case) particular economic opportunities.

While *voters* is a simpler case, given that the features determining group membership are stipulated by an authoritative body, group ascriptions concerning symbolic or non-stipulated

²⁵ Dissenting Opinion of Chief Justice Walter Clark in *Beckett v. Knight* 1915.

social groups, such as *American* or *parents*, can also both be true and unjustly confer or deny social status. Consider the likelihood that, in many communities, the terms referring to these groups are both *taken* to and *used* to refer to a group of people that is distinct from the group of people who plausible ought to be conferred the associated social status.²⁶ For example, hold fixed the social status attached to the group *parents*, and assume that the word ‘parent’ picks out the social group associated with this status. If we are in a linguistic community that takes and uses the term ‘parent’ to refer only to immediate progenitors of a child, or only to legal guardians of a child, then in that context, it is highly plausible that many ascriptions about who is or is not a parent would both be *true* and unjustly confer or deny *parent*-status to certain persons. (Again, I am assuming that adoptive parents as well as primary caregivers for a child who are not that child’s legal guardians deserve this status.)²⁷

At the very least, I think the burden of proof is on someone who wants to show that dominant conceptions and use of a social group terms can significantly diverge from correct use of that term. That is, if social groups are created through our practices of treating certain features as socially significant and grouping persons with those features under a label, how could there be widespread error in our ascriptions? When one want to know what features determine membership in social groups, it is difficult to see what more one should look to other than how people deploy social group terms, and what they take these terms to mean.

That said, you might think that social groups are normative kinds, such that who belongs to a group just is who (morally and politically) *ought* to belong to these groups. On this view, both conceptions and uses of group terms could misalign with their reference, because these terms

²⁶ That is, on both intensional and extensional accounts of these terms’ meanings, the terms will refer to a group that is not co-extensive with the group of persons who ought to have the associated status.

²⁷ Adding to this is Eli Hirsch’s (2009, 240) view that the principle of charity asks us to interpret a linguistic community such that shared use is as reasonable and accurate as possible.

correctly apply only to those who they *ought* to apply to. (E.g., “Ron is Simone’s parent” is true only if Ron ought to be counted as Simone’s parent.) This view is motivated by the thought that social group terms are, in some sense, normatively ‘thick’.

An example of such a view is found in Esa Diaz-Leon’s work on the meaning of ‘woman’. She argues that “moral and political considerations” within a given context can be relevant for determining whether utterances of gender terms in that context are true, because these considerations are often needed to determine whether someone is a woman “for practical purposes.”²⁸ That is, supposing that we are in a context where we want to know whether someone is a woman for the purpose of (e.g.) knowing whether they *should* be allowed to use the women’s bathroom, Diaz-Leon argues that moral and political considerations are relevant criteria for determining whether someone is a woman.²⁹

On such a view, extended to other social group terms, there will never be mismatch between the persons who can truly be counted as social group members and the persons who should be treated as members. Claims like “Ron is not Simone’s parent” may reinforce the unjust denial of social status, but on this view, these claims must be false and so could not express an oppressive truth.

I think there is a serious problem facing this line of reasoning: it appears to trade on an ambiguity in its use of ‘should’. Perhaps a claim such as, ‘Claire is a woman’ is true only if, in *some* sense, Claire should be recognized as a woman given the purposes at hand. But why would ‘should’ here mean anything more than ‘should according to the prevalent ideology in that social context’? That is, why think that it is a moral ‘should’ or a justice-oriented political ‘should’? This view requires independent motivation. Moreover, we have reason to be highly skeptical of

²⁸ Diaz-Leon (2016), 249.

²⁹ Ibid., 249-50.

it: if the view were true, a huge number of ascriptions we take to be true would be false. For example, consider historical group ascriptions such as, ‘He is a slave’, asserted for the purpose of determining who should be considered another human’s property. On a view such as Diaz-Leon’s (but applied to the term ‘slave’), moral and political considerations would determine the truth-value of such claims. The result? No one is a slave. After all, no one should – in a moral and political sense of should – be considered another human’s property. In other words, on a normatively loaded view of social group terms, someone could truly be counted as belong to a social group only if they should be conferred the social status associated with that group. But this would result in a massive error theory with respect to many of our group ascriptions, as it shields us from moral and political error in the construction of social groups.

But suppose you aren’t convinced, and you still are sympathetic with the idea that, as a matter of fact, who belongs to a social group always aligns with who should have the status associated with that group. (That is, you deny the possibility of oppressive truths, in my sense.) In that case, I want to appeal to a different, pragmatic reason for you to adopt a concept of oppressive truths – namely, for the purpose of productive public discourse. Consider the ongoing public controversy over which persons we should count as men and as women. Frequently, this controversy manifests as disputes over questions like *who is a man (or woman)?* Or *what makes someone a man (or woman)?* If I am right about the possibility of group ascriptions unjustly denying or conferring particular statuses, then the debate over what makes someone a man or a woman is a smoke screen distracting from the core issue of who should have the social status of a man or that of a woman. Suppose, for example, we believe that transgender women should have the social status of women, including access to women-only spaces, but are confronted with people who think that someone is a woman if and only if they were born with a female

reproductive system. In this situation, the concept I've developed here shows that the important debate to have with this person is not over whether their gender group ascriptions are true, (even if you think they are not), but rather, over whether gender group ascriptions like theirs perpetuate injustice. And again, this is because, in this context, normatively evaluating such claims in this way is more important than evaluating them for their truth.

More broadly, ongoing philosophical work on conceptual ethics suggests that disputes over which persons should count as women are not alone in this regard. Though I don't have space to go into the arguments, I think this literature makes a strong case that many everyday and philosophical disputes over how we should represent (and categorize) various persons and things often are masked by disputes over whether or not certain assertions are true.³⁰ The notion of oppressive truths that I am developing shows clearly that, at least theoretically, these questions can come apart. In recognizing this, we gain a tool for refocusing these disputes onto core normative questions.

While politically controversial in some cases (e.g. in the case of *women*), I take it that nothing I've said so far is extremely philosophically controversial (even if surprising). Essentially, I've argued that social groups can be poorly constructed, by normative standards. The facts about which features determine membership in a social group can misalign with the facts about which features (given the associated social status) ought to determine who is conferred the social status associated with this group. In these cases, the ontological facts about who belongs to a group are such that, by correctly ascribing group membership, we establish or perpetuate a social injustice: people who rightfully should have (or not have) a particular social

³⁰ See, e.g., Plunkett & Burgess (2013), Plunkett (2015), Thomasson (2017).

status are nevertheless denied (or conferred) this status. Such ontological facts, then, are oppressive truths.

2.4.2 Reinforcing Ontology via Ideology

Second: Social group-term deployments, including ascriptions of membership (or non-membership) in a social group, can reinforce the facts about which features determine membership in that group.

Recall that social groups are not simply collections of people with shared features – rather, they are collections of persons whose features have the same (or similar) social significance. In brief, then, when deployments of a social group-term reinforce the practice of taking certain features as having a particular social significance, they also reinforce the fact that these features (rather than others) determine someone’s membership in that group.³¹

More specifically, how social group-terms are deployed can perpetuate shared practices of taking certain features to have a particular social significance – which, in turn, reinforces which features determine membership in the corresponding social group. To first illustrate this phenomenon, return to the toy example of the Sneeches. Suppose the Sneeches develop a shared practice of deploying the term ‘star-bellied’ only to pick out Sneeches that have a star on the *right* side of their bellies (rather than, as previously, those who have a star anywhere on their bellies). What would happen in this case? Recall that a social group is a collection of person

³¹ This is not to say that the features *determining* group membership are constitutively socially constructed. They might be objective physical features, such as height or skin color. But it is to say that social groups are constitutively socially constructed, insofar as a necessary ingredient for their construction is the shared practice of treating these features as socially significant.

whose features have the same or similar social significance. If this shift occurred, then, the feature that previous united the social group *star-bellied Sneeches* (namely, *having a starred belly*) also would shift, and become the feature *having a starred belly on the right side*. This indicates that there is a complex relationship between the ontological facts about a group membership and corresponding group-term deployments: while group ascriptions are true or false based on the ontological facts about a social group, our practices of group ascriptions (and so, of conferring status) set the facts about which features ground group membership.³²

Return to the *parent* example. Suppose there is a shared practice of taking the feature *biological progenitors of a child* to have a unique social significance, and that persons with features carrying this significance are called ‘parents’. In this case, then, someone is a parent of a child just in case they are the immediate progenitor of that child, and to call someone a ‘parent’ is to signal that they have features with this significance. In this situation, when people claim that Ron and Nellie are not Simone’s parents, or when they ask an adopted or foster child, ‘Who are your real parents?’ they reinforce the practice of taking the feature *biological progenitors of a child* as determining whether someone has features with this social significance. When this practice is reinforced, so too are biological boundaries around the group *parents* – that is, because the practice establishes which features determine membership in that group, to reinforce the practice is to reinforce the boundaries around group membership.

But substantive questions remains: What does it take to reinforce a social practice? And how does this reinforcing occur? That is, what is the mechanism by which group-term deployment reinforces practices of taking certain features (but not others) as having certain social significance?

³² For more on the grounding of social kinds, see Epstein (2016).

This mechanism, I think, is the same mechanism that Eric Swanson has argued accounts for at least one of the primary harms caused by slurs.³³ On Swanson's view, this harm comes about in two parts: First, using the slur – like using any word – carries an 'acceptability implicature' which is incredibly difficult to cancel. Specifically, it implies that using the word in question in *that* sort of context and in *that* sort of way is acceptable. And second, in using a slur, a speaker can reinforce a harmful ideology by 'cuing' that ideology for some purpose. That is, a speaker can reinforce a harmful ideology by putting the ideology in place to fulfill some purpose and signaling that the ideology *should* be used to fulfill that purpose.

I will here follow Swanson in taking an ideology to be a "temporally persistent and socially extended cluster of mutually supporting beliefs, interests, norms, practices, values, affective dispositions, and ways of interpreting and interacting with the world."³⁴ Importantly, this means that ideologies are complex and are typically assumed rather than explicitly acknowledged, meaning that they are a robust shared lens through which people communicate and view the world.

On Swanson's account, ideologies are not inherently misleading or harmful, but they can be both. Swanson's focus is on slurs that operate upon background *harmful* ideologies. Suppose a coach describes an opposing team as 'nerds' in order to motivate his players. In this case, Swanson argues, the coach cues an ideology on which it is normal, acceptable, or even praiseworthy to poorly treat persons associated with academic success and social awkwardness. Moreover, he puts this ideology in place to motivate his players and signals that this ideology

³³ While Swanson discusses the *mechanism* by which this reinforcement occurs, he (and others) assume that it does occur. For discussion of the empirical evidence of this reinforcement, Krauss & Chiu (1998), section "Language, Cognition, and Culture". Among other things, Krauss & Chiu discuss how language can bias both hearers' and speakers' representations of a person or object.

³⁴ Swanson (ms), 5.

should be used for this purpose. In so doing, the coach – along with others who use a slur – does three things:

- (1) He further emboldens himself to enact the ideology cued by the use of the slur;
- (2) He signals to others in the linguistic community that he consents to that ideology; and
- (3) He emboldens others to consent to and enact the ideology as well.³⁵

These effects strengthen (and perpetuate) the ideology in question, including the beliefs, subdoxastic attitudes, habits, norms, concepts, and practices that belong to that ideology.³⁶ More generally stated, to use a slur implies that using the word in the speaker's context and for the speaker's purpose is acceptable, and in doing so, reinforces (and encourages others to reinforce) the ideology cued by that word.

While most social groups terms may be importantly different than slurs, I think that Swanson's account applies far beyond the particular case of slurs: it details how social language (which is not limited to slurs) interacts with and reinforces ideological backgrounds. For this reason, the account is helpful for understanding how oppressive truths are reinforced. How we deploy social group terms can reinforce ideologies that include (among other things) beliefs and

³⁵ Swanson (manuscript) 14.

³⁶ For more on ideologies, including ideological perpetuation, see Haslanger (2017) and Shelby (2003), among others.

sub-doxastic attitudes concerning which features have social significance and what social significance they have, as well as the subsequent practice of grouping together people with these features under certain terms. Apart from whether or not they are used correctly, then, social group terms (like slurs) have uses and effects beyond simply describing persons' features.³⁷ For example, Trautwig's claim that Ron and Nellie are not Simone's parents did not simply describe Ron and Nellie as not being Simone's biological progenitors. It did two additional things: First, it implied that it is acceptable to deny Ron and Nellie the status of being Simone's parents in an everyday context. And second, it cued an ideology for the purposes of describing Simone's relationship with Ron and Nellie. On this ideology, there is a significant or notable difference between families in which primary caregivers of children also are their immediate progenitors and those in which they are not – a difference suggesting that the former type of family is in some way more valuable than the latter.³⁸ In cuing this ideology, Trautwig also reinforced it – and with it, the exclusion of persons like Ron and Nellie from the social group *parents*.³⁹

Importantly, a speaker need not *intend* to reinforce an ideology in order to do so. In fact, someone could intend to *not* reinforce that ideology or have no intentions one way or the other, and still do so. As Swanson also points out, what ideologies language cues depends on social factors that are independent of the speaker: specifically, it depends on the “social, causal, and historical facts about the speech situation.”⁴⁰ Even if Trautwig does not *intend* to reinforce a

³⁷ “When we seek to understand a word, what matters is not the direct meaning the word gives to objects and emotions—this is the false front of the word; what matters is rather the actual and always self-interested *use* to which this meaning is put and the way it is expressed by the speaker, a use determined by the speaker's position (profession, social class, etc.) and by the concrete situation.” (Bakhtin, 1935, 401)

³⁸ As was pointed out to me by an audience member at Tufts, it also cues an ideology on which black families are broken and so less valuable than the traditional nuclear white family.

³⁹ Similarly, claims such as “immigrants threaten American workers” both imply that it is acceptable to deny immigrants the status of being Americans, and cues an ideology on which immigrants are not treated as Americans. In cuing this ideology, a speaker emboldens themselves and other to treat immigrants as non-Americans, and signals to others that the speaker consents to such a practice.

⁴⁰ Swanson (ms), 11.

system of beliefs, practices, and concepts that establish facts about group membership in *parents* such that Ron and Nellie are not Simone's parents, this does not prevent his speech from doing so.

That said, notice I've stated that group-term deployment *can* rather than *will* reinforce the facts about which features determine group membership. This is for two reasons. The first reason is because I take it that ideology can be cued for the purpose of resisting or criticizing that ideology. For example, in the case of slurs, someone might intentionally use a slur – and so cue the ideology associated with that slur – for the purpose of emphasizing the harmful effects of that ideology, or raising the ideology up for criticism. Similarly, I think, someone can intentionally cue a harmful ideology through a true group membership ascription in order to highlight the injustice in the fact that certain persons are (or are not) members of that group. For example, consider the fact that, in many states, smoking marijuana is a crime, and so everyone who has smoked marijuana is a criminal. One can well imagine asserting, "Under US law, half of US citizens are criminals!" in order to emphasize the absurdity of the ideology on which half of US citizens belong to the group *criminals*.⁴¹

Second, I am open to the possibility that there are contexts – call them 'canceling contexts' – where a speaker *is* able to deploy group-terms without cuing an ideology. One example of a canceling context may be medical or legal contexts where the purpose of the ascription is narrow and well defined. For example, suppose that you are filling out paperwork to receive a test for a genetic condition, and the form uses the term 'parents' in such a way that it clearly is asking for information regarding your biological progenitors. The clearly narrow purpose of term

⁴¹ See also, e.g., the following line of Gabriel Rameriz's (2016) poem "On Realizing I am Black": "I am black, and full of stars. I am not the absence of light." Here too, Rameriz cues a problematic ideology – in this case, one associated with blackness – in order to criticize it. Without assuming this ideological background, the line loses its poetic and political force.

deployment *might* avoid cuing larger background ideologies of parenthood such as the ideology cued by Trautwig.

Such contexts aside, most group term deployment will cue *and* reinforce an ideology. These ideologies, as we've seen, can be harmful, insofar as these ideologies establish and perpetuate unjust boundaries around social groups.

So far, I've argued for the existence of oppressive truths, as well as the reinforcement of these truths via true group ascriptions. I now turn to examine three normative upshots of this view for (respectively) the implications for particular group ascriptions, the methodology of feminist metaphysics, and addressing disputes within identity politics.

2.5 Applications

2.5.1 Normative implications for ascriptions

You might rightly be worrying that an enormous amount of our social group ascriptions reinforce oppressive truths. For example, suppose that who currently falls within groups like *citizens*, *spouses*, or even *philosophers* excludes persons who deserve the associated status. Does this automatically mean that we should attempt to revise the boundaries around these groups by revising ascriptions? Or, consider cases where it is unjust that *anyone* has some group status: for example, the social status associated with *women*, *persons of color*, or *disabled persons*. Given that these statuses are oppressed, it appears to be an oppressive truth that *anyone* belongs to these groups. And if we reinforce the dominant ideologies surrounding these groups whenever we ascribe membership in them, aren't we reinforcing oppressive truths by counting *anyone* as a

woman, as a person of color, or as disabled? Does this mean we should stop ascribing these group-memberships altogether?

The first thing to say here is that, yes, we *do* constantly reinforce oppressive truths through our ascriptions, just as we constantly reinforce oppressive business and political practices through what we eat and what we wear. Our lives, including our communicative practices, are wrapped up in oppressive social structures. A lot of the things we do are complicit in and even reinforce these structures.

But that said, the fact that certain ascriptions will reinforce an oppressive truth does not entail that we should revise those ascriptions. The basic reason why is this: to say that we should change our ascriptions in order to avoid reinforcing oppressive truths assumes that, by revising our group membership ascriptions, we would be able to avoid reinforcing injustice in the distribution of social status. And I think that, in many cases – perhaps even the majority of cases – we are *not* able to do this.

This point is in fact commonly made in feminist and anti-racist philosophy. It is generally offered in response to those who support eliminativism about gender or race – that is, doing away altogether with race and gender groups, and so eliminating gendered and racial terms. Even those who agree that an ideal world would be one without (e.g.) racial groups have argued that, given that racial ideology is so deeply embedded, and given that this has created significant material consequences for people based on their racial group, no longer ascribing race would not correct racial injustice, and might even enable this injustice by taking away conceptual tools that help us articulate its manifestations and targets. That is not to say that our racial groups should not undergo *any* revisions – but just to say that they should not be done away with altogether in one fell swoop. Again, this is because doing away with these groups would not actually work

against large-scale harmful practices surrounding race and gender. In this case, what likely would be most effective is working to change the social significance of racial features, rather than changing which features count as particular racial features.

Another case in which changing group ascriptions would be not correct the relevant injustice is when the revisions are too local to have substantial impact.⁴² For example, suppose that I and a few of my friends think that all persons in a long-term relationship should be treated as spouses, and so we ascribe membership in the group *spouses* to all persons in long-term relationships. Most likely, even though we might have a small, local manifestation of this group that avoids certain injustices, our ascriptions would not in fact address the most pernicious injustices. Because our ascriptions in this case are so local, they would be unintelligible to those around us rather than effectively resist oppressive truths – most likely, others would think that we were either confused about those persons’ marital status or simply don’t understand what the word ‘spouse’ means.

Both of these sorts of cases indicate that the conditions under which we should revise ascriptions that would reinforce oppressive truths are the conditions under which changing our ascriptions could *effectively* combat the relevant injustice.⁴³ I suggest one such condition is when our revised ascriptions would cue an *available* alternative ideology that revises the boundaries around the relevant social group and, in so doing, effectively addresses an unjust distribution of

⁴² This relates to a larger point, also made by Sally Haslanger (2017), p. 7, that ideology is a *cultural* phenomenon, and “so a change of culture, and not just a change of individual attitudes, is required to achieve social justice.” That is, in order for ideological change that achieves social justice, individual and small local changes will not suffice – widespread cultural change is required.

⁴³ Some might worry that, in saying this, I am explicitly rejection a knowledge norm of assertion. While there are many formulations of this norms, it roughly can be taken to say that according to which S should only assert that p if S believes p to be true. However, while this norm is highly contested, there are a variety of ways to make my claims compatible with it. Here are two: First, one might, along with Basu (ms) and others, take it that there are normative constraints on knowledge, such that even if p is true, one might not know p if normative considerations weigh against justification in believing that p. Second, one might understand ascriptions and similar speech acts not as assertions, but rather as verdictives. (Thanks to Cat Saint-Croix and Ásta for discussion.)

social status. (By ‘available’ I mean that there is a mutually shared recognition of the cued alternative ideology.)

It is important to emphasize that having an available alternative ideology is no trivial matter. For example, if I were to say to you, “When I say ‘parent’, I mean ‘any adult in the household’. And Ron is Simone’s parent!” this may be comprehensible to you, but will *not* cue an available alternative ideology. Ideologies, as mentioned earlier, are complex, robust, and shared. In order for an alternative ideology to be available, then, there must already be a culture or sub-culture in which that ideology is operative, and both the speaker and hearer must be aware of this community.

This is perhaps best clarified with an example. In “Trans-Women and the Meaning of ‘Woman’”, Talia Bettcher argues that gender terms like ‘woman’ do not have a single meaning – rather, they change in meaning across contexts. Trans activist subcultures, she writes, have “developed different gender practices (including the use of gender terms such as ‘woman’ and ‘man’) that do not always accord with more mainstream ones.”⁴⁴ In such communities, for example, the question *are trans women really women?* does not make sense, as trans women are paradigm cases of women rather than borderline or otherwise opaque cases. I take Bettcher’s description to indicate that, in trans activist subcultures, there are operative ideologies surrounding gender that are *alternative* and even *resistant* to dominant ideologies of gender in the broader culture. That is, it is not simply that the members of these subcultures use the words ‘woman’, ‘gender’, etc. differently – their entire system of beliefs, norms, practices, and subdoxastic attitudes surrounding gender are different than the dominant system. In a case like this, someone in a context in which the dominant ideology is operative may be able to use gender

⁴⁴ Bettcher (2013), 235.

terms in such a way that they can successfully cue this alternative ideology. (At the least, I take this to require that both speaker and hearer are somewhat aware of the alternative ideology.) The availability of an ideology, then, requires collective momentum – in the case of gender ideologies in Northern countries, however, it seems that this momentum is quickly amassing.

2.5.2 Motivating the Ameliorative Project

One might worry that my argument thus far has a troubling implication: namely, that whenever the extension of a social group misaligns with the extension of persons who should have the associated status (and the conditions in 2.1 are filled), we should just *change the subject* when we talk about that group. That is, rather than talk about the way the world is, we should talk as if the world were a different (albeit better) way, which would not only suggest a blatant naivety, but also would result in talking past the majority of people.

To best understand my reply to this objection, it is important to first look at prominent place where the same objection arises: in response to ontological analyses of gender and racial groups that do not align with everyday concepts of those groups, and especially against *ameliorative* analyses of such groups – that is, normatively guided analyses of these groups. An ameliorative project, as standardly understood in social ontology, begins by asking what our “legitimate purposes” are in categorizing people into certain social groups (e.g., race and gender groups), and then develops an account of those groups that helps achieve these purposes.⁴⁵ One example of an ameliorative project is Sally Haslanger’s analysis of gender and race. On her account, membership in these groups is not determined by one’s bodily features, but rather, is determined

⁴⁵ For more on the ameliorative project, see Haslanger (2012), 366.

by one's social status – in particular, whether one is privileged or subordinated *because* of one's bodily features (among other things).

The charge of 'changing the subject' is frequently raised against such ameliorative projects. These ameliorative, social constructionist accounts, some worry, change the subject of analysis, rather than in fact analyzing the things people are talking about when they talk about social groups. Typically, this objection is raised for ameliorative accounts of social groups that are widely taken to be (in some sense) natural categories, rather than socially constructed ones.

Some philosophers have responded to this charge by arguing that these analyses are *not* changing the subject, and that the groups we are talking about when we talk about social groups are often quite distinct from what common understanding of these groups would suggest we are talking about. Sally Haslanger is perhaps best known for this line of response, arguing that:

[Some] argue that [the ameliorative approach] amounts to changing the subject.

Our notion of parent (or *our* notion of race) is of a biological category, and any modification that disrupts that assumption replaces *our* concept with a difference one. In other words, 'parent' just means *immediate progenitor*, and if we start using it [differently], then we have changed the meaning. So social constructionists are wrong to say that *parent* is socially founded...

Social constructionists can rely on externalists accounts of meaning to argue that their disclosure of an operative or a target concept is not *changing the subject*, but better reveals what we mean. By reflecting broadly on how we use the term 'parent', we find that the cases, either as they stand *or adjusted through*

ameliorative analysis, project onto an objective social, not natural type... This is not to propose a new meaning, but to reveal an existing one.⁴⁶

In some cases, this might be the best response. Maybe in some cases, normatively guided metaphysical analysis will reveal what we are talking about, rather than changing the subject. And, more specifically, maybe in some cases, social groups we all take to have (in some sense) natural boundaries in fact are socially constructed.

But, that said, it would be nice if another response were available to ameliorative theorists. For one, I think we can imagine cases where this response would be a hard sell, because the result of ameliorative analysis would be distant from dominant understanding and practices surrounding a social group. But also, it would be helpful to have theoretical motivation for developing ameliorative analyses that aim at developing concepts that, if taken up within an alternative ideology, would *change* social groups, rather than simply describing these groups.

The concept of oppressive truths provides this alternative response. While an ameliorative analysis might advocate changing what we are talking about when we talk about a social group, this change may be called for. It provides motivation for developing a theory of a social group with explicitly normative goals in mind. Such projects would not be beholden to ‘getting it right’ about what our social group terms currently mean, or who currently belongs to those groups. Rather, they would focus on developing practically and normatively guided theories of social groups in hopes of revising these groups – that is, theories aimed at narrowing the gap between who is and who should be treated as members of a particular group.

⁴⁶ Haslanger (2012), 393, 398.

In addition to motivating the ameliorative projects, the concept of oppressive truths also targets the seemingly widely held assumption that, while we can mistreat members of certain social groups, *who* we should treat as members of those groups is set by the facts about who is a member of these groups, and those facts aren't up to us. That is, it seems to be taken for granted that not only is there a fact of the matter about who *is* a woman, who is disabled, and so on, but also, these facts determine who should *treated* as a woman, as disabled, and so on. This assumption perpetuates ignorance about the contingencies of who is taken to be members of a certain social group in our time and place, and so precludes understanding the important relationship between who is counted as and who is a member of a social group. In particular, it precludes recognizing that, not only are we sometimes called to change who is *counted* as a member of a group, but also who *is* a member of that group.

How we use words is important (even if not sufficient) for enacting such change. This point, well recognized in the work of Foucault and Ian Hacking, as well as in contemporary discussions of epistemic injustice, draws from the idea that many social facts are in some significant sense *up to us*.⁴⁷ In particular, facts about social groups are up to us: while the features that determine group membership (*de re*) may not be socially constructed, the ideologies attributing social significance, labels, norms, and so on to those with these features – i.e., the ideologies resulting in the construction of social groups -- are inherently social. Of course, that is not to say that these ideologies are easy to change, or that using words differently is, of itself, enough to create ideological change. Rather, this is to say that the words we use and the way we use them draw from and interact with background shared ideologies; that our speech can resist, adjust, or

⁴⁷ See, e.g., Fricker (2007), 147: "In the hermeneutical context of social understanding, it is also clear that, at least sometimes, if understandings are structured a certain way, then so are the social facts." Understandings, in a hermeneutical context, are highly reliant on there being available words to articulate one's experience. See also Foucault (1980) and Hacking (2000).

reinforce these ideologies. Sometimes, I am suggesting, this power of language can be used to adjust or question ideologies behind social groups, and in doing so might, as Linda Alcoff puts it, “yield new worlds, and not merely new words.”⁴⁸

This does not mean that there is no reason to investigate our social groups. Nor does it mean that philosophical accounts of social groups must articulate justly constructed social groups. In this, I disagree with Stephanie Kapusta, who argues that philosophical deployment of gender group-terms are unacceptable (from a transgender standpoint) if, were that deployment broadly implemented, it would be “an oppressive or harmful gender-term deployment.”⁴⁹ To translate this into terms of oppressive truths, one might think that philosophical deployments of group-terms that align with oppressive truths are unacceptable. This would immediately rule as unacceptable (e.g.) philosophical analyses of parenthood that excludes Ron and Nellie or analyses of gender that exclude transgender persons.⁵⁰

While I endorse the merit of philosophical projects that have prescriptive ontological aims, I also endorse the value of projects that *illuminate* oppressive truths. That said, the framework of oppressive truths highlights the fact that these projects must not be done in a supposedly value-neutral way – such neutrality, we’ve seen, is not possible. Unless such projects highlight the ways in which these ontological facts contribute to oppression, they will reinforce oppressive truths. This means that social ontological projects⁵¹ must take seriously the ability of their language to reinforce oppressive truths, and be open in taking a value-laden stand against oppressive social systems. (After all, to say nothing on the matter is to covertly take a value-laden stand *for* them.) Such projects, though, remain immensely important: far from being

⁴⁸ Alcoff (2010), 136.

⁴⁹ Kapusta (2016), 511.

⁵⁰ Haslanger’s (2000) famous analysis of gender has been criticized for this reason.

⁵¹ I suspect this extends to ontological projects in general, but I only commit to this restricted claim here.

‘unacceptable’ because they supposedly conflict with prescriptive projects, they pave the way for such projects.

2.5.3 Dissolving Identity Puzzles

Understanding oppressive truths is not only useful for academic philosophers; it is also useful for diagnosing ongoing public disputes over self-identity and social recognition. While such disputes are expansive, I will focus on the debate over how to count the gender or race of persons who identify with a gender or a race other than the one they were assigned at birth – (e.g., someone born with male-typed anatomical features who identifies as a woman, or someone whose progenitors are white who identifies as black).

Let’s begin by getting two concrete cases on the table. Let’s consider Janet Mock and Rachel Dolezal. Janet is a popular writer, transgender rights activist and TV show host. She was assigned male at birth, but identifies as a woman. Rachel Dolezal was an instructor in Africana Studies and the former president of her local NAACP chapter. She is the child of two white persons, was assigned the race ‘white’ as a child, but identifies as black. Both Janet and Rachel, we can assume, genuinely feel as though they were mislabeled at birth, and genuinely feel as though they are members of the groups with which they identify, rather than the groups they were identified as belonging to as children. Neither person is intentionally being deceitful or dishonest in expressing their identities.

Let’s now consider the following ascriptions:

Ascription 1: Janet Mock is a man.

Ascription 2: Rachel Dolezal is white.

How should we evaluate these ascriptions? Many people take it that we should not make Ascription 1 but may make Ascription 2. At least on its face, this set of beliefs is puzzling. Why do we evaluate these ascriptions differently? Why say that Rachel Dolezal is white, but not say that Janet Mock is a man? This puzzle is even starker for those who think that the explanation is a difference in these ascriptions' truth values, particularly when one takes into the account the large number of gender and race theorists who take these groups to have a similar metaphysical status. For example, if you think that both gender and race are biologically-based categories, it is puzzling why someone could not change their biological features to change races, but could do so to change gender. (This puzzle is exacerbated when you consider that many people would take Ascription 1 to be false even if Janet Mock did not change her physical features by receiving hormone therapy or surgeries.) And for those who think that gender and race are both are social-position based categories, it is puzzling why someone could change their social position to change genders, but not to change races. For example, on a view where gender is the social interpretation of perceived reproductive features, and race is the social interpretation of perceived ancestral features, it is puzzling why someone could not change races by changing what ancestral features they are perceived to have.

We could explore what makes someone a woman and what makes someone black, and whether the answers suggests that someone can change genders or change races. In fact, this has been the dominant approach – from both the political right and left – to debates over whether the following conditional holds: *If we should accept that Mock is a woman, then we should accept*

that *Dolezal is black*. As Rogers Brubaker notes, careful analysis of Dolezal's case, and its comparison to that of a transgender woman such as Mock's, has been "largely subordinate to efforts to validate or invalidate the identities claimed by [Mock] and Dolezal."⁵² He suggests that these effects are part of a much broader current cultural tendency to police or, alternatively, justify particular identity claims using notions of "authentic, objective, and unchosen identities."⁵³ Indeed, the majority of debates regarding the identities of persons such as Mock and Dolezal appear obsessed with drawing an asymmetry between their identities along such lines.⁵⁴

I will not take a position here on whether this asymmetry exists. Instead, I want to highlight the way in which my framework of oppressive truths offers an alternative path to dissolving this puzzle. We don't have to figure out the metaphysics of gender or race. We can reinterpret the puzzle as a *normative* (and political) puzzle rather than a metaphysical one. This opens up a new path for exploring a different sort of asymmetry in Ascription 1 and Ascription 2 – specifically, an asymmetry in our answers to the following questions:

Q1: Does ascription 1 avoidably and unjustly confer social status?

Q2: Does ascription 2 avoidably and unjustly confer social status?

⁵² Brubaker (2016), 415. Brubaker references Caitlyn Jenner, rather than Janet Mock.

⁵³ Ibid.

⁵⁴ Most common among progressive writers is an attempt to locate, in nature, history, or medicine, a reason why transgender identities are legitimate, while Dolezal's is not. See Urquhar (2015): "If race and gender are both social constructs, and if both have been built around observable biological traits, then what is the crucial difference that makes a felt gender identity a true one, but a felt racial identity fraudulent? The short answer is that most trans people and their allies suspect that transgender people are born that way." See also Talusan (2015) "Dolezal identified as black, but I *am* a woman, and other trans people *are* the gender they feel themselves to be." and Allen (2017) "Transgender people can undergo medically-proven treatments developed over decades to alter their sex characteristics; ... Rachel Dolezal told Matt Lauer in 2015 that she "certainly [doesn't] stay out of the sun.", among many others.

If we can show that there is an asymmetry in Q1 and Q2, such that ascription 1 does and ascription 2 does not unjustly confer social status, this would account for why many people evaluate these ascriptions differently without appealing to a difference in their truth-values. More specifically, we can account for an asymmetry in evaluations of Ascriptions 1 and 2 if we can show one or both of the following:

- (A) Were Ascription 1 true, it would reinforce an oppressive truth, but were Ascription 2 true, it would not.
- (B) Were Ascription 1 to reinforce an oppressive truth, it would do so avoidably, but were Ascription 2 to do so, it would be unavoidable.

That is, if we could establish that it would be oppressive if the group *women* did not include Mock, but not oppressive if the group *black persons* did not include Dolezal, this would be enough to show that – *regardless of truth value* – these ascriptions are normatively asymmetrical.⁵⁵ Alternatively, we could establish that, regardless of any relevant oppression, there is no available alternative ideology of race that would include Dolezal in the group *black persons*.

I haven't settled on a preferred explanation for this asymmetry. However, drawing from Diaz-Leon, we can briefly consider how one might use my framework to establish an

⁵⁵ In this, I agree with Brubaker (2016), 416 in taking the main significance of the Dolezal case to be the “opening it provides for the development of a more nuanced comparative analysis of the micropolitics of sex/gender and racial/ethnic identity and difference.” That is, it invites us to consider *practical*, and not simply theoretical questions about the similarities and differences between these identities.

asymmetry:⁵⁶ Grant that there is value in our having and using an ancestral concept of race as well as a value in our having and using a biological concept of gender. For example, an ancestral concept of race allows us to easily talk about a historical pattern of injustice and disadvantage facing those who are or who descend from colonized or enslaved people groups. And a biological concept of gender allows us to easily talk about a historical pattern of injustice and disadvantage facing those with female reproductive features. Nevertheless, one might argue, any harm done by revising our concept of gender to one based on self-identity rather than biology is outweighed by the need to address the violence and exclusion facing transgender persons – harms that would be alleviated by affirming their identities. Such a balance, arguably, doesn't translate to the racial case. Harm caused by revising our concept of race away from one based on ancestry is not outweighed by a need to address harm to persons like Dolezal. For this reason, even if Ascription 1 were true, it would avoidably reinforce oppressive truths, whereas if Ascription 2 were true, it would not avoidably reinforce oppressive truths.

Again, I'm not offering a complete argument for this asymmetry. My central point is that we can address the puzzle over these gendered and racial identities without appealing to a metaphysical asymmetry – my framework allows us to focus on establishing a normative asymmetry.

I've focused on the case of persons who do not identify with the gender or race they were assigned at birth, but the framework of oppressive truths can be applied to a host of controversies over which social identities we should acknowledge. The framework lends desperately needed nuance to current debates within identity politics, which often contain flatfooted rhetoric that suggests that people *always* or *never* have a right to be ascribed whatever

⁵⁶ Diaz-Leon (2015).

identities they claim. This reasoning is a blunt tool – far too blunt to make sense of important difference between cases where accepting someone’s identity would cause avoidable harmful to others, and cases where it would not. The framework of oppressive truths focuses these debates on the structural normative inquiries that matter most when considering how to pair autonomy and self-determination with social justice.⁵⁷

⁵⁷ Many thanks to audiences at the Penn-Rutger-Princeton Social Epistemology Conference, the Feminist Utopias Conference, University of Michigan (Ann Arbor), Princeton University, Tufts University, Yale University, and the Eastern APA (2017) for feedback during the development of this chapter.

Chapter 3

Real Talk on the Metaphysics of Gender

3.1 Introduction

North Carolina recently embroiled itself in controversy by passing the “Public Facilities Privacy and Security Act.” According to this law, it is illegal for someone to enter a men’s restroom or locker room unless they were assigned the sex ‘male’ at birth, and illegal for someone to enter a women’s restroom or locker room unless they were assigned the sex ‘female’ at birth.¹ This law, in effect, insists that someone assigned male be socially interpreted as a man, and someone assigned female be socially interpreted as a woman, effectively yelling ‘No!’ at trans women’s and men’s explicit and implicit requests to be ascribed the gender corresponding to their self-identity. A sea of outcry soon followed the passing of this bill, and heated debate over this and similar legislation continues across the country.

Clearly, the issue of which persons we should acknowledge and treat as men and women is extremely controversial. And, frequently, these controversies manifest as arguments over metaphysical questions, such as *who is a man (or woman)?* Or *what makes someone a man (or woman)?*² But this manifestation only makes sense if one assumes that what gender someone *in fact has* determines what gender we *ought* to ascribe to that person – that is, e.g., if one assumes

¹ North Carolina State Assembly (2016).

² While I do not assume that there are only two gender categories (men and women), I will often restrict my discussion to these categories for simplicity’s sake.

that one ought to be ascribed *woman* only if one is, in fact, a woman.³ This assumption is pervasive. In addition to everyday contexts such as the streets of North Carolina, it is also found in recent philosophical discussions of the metaphysics of gender. There, authors often take for granted that the correct metaphysical theory of gender (if there is one) will tell us which persons we should acknowledge as (e.g.) men or women.⁴ In short, both popular and philosophical discussions proceed as if the metaphysical facts about who belongs in which gender category constrain the facts about which gender ascriptions we ought to make.⁵

I argue against this assumption. That is, I argue that we should not look to metaphysical theories of gender to settle disputes over gender ascriptions, and we should instead focus on normative questions about how we ought to treat one another. This is because, I argue, the metaphysical facts about gender do not constrain the normative facts about gender ascriptions -- i.e., who we *should* recognize as a man, a woman, genderqueer, etc.⁶ As a result, it might be the case that we *should not* make true gender ascriptions, and *should* make false ascriptions.

After contextualizing and motivating my project in section one, I turn in section two to my first defense of the claim that the normative facts about which gender ascriptions we *should* make are not constrained by facts about what genders persons in fact have: the defense from *social implications*. This argument proceeds in two steps. First, I outline many of the usual social

³ In what follows, I target an assumption that the *metaphysics* of gender (i.e., the real definition of *gender*, or of *women*, etc.) constrains ethical gender ascriptions. Such views seem, to me, more common than the assumption that the *semantics* of gender (i.e. the correct meaning of ‘gender’, or of ‘women’, etc.) constrains these ascriptions. However, I take it that my criticisms of the former view would analogously apply across to the latter view.

⁴ For more on assumed connections between gendered language and the metaphysics of gender (and a critique of one such assumption), see Barnes (manuscript).

⁵ The landscape of these discourses is, I think, nicely captured by sociologist Rogers Brubaker (2016), 415, who writes: “The massive destabilization of long taken-for-granted categorical framework has significantly enlarged the scope for choice in the domains of race, ethnicity, sex, gender, and sexuality. Yet the enlargement of choice has generated anxieties about unnatural, opportunistic, exploitative, or fraudulent identity claims. *And this, in turn, has generated effort to police questionable claims in the name of authentic, objective, and unchosen identities, as well as attempts to justify unorthodox claims in the name of such identities.*” (Emphasis added.)

⁶ I here follow convention in using ‘genderqueer’ as an umbrella term for a variety of gender identities that fall outside the gender binary - e.g., agender, gender fluid.

implications – and more particularly, social statuses – that come with gender ascriptions. Second, I argue that, if the facts about how to justly distribute gender status come apart from the metaphysical facts, the implications are what should (and probably do) matter when considering what gender we should ascribe to someone. That is, if the conditions under which it is just to place on someone the implications of being ascribed *woman* come apart from the conditions under which someone *is* a woman, then we should falsely ascribe *woman* to them. Or, in other words, the social implications of ascribing (e.g.) *woman* or *man* to someone should guide our gender ascriptions, regardless of the metaphysical facts about their gender.

Section three presents another defense of the idea that gender ascriptions should be guided solely by normative considerations (and not by metaphysical considerations): the defense from *oppressive truths*. Here, I argue that, even if it were true that (e.g.) trans men are not men, this would be an oppressive truth. That is, it would be an ontological fact about the social group *men* such that accurately ascribing group membership would lead to an unjust distribution of social status and would reinforce this ontological fact. Moreover, we have the opportunity to *revise* our gender groups by revising our gender ascription practices, reshaping them to better suit our normative goals. This possibility provides yet another reason to think that what ascriptions we *should* make is not constrained by their truth-value.

Combining these two defenses of my core claim, then, we arrive at the idea that we have (at least) two good reasons for thinking that facts about which gender ascriptions are just are not constrained by the metaphysical facts about gender: (1) the social implications of ascribing gender might require us to falsely ascribe genders, and (2) these false ascriptions, with broad uptake, could beneficially revise our gender categories.

This has important implications for approaching disputes over gender ascriptions, such as the one in North Carolina. What weight should we give appeals to the ‘traditional understanding’ of (e.g.) *woman* in order to justify the view that that membership in the group *women* (or *men*) should not be ascribed to transgender women (or men)? And what place (if any) does metaphysical theory have in resolving these disputes? If I am right, we should give no weight to appeals to traditional understandings, because metaphysical truths about gender do not answer normative questions concerning gender ascriptions – it may be unjust to make true gender ascriptions.⁷ This also means that metaphysical theories of gender cannot settle disputes over which gender ascriptions are just.

This last point has two further implications, which I discuss in section four. First, given that the just gender ascriptions are not constrained by truth-value, a metaphysical theory of gender should not be criticized on the basis of excluding a group from (e.g.) the group *women* that we feel *ought* to be ascribed membership in this group. It might be that – whatever the boundaries around this social group are – they are part of an unjust, trans-exclusionary social structure. And second, the category of unethical truths reveals a new methodology for metaphysical theorizing about social categories, including but not limited to gender categories: specifically, one that combines *describing* operative social categories with *prescribing* revision to which social categories are operative.

⁷ Barnes (manuscript) argues that the metaphysics facts about gender can also come apart from the truth conditions for our gender terms. This is consistent with my thesis, as I do not take a stance on the truth conditions for our gender terms, but only argue that *even if* the truth conditions aligned with the metaphysical facts, this would not answer the question of how we *ought* to ascribe gender terms.

3.2 What constrains just gender ascriptions?

3.2.1 The Alethic View

Let's begin by setting out a working definition of 'gender ascription'. Importantly, by 'gender ascription', I don't just mean using particular gender terms such as 'woman' or 'genderqueer' to refer to someone. Rather, I take ascription to be a broader phenomenon, which includes the variety of ways in which we place social labels on each other. For my purposes, I characterize gender ascription along the following lines:

Gender Ascription: S is *ascribed* gender X by an individual R or collective C when R or C **(a)** verbally attributes the property *being an X* to S, or **(b)** treats S as having the social status associated with being an X.⁸

For example, suppose that I walk into a women's restroom. By the above definition, there are at least two ways I might be ascribed woman in this situation. First, someone might refer to me, directly or indirectly (such as through pronouns) as a woman, thereby verbally attributing to me the property *being a woman*. Second, someone might (e.g.) take me to permissibly occupy that women-only space (and to impermissibly occupy a men-only space), thereby treating me as having the social status of a woman. As I will use the term, either of these events would ascribe *woman* to me. Moreover, this working definition leaves open that one might be ascribed multiple genders (in different ways) by the same individual or collective. This phenomenon is common –

⁸ This is not to suggest that (e.g.) being a woman has a single associated status. Intersectional considerations make this quite implausible. I find it most helpful to think of this status as a determinable having many context-sensitive determinants.

for example, it occurs whenever a transgender or intersex athlete is called a woman, but simultaneously barred from competing in women's sporting events.⁹

Now that we have a rough idea of what gender ascriptions are, we might wonder how we can tell what gender we *should* ascribe to someone (assuming we will ascribe some gender or other). One popular response to this question, in both everyday and academic discourses, is that just gender ascriptions are constrained by the facts about who is a man or a woman.¹⁰ Call this the **alethic view**. On this view, just gender ascriptions are *truth-normed* in the sense that the gender ascriptions we *should* make are constrained by the metaphysical facts about what genders people *in fact have*.¹¹ That is, on this view, there is a connection between the metaphysical and normative facts such that we *should* ascribe (e.g.) the gender *woman* to S *only if* S is a woman.¹²

To illustrate the prominence and importance of this view, let's look first at everyday disputes about gender ascriptions, such as those ongoing in North Carolina. In these contexts, we find social conservatives who think we should not ascribe to transgender persons the genders with which they identify. These persons typically base their position on a biological-based understanding of gender categories, which says that the way one's body is determines what gender one has. They then build in an additional assumption: what gender one has determines which gender one ought to be ascribed. This then amounts to the claim that the way one's body is determines which gender one ought to be ascribed. Consider, for example, rhetoric employed

⁹ See, for example, the saga of intersex runner Caster Semenya.

¹⁰ This aligns with Brubaker's (2016), 421 observation that, in response to the destabilization of gender (as well as other) categories, we see "not the disappearance but...the *strengthening* of essentialist, objectivist, and naturalist reasoning."

¹¹ Someone could hold this view because they believe that the metaphysical facts *ground* the normative facts about gender ascriptions. But this is just one option -- someone who holds the alethic view need not commit to any particular explanatory relationship between the normative and metaphysical gender facts.

¹² I take it this assumption is motivated by the initially plausible idea that we always should try to track the metaphysical facts with our concept and language deployment, combined with the assumption that our gender concepts and language indeed refer to real categories. My paper criticizes the first assumption -- see Barnes (manuscript) for criticism of the second assumption.

by opponents of a city ordinance in Charlotte, North Carolina, that (among other things) forbid discrimination on the basis of gender identity, including discrimination with respect to accessing restrooms and locker rooms aligning with one's gender identity:

Keeping separate facilities validates the uniqueness of each gender. Male and female were created distinctly to fulfill vital roles in our society. [We must] protect the safety and sanctity of our men, women, boys, and girls.¹³

White woman identifies as black... Sighted woman identifies as blind. Man identifies as woman. ...We're going to throw out science and DNA?¹⁴

While tortured, the logic in this rhetoric is clear:

- (a) One's biological makeup determines one's gender.
- (b) Someone should be ascribed the gender that they are.
- (c) *Therefore* transgender persons should not be given access to (e.g.) the bathroom corresponding to their gender identity.

In this line of reasoning, (b) assumes a particular connection between the metaphysical facts about gender and the normative facts about gender ascriptions. But rather than question this assumption, progressives -- including progressive philosophers -- by and large focus on attacking (a). That is, their efforts take (b) for granted, and instead target the claim that the metaphysics of

¹³ Video. "Charlotte City Council approves LGBT protections in 7-4 vote" Charlotte Observer, Feb. 22, 2016. <http://www.charlotteobserver.com/news/politics-government/article61786967.html>

¹⁴ Ibid.

gender is such that gender is determined by biology. This focus is apparent in recent literature on the metaphysics of gender, where philosophers sometimes employ the following sort of argumentative schema:

1. Persons in Group X are women.
2. Theory T implies that persons in Group X aren't women.
3. Therefore Theory T is false.

For example, one version of this schema would go as follows:

- i. Transgender women are women.
- ii. Biology-based theories of gender imply that transgender women aren't women.
- iii. Therefore, biology-based theories of gender are false.

The schema is valid and, on the surface, seems to provide a powerful tool against particular views of gender. Employing counter-examples is, after all, a tried and true method of philosophical analysis. But, within philosophical discussions about gender, claims taking the form of (1) often appear normatively motivated, suggesting that the schema in fact deployed is something more like the following:

- 1*. Persons in Group X *should be ascribed woman*.
2. Theory T implies that persons in Group X aren't women.
3. Therefore Theory T is false.

In other words, philosophers have assumed particular stances about what groups of persons *should* be ascribed (e.g.) *woman*, and then taken whether a theory implies that these persons in fact have that gender to count for or against that theory. This suggests that many metaphysicians currently writing about gender operate with (b) as a background assumption -- they assume that the metaphysical facts about gender *constrain* the normative facts about gender ascriptions. In doing so, they take the correct metaphysical theory of gender to *reveal* correct answers to questions such as, “It is unjust to say that trans women are not ‘really’ women?” Or “Should we acknowledge some persons as being neither women nor men?” And so, if a particular metaphysical theory (e.g.) does not include transgender woman in the category of women, the theory rejected for supposedly implying that we should not ascribe *woman* to transgender women.

We find this assumption explicitly endorsed or implicitly employed throughout metaphysicians’ discussions of gender. Elizabeth Barnes, for example, while addressing what she calls ‘deflationist approaches’ to the metaphysics of gender -- that is, approaches on which the metaphysical boundaries of gender categories can be stipulated to ensure alignment with our normative interests -- argues this deflationist approach “seems to confuse direction of explanation for some questions of social justice.”¹⁵ This is because, according to Barnes, they take normative questions about (perhaps among other things) gender ascriptions to explain the

¹⁵ Barnes (2016).

metaphysical facts about gender, rather than taking metaphysics facts to explain the normative facts about gender ascriptions. Barnes argues:

A successful account of gender ought to say that trans women are women; it would be *unjust* not to classify trans women as women. But at least part of that injustice, **on most accounts**, consists in failing to treat trans women as what they are. That is, **it's unjust to say that trans women aren't women because trans women really are women.**¹⁶

Barnes here acknowledges that, on 'most' metaphysical accounts of gender, the facts about whether or not we *should* say that trans women are women are constrained (perhaps even *explained*) by the facts about who is a woman.

This background assumption – the assumption that what gender we should *ascribe* to someone is determined by what gender they in fact have – is the alethic view. This view also is echoed in other recent theories on gender, and particularly ones responding to *social position-based* theories of gender. This family of views says that to be gender X is to occupy a particular social position. For example, on Sally Haslanger's (2000) proposed analysis of gender, to be a woman is to be subordinated because of being perceived as female.¹⁷ Haslanger reaches this view by taking gender to be the thing that explains systematic hierarchical differences between persons perceived as having stereotypical male or stereotypical female reproductive features.¹⁸

¹⁶ Barnes (2016).

¹⁷ Haslanger (2000), my paraphrase.

¹⁸ Admittedly, it is a bit difficult to pin down exactly whether Haslanger takes herself to be giving a revisionary or non-revisionary analysis of gender. There are comments throughout her work, however, suggesting that although she would not argue that there is *one* thing that gender really is, analyses that aim to explain normative truths pick out at least one of the privileged ontological categories eligible to be called 'gender'.

Social position-based views, and particularly Haslanger's view, have been criticized for excluding trans women who are not perceived as female from the category of women. Katharine Jenkins, for example, objects to Haslanger's view on the grounds that many transgender women do not occupy a woman's social position, but nevertheless should be counted as women. Specifically, she writes that it is "an important desideratum of a feminist analysis of gender concepts that it respect [trans] identifications by including trans people within the gender categories with which they identify, and not including them within any categories with which they do not identify."¹⁹ Jenkins, that is, seems to suggest that we should prefer a metaphysical theory of gender that will align with the normative facts about gender ascriptions. For example, if we *should* count as women persons who identify as women, then our metaphysical theory of *woman* must guarantee that persons who identity as women *are* women.

While Jenkins is fairly explicit about assuming the alethic view, this assumption is often more implicit. Mari Mikkola, for instance, also writes in criticism of social position-based theories of gender (like Haslanger's) that:

[Taking] gender to depend ontologically on oppressive social conditions... would not only unwittingly eradicate women and men: doing away with gender should be feminism's political goal. As I see it, both implications are problematic and ones that many ordinary social agents are unwilling to accept... For ordinary thinking, then, the idea that feminists should aim to do away with women and men harbors undesirable ontological commitments.²⁰

¹⁹ Jenkins (2015). In fact, this constraints on a metaphysical analysis of gender has become so widespread that it is known as the "commonality constraint," and is also found in Bach (2012), Tiechman (manuscript), and McKittrick (2015).

²⁰ Mikkola (2011).

This criticism centers on Haslanger's view's implication that, under improved social conditions of gender equality, we would 'do away' with the categories *women* and *men*. Since Mikkola surely does not object to doing away with systematic subordination of persons who are perceived as female, it seems that at least a plausible interpretation of the objection is based on a reading of Haslanger as implying that we should do away with gender ascriptions, which would prevent persons from finding important political solidarity under a common label. Because Mikkola finds this implication problematic, she takes it to provide good reason to reject Haslanger's view. But again, if my interpretation is correct, Mikkola takes the metaphysical facts to determine the normative facts about gender ascription, insofar as she assumes that Haslanger's aim of eliminating our *current* categories *women* and *men* implies that we should stop ascribing gender altogether.²¹

In case you are not yet convinced, Jennifer McKittrick also critiques social position-based theories of gender (such as Haslanger's) on the ground that they "have problematic implications for transgender."²² In order to account for the intuition that transgender women should be counted as belonging to the category of women, McKittrick advocates a theory on which gender is constituted by complex behavioral dispositions. Admittedly, it is unclear whether McKittrick takes this intuition to be a metaphysical or normative intuition. However, it seems uncharitable to read it as a metaphysical intuition, as this would simply be question begging against (or, at the least, talking past) those who advocate biology- or social role-based theories of gender. Yet if it is a normative intuition, then we see yet another example of a philosopher assuming that the gender ascriptions we should make are constrained by the metaphysical facts about what genders persons have.

²¹ The same argument against social constructionist accounts of gender is also made in Bach (2012).

²² McKittrick (2015), 2575.

I could continue, but hopefully the trend is clear. Recent work on the metaphysics of gender often seems to *assume* that the metaphysical facts about who has which gender determine the moral facts about the rightness or wrongness of certain gender ascriptions.²³

3.2.2 The Normative View

The alethic view is the view that the metaphysical facts about gender constrain the normative facts about the rightness/wrongness of particular gender ascriptions. The view I defend, the **normative view**, is the rejection of this assumption. On the normative view, the metaphysical facts do not constrain the normative facts about gender ascriptions -- rather, the normative facts about gender ascriptions are constrained by further normative facts about how we ought to treat others.

Importantly, the normative view is neutral on whether there are any metaphysical categories corresponding to our gender language, as well as the question of how we should go about determining what these categories are, if there are some. Perhaps we must look for the ontological category that makes the most sense of the historical and current purposes for our gender talk.²⁴ Or perhaps we must look at English gender words and their current use. Perhaps the correct metaphysical theory of gender (if there is one) does not provide truth conditions for our natural language gender terms (e.g., ‘woman’).²⁵ Perhaps the metaphysical facts about gender categories are independent of the normative facts about gender ascriptions. Or perhaps

²³ One might be tempted to interpret these projects as simply *different* projects, rather than projects in disagreement about the correct metaphysical account of gender. But the fact that these authors present themselves as disagreeing with each other, rather than just stipulating different uses of the term ‘gender’, makes this implausible.

²⁴ Such a method would make sense, for example, for someone who assumes both semantic externalism and a version of scientific essentialism. See Haslanger (2012), 14 for more on this subject.

²⁵ See Barnes (manuscript).

the correct metaphysical facts are constrained by these normative facts. For the purposes of this paper, I remain neutral on all of these issues.

The point I want to make is just this: What morally matters in making gender ascriptions is not the metaphysics of gender -- it is a normative issue, and one that cannot be answered by metaphysical analyses of gender. Maybe there is a single privileged answer to the question *What makes someone a woman?* or *Is gender binary?*. Maybe there isn't. According to the normative view, the answers to these questions will not reveal what we should care about when we ask *Should we refer to trans men as men?*, *Should we recognize gender categories such as genderqueer?*, and so on. Concern for justice, and not metaphysics, is needed to answer these questions.

Before moving on, I think it is important to clarify two projects I am *not* pursuing. First, I do *not* (and, to be honest, dare not) take a stand on metaethical disputes concerning ethical naturalism and non-naturalism. That is, my thesis should not be read as having any implications for the much broader question of whether all normative facts are wholly grounded in non-normative facts. I simply am making the restricted claim that the normative facts about gender ascriptions are *not* wholly grounded in non-normative facts about the membership conditions of gender categories. Second, and relatedly, I am not suggesting that the metaphysical facts about gender do not explain or constrain *any* normative facts. Consider again Haslanger's account of the membership conditions in the category *women*. On her view, someone is a woman if and only if they are systematically subordinated by men (among other things). If this is right, then this at least seems to wholly ground the normative fact that we ought to eliminate or drastically revise this category. None of this is in conflict with my present thesis. Or, in other words, even if

metaphysical facts about gender wholly ground *some* normative facts, this does not rule out that they do not constrain normative facts about gender ascriptions.²⁶

3.3 Social Implications

As discussed above, it is commonly assumed that metaphysical facts about gender constrain normative facts about gender ascriptions. In this section, I hope to persuade you that this is not the case, because the normative implications of ascribing someone a particular gender should outstrip concerns about whether that person meets certain gender category membership conditions. The central claim I will defend is simply this:

- (A) The metaphysical facts about gender do not constrain the normative facts about gender ascriptions.

To motivate (A), I return my focus to ordinary discourse surrounding gender. In particular, I argue that:

- (B) Normative facts about how we should treat others (and not metaphysical accuracy) are what matter when determining which gender ascriptions are just.

²⁶ Thanks to Elizabeth Barnes for pressing this point.

I hope this thesis strikes you as intuitively correct. After all, it seems unlikely that, with so much on the line with respect to transgender persons' civil liberties and stigmatization, the many ongoing disputes about gender ascriptions are determined by independent or prior facts about who is (e.g.) a man or a woman, rather than facts about who should have which civil liberties and gendered resources (and when), and who may be subject to certain gender norms (and when).

Before delving into the issue, though, let's warm up to it by considering an analogous issue: marriage. Let's suppose, for the sake of argument, that in the midst of national debate over legalizing same-sex marriage in 2014, a philosopher presented an extremely compelling argument that, e.g., given empirical, historic, and linguistic considerations, the category *marriage* is indeed an intended life-long union between a cisgender man and a cisgender woman.²⁷ Of course, I cannot definitively say that this would not have ended gay rights activists' efforts to extend marriage ascriptions to same-sex couples. But hopefully you will share in the intuition that such a reaction would be extremely unlikely, and moreover, that activists *should not* have this reaction. And this is because what matters in how we answer the question of whether to ascribe *marriage* to certain same-sex relationships is not, at bottom, the metaphysical theory of what the category *marriage*. What matters is *justice* for same-sex couples -- both with respect to the pragmatic rights afforded to individuals whose relationships are ascribed membership in the category *marriage*, as well as with respect to cultural endorsement and acceptance of these relationships.²⁸ What matters, in other words, are the *social implications* of

²⁷ If this seems too unlikely, imagine that God appears and informs us that 'marriage' rigidly designates this natural kind (but makes no claim about the morality of same-sex unions). Also, by 'cisgender', I just mean a person who identifies with the gender traditionally associated with the sex assigned to them at birth. (E.g., Someone assigned 'male' at birth who identifies as a man.)

²⁸ Gideon Rosen, responding to the objection that same-sex marriage would 'redefine' marriage, also made a nearby point. "If opponents of SSM cannot explain the charge of redefinition and its moral force, they should drop it. The only defensible claim in vicinity is that a move to SSM would involve a change in the law and in society. But when the point is put in this way, it no longer sounds like an objection. The substantive question is whether this change

having a relationship count as a marriage, regardless of whether these implications push in the direction of revising the conditions under which a relationship is a marriage.

Similarly, we can suppose a brilliant philosopher comes up with an argument that nearly definitively shows that, given empirical and linguistic considerations, the category *women* is indeed biologically based, such that one must have particular biological features in order to be a woman.²⁹ Should this end debates between (e.g.) those who advocate for biological-based, social role-based and identity-based concepts of gender? Or between those who advocate for and against trans-inclusive bathrooms? Again, I don't think so. And this is because what is, *at bottom*, at stake in how we ascribe gender are questions about what ascriptions would create socio-political change for the *better*, not questions about which ascriptions would maximize metaphysical accuracy.

Why is this, though? That is, why is it the case that these disputes, which at the surface appear to be about metaphysical facts, often are not and *should not* depend on such facts? In short, I think it is because, with respect to social categories like *marriage* or *women*, the social implications of being treated as a member of these categories should outstrip concerns over whether someone meets certain membership conditions.³⁰ And, as a result, the metaphysical facts about gender will not settle what ascriptions we *should* make. Whatever being a woman *requires* (if there is such a thing), it will not answer questions about how we *should* conceptualize

would be for the better. The rhetoric of redefinition obscures this fact; it has no place in a serious discussion of the issue.” <http://dailyprincetonian.com/opinion/2009/11/rejecting-the-rhetoric-of-redefinition/>

²⁹ An example of such a view (and a fairly compelling one, I think) is found in Bach (2012).

³⁰ For more on this possibility, see Brian Epstein (forthcoming) on the ontology of groups. As Epstein points out, the conditions required for a group of a certain type to exist may not include all the essential features of that group. For example, it is not required that a group be essentially rule-establishing for the group to be of the type ‘faculty committee’, though this might be an essential feature of a particular group of that type. I am not taking a position on whether these social features are essential to gender groups, but simply pointing out that, even if they are, it does not mean that these features are necessary to these groups *being gender groups*.

gender, whether and when we *should* count someone as a woman, or what (if any) social meaning *should* accompany belonging to this gender category.

To support this claim, let's consider some of the weighty associations and implications accompanying gender ascriptions. Once we do this, I think, it will be apparent that the conditions under which these associations and implications are permissibly placed on someone can be determined independently of and may come apart from the conditions under which that persons has the ascribed gender.

So what are these associations and implications, and why do they matter when determining ethical gender ascriptions? To fill out this picture, we can look to the rich body of *culturally shared* associations and implications associated with gender categories, and so with ascribing someone a particular gender. This is most easily approached through the case of *verbal* ascriptions, because it is clear that gender terms have a rich body of cultural, emotional, and political associations. David Plunkett describes such terms as ones where the “associations will often be harder to shake than the specific application-conditions associated with the term itself, or the specific meaning it has at a given time.”³¹ That is, terms like ‘woman’ and ‘man’, like ‘freedom’, ‘rape’ and ‘person’, have robust associations that remain constant across changes in their patterns of use. And, centrally, given the associations of a particular term, that term can be paired with *better* or *worse* patterns of use. The key idea here, then, is that when a term carries meaningful associations that -- at least to some extent -- can float free of the term's pattern of use, it can *morally matter* how we ascribe this term. That is, it can be more or less just *how* and *when* we ascribe membership in these categories. And the just conditions might come apart from conditions under which someone is in fact a member of the category that term refers to.

³¹ Plunkett (2015), 843.

We can fill out this point with the example of ‘parent’. ‘Parent’ is a term that has vivid connotations of (among many other things) having guardianship of a child, being a child’s caregiver, having responsibility for supporting a child, or having had these roles. And, importantly, these associations are welded to the term so strongly that they, to some extent, remain attached to the term regardless of changes in the term’s patterns of use. Given this, it is reasonable to think that our use of ‘parent’ *should not* exclude persons who stand in a relationship to a child that seem deserving of these connotations, even if this conflicts with regular usage. That is, it is reasonable to think that we should look for the pattern of ascribing *parent* that best suits the associations of the term ‘parent’, even if this means revising our patterns of use, and even if this means ascribing this term to someone who is not a member of this category.

And yet, as Haslanger discusses, ascriptions of parenthood often fall short of this goal. Many persons operate with a background ideology on which someone truly is a parent only if they are an “immediate progenitor” of a child.³² That is, they hold something like a biology-based theory of *parenthood*, and therefore ascribe the term ‘parent’ only to biological progenitors of children, excluding persons who have guardianship over, care for, or are otherwise responsible for a child. But such a practice, regardless of whether or not the *parent* turns out to be this biological category, seems *wrong*. It withholds from (e.g.) adoptive parents, stepparents, foster parents, grandparents, and other family members raising children who are not their own biological progeny a status that they *ought* to be given.

Anyone looking will not be hard pressed to find cases where who is a parent (according to either biological or social theories of *parenthood*) comes apart from who we ought to ascribe

³² Haslanger (2012), 389.

count as a parent. For example, in the 2016 summer Olympics, commentator Al Trautwig said on-air that gymnastics star Simone Biles “was raised by her grandfather [Ron] and his wife [Nellie], and she calls them mom and dad.”³³ In response to criticism, he later on Twitter dug in his heels, saying that, “[Her grandparents] may be mom and dad but they are NOT her parents.” Such claims clearly seem *wrong*, but not because – or at least, not primarily because – they are *descriptively incorrect* or *untrue* (even if they are). Trautwig’s claim is criticizable for its implications concerning the importance and nature of the relationship between Simone, Ron, and Nellie, and not for expressing a false theory about *parenthood*. That is, Trautwig’s comment was *morally wrong*, even *if* metaphysically correct, because it denied the associations of ‘parent’ (as well as the purpose of laws and norms surrounding the category *parents*) to persons who *should* be ascribed membership in this category.

Like ascribing *parent*, ascribing gender, including through non-verbal gender ascriptions, confers robust social statuses. These ascriptions are packed with associations related to a given society’s operative stereotypes of masculinity and femininity. This much is indisputable, given the thousands of research articles in linguistics, social psychology and sociology that rely on these associations to test gender bias and stereotype threat.³⁴ Moreover, and as the current debate over recognition for transgender persons illustrates, these statuses are, within a context, welded to gender ascriptions regardless of changes in our ascription practices. For example, when we observe cultural disputes over whether someone assigned a female sex at birth but identifies as

³³ See US Magazine, “Simone Biles Responds to NBC Sportscaster’s Comment About Her Parents”. August 10, 2016. Simone’s grandparents adopted her at age four, and have been her primary caregivers ever since.

³⁴ For example, see Silviera (1980) “Generic masculine words and thinking”, Cadinu, Mara et al.. (2005) “Why Do Women Underperform Under Stereotype Threat? Evidence for the Role of Negative Thinking”. *Psychological Science* 16.7: 572–578, among many many others. See also Pryzgoda (2000), 555: “Hamilton (1988) argued that both hearing and using masculine generics creates male bias in people’s impressions and images of the subject under discussion... Gendered word choices have been shown to influence comprehension, memory (Crawford & English 1984; Harrison 1985), and attitudes toward women (Crawford, Stark & Renner 1998; Miller & Swift 1980).”

male ‘is a man’, this dispute is not over whether we should change the social status of men. (Though certainly we should talk about that too!) Rather, it seems that this disputes is about whether we should revise our dominant patterns of ascribing *man* such that the accompanying status may be conferred on someone not born biologically male. And relatedly, in disputes over when we should ascribe *genderqueer* or *agender*, it seems that part of what we are asking is whether (or how) we should revise our ascription practices such that, for some persons, we refrain from conferring a gender status. In short, given the rich statuses conferred through gender ascriptions, it is far from trivial to ask, ‘What is the all things considered *best* practice for ascribing gender?’.

It is important for this point that these ascriptions need not have obvious practical or legal effects in order for persons to care deeply about when they are deployed. (Though, in section three, I will turn to examine non-obvious practical implications of gender ascriptions.) For example, consider the recent event in which television personality Piers Morgan, in an interview with trans woman Janet Mock, said that she was “born a boy.”³⁵ Morgan was instantly rebuked, and subjected to intense criticism across social media. And yet it is an interesting question why this remark was deeply offensive. It seems that the source of offense could not be denying Mock particular rights, or even current recognition as a woman. Instead, their disagreement seemingly centers on the simple question of whether we *should* withhold ascribing a gender, and all of its accompanying implications, to someone (even a past stage of someone) who does not want this gender to be ascribed.

One reason for this might be that ascribing *boy- or man* to someone implies (or was taken to imply) that the person always will be a man. Or perhaps, in this case, it implied a trivialization of

³⁵ See, e.g., Hess, A. (Feb. 7, 2014). “Piers Morgan’s Interview with Janet Mock Was Not a Failure of Sensitivity. It Was a Failure of Reporting.” *Slate*.

being a trans woman – as if one’s gender is something one changes on a whim, like one’s college major or haircut. Frankly, it is hard to pinpoint the source of offense, given that much of the criticism simply asserts that Morgan was *descriptively* wrong, and that Mock ‘has always been a girl/woman’. It seems clear, though, that this cannot be the source of offense – after all, Morgan could have gotten Mock’s birthplace or alma mater wrong without offense. Something about what it *implies* to call Mock’s infant self a boy, above and beyond issues of descriptive accuracy or pragmatic effects, seems to be the reason why Morgan was wrong to do so.

That said, a second, related reason why we should think that everyday disputes over gender ascriptions convey normative disagreements concerns the clear pragmatic (including legal) implications of gender ascriptions. For example, in asserting, “X is (or is not) a woman,” we do more than make a claim about whether we take X to have certain features, or apply particular feminine associations and stereotypes to X. In addition, we license using certain pronouns for X, taking X to be subject to particular social norms, and providing X with any rights or responsibilities that come along with being classified as a woman. In many cases, in addition to *licensing* these things, our ascriptions themselves can *do* these things. That is, our gender ascriptions often have practical effects. To get a better grip on this phenomenon, consider the following practical effects that (for better or worse) are often caused or licensed by ascribing *woman* to someone (call them ‘Sam’):

- Sam is subjected to social norms that require Sam to use feminine pronouns (she/her/hers), and adopt a feminine gender presentation (shaving legs and underarms, wearing ‘women’s’ clothes, etc.).

- Sam is subjected to social norms requiring Sam to have only male sexual partners, and to engage only in stereotypically ‘feminine’ activities and occupations.
- Sam is subjected to legal and social norms requiring Sam to *not* use men’s spaces, (e.g., men’s restrooms and locker rooms, men’s shelters, men-only clubs and schools), as well as men’s social institutions (e.g., men’s scholarships or government programs, men’s athletic teams).
- Sam may criticized, ostracized, legally punished, or put at risk of physical or emotional violence for violating any of the above legal or social norms.

You get the idea. Regardless of whether the metaphysics of *woman* (if there is such a thing) builds in gender status, commonly shared gender ideology means that this status is conferred when we ascribe *woman* to an individual -- including policed expectations that they will conform to gender-related norms or laws.³⁶ The particular expectations I’ve listed are, of course, not universal, and will vary depending on one’s time, place, and other social identities. Nonetheless, it is widely assumed that persons who are counted as women will dress and behave in particular ways, and will have particular rights, responsibilities and interests. This much is not controversial -- the fact that gender categorization confers social status is heavily evidenced by

³⁶ I use ‘norms’ here in the sense of ‘typical standards’, which need not be viewed as wrong to transgress.

empirical data on (e.g.) stereotype threat, gender discrimination, and childhood development of gender identity.³⁷

Both our shared socio-historical and cultural associations, as well as our shared understanding of the pragmatic implications attached to gender categorization suggest that how we sort others into these categories is incredibly important, and not only for transgender persons. In debates over single-gender spaces (such as restrooms and locker rooms), it seems that a variety of interests are in play – some legitimate (such as decreasing sexual predation) and some illegitimate (such as hate-based discrimination against transgender persons). Given this, how and when we ascribe genders has implications not only for how transgender persons are treated, but also, how the variety of interests at stake are framed and discussed. Moreover, I suspect that the incredible weight of our gender ascriptions at least partially explain why, in LGBTQ and ally communities, gender *identity* is taken to be the primary determining factor of appropriate gender ascription – it might seem wrong to ever confer a gender status on someone without their receptivity to that status.³⁸ And, quite plausibly, gender identities provide that receptivity, whether explicitly or implicitly.

To further illustrate the importance of these implications, consider again the fierce, ongoing debate over bathroom access for trans persons. The rhetoric around these debates seemingly is not over the facts about individuals' biology, gender-identity, or gender-expression, nor the standard usage of 'man' and 'woman'. In fact, it does not seem to be about the terms 'woman'

³⁷ Steensma, et al. (2012). "Most children develop the ability to label their own and others' gender between 18 and 24 months. This ability is related to increased gender typed preferences such as the preference for stereotyped toys... For most children, gender identity is largely congruent with their gender role behaviors."

³⁸ By this, I do not mean that it would be wrong to, e.g., call someone a 'woman' unless they were comfortable with the associations of being ascribed *woman* -- rather, I mean that, given all the implications that *come with* being ascribed *woman* (whether or not we like it), it might seem that we should not call someone a 'woman' unless they consent to being so called. Someone might even consent to being so called *for the purpose* of transgressing what they take to be oppressive accompanying norms (see Jenkins (2015)).

and ‘man’ at all. (One can well imagine a conservative saying “Fine, call them ‘women’ if you insist! But they aren’t *really* women, and I still don’t want them in women’s restrooms!”)³⁹

Rather, it seems to, at bottom, a *normative* debate between those who think that one must have stereotypical female genitalia in order to have a women’s social status, and those who think that it is sufficient to self-identify as a woman to have to have this status. But ascriptions *woman* and *man* (usually via linguistic ascriptions of ‘woman’ and ‘man’) remain the focus of this normative debate insofar as both sides *share* the assumption that being ascribed *woman* confers the right to access to women’s spaces.

More importantly, even if such debates were not about these practical implications of gender ascriptions, they *ought* to be. Consider that whether we ascribe woman to a transgender woman significantly affects whether or not she will have unfettered access to preferred language, women’s spaces and activities, women’s social institutions, and so on, as well as whether she will be viewed as transgressing social norms (and perhaps subject to violence) if she chooses to adopt a feminine gender expression or use these spaces. Asking whether we should grant this social status to a transgender woman should not rest on a metaphysical debate, particularly given that we have no guarantee that we have constructed these categories in a just fashion. Nor should it rest on a lexical debate, or anatomical facts – this question should rest on how we ought to treat others, and what autonomy persons ought to have over their gender status. It should not rest on determining the ontological contours of our gender categories.

³⁹ This sentiment was, in fact, explicitly seen in recent debate over the ‘HERO’ ordinance in Houston, TX, which would have prohibited discrimination based on (among other things) gender identity. Opponents of HERO were quoted as saying both “No men in women’s bathrooms, no boys in girls’ showers or locker rooms,” as well as the more specific, “Anybody with a penis, I don’t want them in the ladies’ restroom.” This shows, I think, that while the term ‘man’ is central to this debate insofar as the conservative group took this term to refer to all and only persons with particular anatomical features, the word was not fundamentally important -- at bottom, this was a dispute over whether certain persons ‘with penises’ should be given access to women’s restrooms. [<https://www.washingtonpost.com/news/morning-mix/wp/2015/11/03/why-houstons-gay-rights-ordinance-failed-bathrooms/>]

This position also makes good sense of why resistance to distinguishing between sex (as an anatomical feature) and gender (as a social or psychological feature) on metaphysical grounds seems to be missing the point. After all, those who advocate for this distinction do so because, they argue, it is important for achieving certain normative goals: e.g., illuminating the subordination of female persons, or making sense of transgender experiences.

One can well imagine social conservatives scoffing at this movement, and pointing out that to be a woman *just is* to be a female human. But this response clearly would be overlooking what *should* matter when considering whether to adopt a sex/gender distinction. What matters is whether revising our understanding of gender to achieve these goals is something we *should* care about doing. So to resist this distinction on the basis of one's preferred metaphysical story about gender is to simply fail engage the important normative issues that should decide whether or not we adopt the distinction.⁴⁰

This point is nicely illustrated in the following quotation from Catherine McKinnon:

Male dominant society has defined women as a discrete biological group forever.

If this was going to produce liberation, we'd be free... To me, women is a political group... Anybody who identifies as a woman, wants to be a woman, is going around being a woman, as far as I'm concerned, is a woman.⁴¹

Here, McKinnon explicitly acknowledges that 'woman' has been "forever defined" as a biological group, as well as the fact that she nevertheless rejects that understanding in favor of a

⁴⁰ Of course, there are other normative reasons one might resist the social position based usage of 'women', such as the concern that it would exclude many trans women. But I take it that, at least in the early waves of feminism, this was not the main source of resistance.

⁴¹ On Century Avenue (2015).

politically motivated understanding. Given this, we should not expect McKinnon, or others with normative commitments at stake in these disputes, to be moved by any metaphysical argument that *woman* is (e.g.) a biological property, or any particular sort of property that conflicts with their normative commitments about who should be acknowledge as women. After all, they recognize, these are not the considerations that should matter when we ask how we should distribute gender status to persons. Rather, we should expect what McKinnon demonstrates: rejection of the dominant view about this ontological category, and a continued insistence that we should adopt (what the speaker takes to be) a politically necessary understanding of gender categories.

This point equally applies to the above discussion of bathroom access for trans persons. Given what is at stake in this disputes, it would be overlooking what matters to look to metaphysical analysis of gender to resolve these disputes. But this is not to say that concepts of gender are unimportant -- because of *shared* cultural understanding that someone who is a woman has access to women's spaces, these disputes take the form of arguing about which concept of gender we should adopt. But it would be myopic to think that, because of this, the dispute fundamentally is about which concept best tracks the metaphysics of gender. What fundamentally matters is whether (e.g.) persons who were assigned male at birth and identify as women *should* be conferred the social status aligning with their identity. What is at stake, in other words, can be stated without ever using gender terms.⁴² And this is why questions about the metaphysics of gender (or semantics of gender terms) are not relevant for settling on a position within these disputes.

⁴² Except, of course, in intensional contexts, such as 'The legislators believe that 'women' should only apply to persons with XX chromosomes.'

Notably, this is not unique to the case of gender: in cases where a category is in dispute, and this dispute has important and immediate implications, the source of the dispute usually is not disagreement about the current ontological boundaries around these categories.⁴³ (For example, see historical and current boundary disputes over who/what should be counted as *marriage*, *sexual harassment*, *black*, or *disabled*.) Similarly, in these cases, metaphysical analysis is not relevant – it should not (and probably could not) be the source of conflict resolution.

Much more, of course, could be said. One might wonder why, if I am correct that what matters in these disputes is normative, there is so much dispute about whether (e.g.) trans women are ‘really’ women, even apart from practical considerations such as bathroom access. To fully address this question, I think, one would need to do a fuller analysis of implicit misogyny and reinforcement of patriarchal social systems. Why, for example, is there more resistance to counting trans women as women than there is to counting trans men as men? How does denying particular gender ascriptions to trans persons reinforce a system that privileges cisgender persons, and particularly cisgender men? I suspect that, once we look more carefully at these questions, we will find evidence that attempts to reinforce patriarchal social systems sometimes manifests in the form of resistance to shifting gender ascriptions to include trans women or men, resistance to making non-binary gender ascriptions, and so on. Of course, an in-depth analysis of this phenomenon is beyond the scope of this paper, but I hope it seems plausible to you that even these disagreements are, at bottom, normative disagreements closely related to conservatism about gender roles, transphobia, and patriarchal social systems. And, for answering *those* questions, the metaphysics of gender has nothing to say.

⁴³ For more on these sorts of disputes, which often manifest as conflicting usage of a contested term associated with the category -- (e.g. ‘She is a woman.’ ‘No she isn’t.’) – see Burgess and Plunkett (2013a) & (2013b).

3.4 Oppressive Truths

In the previous section, I argued that everyday disputes over gender ascriptions should not and likely could not be settled by looking to the metaphysical facts about gender (even if there are such facts). These disagreements, I suggested, arise because of *shared* understanding of the statuses conferred by gender ascriptions, and disagreement over *which persons* should have which gender status. As such, illumination of the metaphysics of gender should not settle disputes over gender ascriptions. Rather, the *social implications* of our gender ascriptions provide us with a moral reason to make ethical gender ascriptions...even if they are false.

This position, if correct, has striking implications for ongoing political (or everyday) disputes concerning gender, as well as for philosophers working on the metaphysics of gender. First, it implies that the *truth* about gender ascriptions does not matter when determining appropriate gender ascriptions. Instead, what matters is whether our ascriptions are *just*. Second, it implies that metaphysicians working on gender should not dismiss a view simply because it implies or describes an oppressive state of affairs -- e.g., that trans women are not within the category *women* or that gender is binary.⁴⁴

I will return to the latter implication in next section. For now, I want to address those who remain squeamish at the idea of making false gender ascriptions simply because they avoid unjustly conferring gender status. After all, striving after *truth* is a foundation of intellectual inquiry. We often assume that the ‘truth will set us free’, and recoil at attempts to achieve a good end by spreading falsehoods. And yet my previous argument suggests the opposite: certain truths about gender could be oppressive and, if they are, they should be suppressed. To defend this

⁴⁴ I take it that this point is further reinforced by Barnes’s (manuscript) argument that metaphysical and semantic questions about gender and gender terms (respectively) come apart.

claim from a different angle – one that might be more persuasive to the champion of Enlightenment ideals – we can consider a second argument: the argument from *oppressive truths*. This argument says, in effect, that ontological facts about gender groups may be *oppressive* in the sense that true group ascriptions – that is, ascriptions that align with the ontological facts – create an unjust state of affairs, and also *reinforce* these oppressive facts. Moreover, if these facts are oppressive, then we can shift the ontological facts about gender social groups by making false gender ascriptions.

Let's begin by considering the possibility that there *is* a single correct metaphysical analysis of gender (or of particular genders). In fact, we can go ahead and stipulate, for the purpose of the present argument, that gender categories are trans-exclusive, such that the ascriptions "Caitlyn Jenner is not a woman" or "Aydian Dowling is not a man" are true.⁴⁵ Should this affect our resistance to trans-exclusive use of gender terms, or our support for altering gender language use to be trans-inclusive?

In the previous section, I argued that the answer univocally is *no*, it should not. And this is because the metaphysics of gender does not determine the ethical ways to represent and discuss individuals' genders – the normative implications of these ascriptions might come in conflict with the metaphysical facts.

Now, focusing on *verbal group ascriptions*, I wish to address those who still worry that we have a *prima facie* obligation to assert only true things, and this obligation is not outweighed in the case of gender ascription. We can continue to load the scales on the side of making ethical ascriptions regardless of their truth-value. In order to do this though, I need to establish a category for a particular sort of truth: the category of *oppressive truths*. By establishing this

⁴⁵ Jenner is a transgender woman; Dowling is a transgender man.

category, and showing that the claim ‘trans women are not women’, if true, would be an oppressive truth, another reason to make only just gender ascriptions becomes clear.

Oppressive truths, as I will use the term, are truths about social group membership such that *correctly* ascribing social group membership – that is, counting someone as a group member in accordance with the ontological facts about that group – unjustly distributes social status. So, for example, consider the social group *slaves*. In the 1850’s, it was true that some people were slaves. But this truth is, on my view, an oppressive truth: Truly ascribing membership in this category unjustly conferred a particular social status on persons.

In order to understand oppressive truths, I should clarify how I here understand social groups such as *women*. These groups are established when, according to a shared ideology (roughly, a robust system of beliefs, attitudes, and behaviors), particular features have social significance and so persons with those features are grouped together under a label. When someone is systemically recognized as a group member, they are thus conferred the social status associated with that group.⁴⁶ So, for example, the feature of having a female body, on many shared ideologies, is highly socially significant. In many cases, persons perceived to have female bodies are grouped under the label ‘women’, and persons who are systemically recognized as women conferred a particular status. (We previously explored some of the features of this status.)

Importantly, this point is neutral on the debate over the semantics of terms such as ‘woman’ – it only requires the quite plausible claim that being systematically called a ‘woman’ (or a translation thereof) amounts to being systematically conferred the social status associated with the group *women*. This is because, regardless of what the term ‘woman’ *describes*, someone who is systematically referred to as a ‘woman’ is conferred the social status associated with the group

⁴⁶ I say ‘when recognized’ in order to allow for cases of passing.

women. So even if the semantics of the term ‘woman’ is, in a context, such that ‘woman’ just means ‘person with XX chromosomes’, labeling someone a woman still ascribes a group membership and so (taken systematically) confers on them the associated status. What these words can *do*, in other words, may go beyond what they *mean*.⁴⁷

The idea behind oppressive truths, then, is that group ascriptions such as ‘Sam is a woman’ could be *true* (given the boundaries of the social group *women*) while unjustly conferring the associated social status on Sam.

We’ve already looked at the ways in which gender ascriptions can unjustly confer gender statuses. My goal for the remainder of this section, then, is to convince you that (1) such gender ascriptions can be true and can reinforce unjust ontological facts about who belongs to gender groups, and (2) false gender ascriptions can work to combat this injustice by revising the facts about who belongs to these groups.

First, why think these ascriptions can be true? I want to avoid getting into the weeds of the various semantic views about ‘woman’.⁴⁸ Instead, I want to step back and think about how we might go about determining the truth-conditions of terms like ‘woman’ and ‘man’. The two most obvious ways we might do this are to either look at how competent language speakers *use* these terms, or to look at what competent language *take* these terms to mean – that is, to adopt either an extensional or intensional approach to these terms’ meanings. On either approach, though, one is almost certain to find linguistic communities where both the extension and intension of these terms is such that these terms must refer to a group that is not co-extensive with the group of persons who ought to have the associated statuses.⁴⁹

⁴⁷ For more on this distinction, and its importance for feminist philosophy of language, see Bauer (2015).

⁴⁸ But see, e.g., Saul (2012) and Bettcher (2013).

⁴⁹ A more thorough argument for this claim is provided in Chapter 1.

For example, suppose consider linguistic communities that *take and use* the term ‘woman’ to refer only to persons who are perceived to have female bodies. Even in these contexts, the term ‘woman’ has the function of ascribing membership in a social group – namely, the group that has a particular social status as a result of having female bodies. In these contexts, then, it is highly plausible that many assertions about who is or is not a woman would be *true*, would unjustly confer this status on genderqueer persons and transgender men, and would unjustly deny the status to transgender women who are not perceived as female. That is, such assertions plausibly are true in some contexts, even though they unjustly distribute the social status associated with gender groups.

My next point is that such ascriptions additionally can reinforce oppressive truths: that is, group ascriptions can reinforce unjust ontological facts about social groups. At first, it might seem strange and uncomfortable to think that truth and ethics can come apart to the point that expressing truths would reinforce oppression. But this strangeness dissipates once it is clear that the oppressive truths I am presently concerned with are facts about *constitutive* socially constructions.⁵⁰ By *constitutively* socially constructed, I follow Haslanger in adopting the follow definition:

Constitutive [social] construction: Something is constitutively [socially] constructed iff in defining it we must make reference to social factors.⁵¹

⁵⁰ I want to leave open that there may be unethical states of affairs about natural categories that we have *causally construction* that we also may have a duty to not assert, at least in some contexts. (Perhaps an example of this would be, ‘Males are stronger than females’, assuming this biological fact is due to gendered distribution of resources and activities across human history). Of course, this raises many questions of what (if anything) distinguishes natural and social categories. I don’t take a stand on that here. I only mean to leave open that, if there *are* natural categories like these, there may be oppressive truths about those categories. In these cases though, I suspect the truths will be unethical to assert only because of their causal consequences, and not because these assertions partly constitute the unethical state of affairs.

⁵¹ Haslanger (1995).

As I interpret it, this amounts to the claim that, if something is constitutively constructed, facts about its existence are grounded in facts about social factors – for example, patterns of collective intentions or actions within a given environment. With respect to gender ascriptions, then, the relevant social construction is the category of persons who belong to the social group *women* -- that is, those whose have a particular gendered social status as a result of being systemically counted as women (for example, by being called ‘women’).

Importantly, I think we get this result regardless of one’s view on the semantics of gender terms like ‘woman’. For example, if the term ‘woman’ just means ‘an adult human with a female body’, the claim ‘Trans women are not women’ might, strictly speaking, only say that trans women are not adult humans with female bodies. But such claims, even if this is one’s semantic view of ‘woman’, clearly do more than simply *describe* trans women as having a certain sort of body. Because this is a socially significant feature, these claims undeniably also have the pragmatic effect of *implying* that trans women do not belong to the social group *women*. That is, they ascribe non-membership in the group *women*, as they indicate that trans women do not have features with the social significance of *women’s* features.⁵² And this is an unjust state of affairs since, as I argued in section 2, trans women ought to be included in this group and so have the social status of group members.

Gender ascriptions are just one example of claims that, even if true, either describe or imply some unjust constitutive social construction. Consider other examples:

⁵² On alternative views of the semantics of ‘woman’ – for example, ones where ‘woman’ picks out persons who have the social status associated with this physical feature, such claims (if true) would *describe*, and not merely imply that trans women do not belong to the social group *women*.

- i) Blacks are not voters.
- ii) Marriage is the union of one man and one woman.
- iii) Non-consensual marital sex is not (legal) rape.

These claims do not directly refer to human actions. Nevertheless, they describe or imply an unjust and constitutively constructed state of affairs. This is because each claim either *describes* an unethically constructed social category (e.g., *voters*), or because it *implies* an unethically constructed state of affairs (e.g., non-consensual marital sex not being a punishable crime).

More particularly, consider (i). This sentence once expressed something that was made true by an unjust social system that *persons built together* – namely, a governmental system that systematically privileged white persons over black persons. The same is true of (ii): like *voter*, *marriage* is a constitutively constructed category that we built in an unjust fashion, systematically privileging straight persons (and particularly males) over gay persons. Sentence (iii), in contrast, implies but not does describe an unethical state of affairs. Stipulated legal definitions of rape, after all, do not refer to constitutive constructions – they describe features of actions that one human being might do to another, and give any actions with those features the label ‘rape’. That non-consensual marital sex does not have one of those features is not itself unjust. But an unjust state of affairs is *implied*: namely, that non-consensual marital sex does not have the social significance of recognized forms for rape. And this category – the category of actions with the social significance associated with rape – is constitutively constructed. It is up to us which actions have this significance.

All of this shows why seemingly non-normative ontic claims can in fact describe or imply normatively evaluable states of affairs – namely, constitutively constructed ones. And this is

closely related to the second feature of these claims: expressing them in certain contexts (verbally or otherwise) may *reinforce* the unjust states of affairs at issue.

To illustrate this point, consider the example of a border. Borders, clearly, are constitutive social constructions. While a stream exists independently of what we do, if that stream also is a border, it is a border by virtue of our collective attitudes and actions within that particular material environment. Now consider an ethically contentious border, such as the Berlin Wall. Persons understanding, acknowledging, and treating the Berlin Wall as a border made it such. (Of course, these attitudes and actions are complex and varied, and involved attitudes toward powerful governing bodies' pronouncements and actions, but for our purposes, we can gloss.) And this means that speech actions such as, "The Berlin Wall is the border separating East and West Germany," soldiers guarding the wall, citizens not approaching the wall without documentation or intending to covertly climb the wall, and a vast number of other intentions, behaviors and actions all together, with the material reality, made the Berlin Wall a border.

Now, further, let us (safely, I trust) assume that there being such a border was an unjust state of affairs. In this case, many ascriptions of *border* to the Berlin Wall not only implied an unjust state of affairs, it also *reinforced* this state of affairs, even if only in imperceptible ways.⁵³ That is, because borders are constitutive social constructions, borders become and remain borders through (among other things) humans collectively ascribing (whether verbally or through intentions or behaviors) the property of *being a border* to them. And so, there is a clear sense in

⁵³ There is an important question here whether the result of collective action (the border) decomposes to individual actions, such that each individual act is indeed part of what creates the border. This question falls squarely within the much larger problem of collective action, to which a variety of solutions have been posed. I believe any of these solutions will do in the case of relating individual acts to constitutive constructions as well. I here gesture to the solution in Kagan (2011). (If you think there is no solution to this problem, this also means that the cause of constitutive social constructions cannot be decomposed to the individual level.)

which these ascriptions had the effect of reinforcing the very same unjust state of affairs that they described or implied.

I think it is possible – and maybe even likely – that there are social contexts in which truths about gender groups are oppressive truths. That is, I think it is possible that true gender ascriptions imply an unethical and constitutively socially constructed state of affairs, such that some accurate gender ascriptions work to reinforce unjust ontological facts concerning gender groups. Why think this? Mostly because I think it is possible if not plausible that gender groups, as we have constructed them, are not sensitive to persons' self-identifications, and so accurate gender ascriptions might force the gender norms discussed in section two upon persons without their consent.⁵⁴

Again, even if gender terms currently pick out persons purely on the basis of biological features that are not socially constructed, our ascribing a gendered status to persons on the basis of their biological features is up to us. Inasmuch as it is up to us, and given the importance of these statuses, we would do well to change our concepts of gender, and subsequently, when and why we deploy gender terms.⁵⁵ Such a conceptual change may result in a reference-shift of gender terms; if so, it is a much-needed shift.

At this point, two clarifications are in order. First, I am not claiming that *all* true gender ascriptions that describe or imply oppressive truths about gender groups (if there are any) reinforce these oppressive states of affairs. We might distinguish between *reinforcing* and *nonreinforcing* ascriptions. Reinforcing ascriptions are ones where group ascriptions that describe or imply an unjust constitutively constructed state of affairs will *reinforce* that state of

⁵⁴ Of course, someone aiming to eliminate gender groups could go for the more radical position that all ontological facts about gender groups are oppressive truths.

⁵⁵ We might compare here our shared concept of *persons*. It is plausible that, given the weighty implications of classifying creatures as persons, we ought to ensure this concepts includes certain non-human animals.

affairs. Nonreinforcing ascriptions are ones where an ascription of this sort will not reinforce that state of affairs. For example, the ascription ‘Blacks are not voters’, when voiced in a legislature for the purposes of justifying not having polling places in black neighborhoods would be a reinforcing ascription. Alternatively, this ascription, when followed immediately by a criticism (e.g., “This is a travesty that must be corrected!”) appears to be a nonreinforcing ascription. Similarly, we can distinguish between contexts such as town halls in North Carolina, where asserting claims such as ‘Trans women are not women’ would *reinforce* trans-exclusive gender groups (and so the state of affairs in which trans women are not ascribed woman) with contexts such as feminist philosophy classrooms, where the same claim might be expressed for the purposes of criticism.

Second, any harm done by true gender ascriptions would not be limited to reinforcement. Ascribing *woman* to a transgender man behind his back might only reinforce an unjust state of affairs in which transgender men are not ascribed *man*. Ascribing *woman* to that person in a different context might reinforce this state of affairs *and* force gender norms on that individual without their consent. (For example, in the context of a college board determining where transgender students will be housed.) And in yet another context, such as making this ascription to the person’s face, one might additionally cause harms of disrespect, rudeness, etc. As mentioned above, some ascriptions may cause none of these three harms. Some might cause any variety of them. For example, even imagining that it *were* true that Mock was “born a man,” I think it is highly plausible that Morgan’s saying so was unethical at least in part because it reinforced an unjust state of affairs *and* was extremely rude. All this is to say that the reasons why we have a duty to not make a particular gender ascription in a given context might be

limited to its reinforcing effect on an unjust social structure, or might (perhaps in addition) concern other, more direct harms.

What does all this mean for the ongoing philosophical debates over the metaphysical facts (if there are any) about *women*? I think it means this: The social group *women*, supposing there is one, might be trans-inclusive. But it might not. And, if it isn't, we should not take this as much more than yet more evidence that society is unjust. Certainly, in most contexts, it should not affect our gender ascriptions. That is, even *if* (e.g.) 'Trans women are not women' were true, it would describe or imply an oppressive truth, and assertions of this claim would reinforce that oppressive truth.

In sum, we can either perpetuate and reinforce or, alternatively, avoid reinforcing unjust gender structures through our gender ascriptions.⁵⁶ And so, given our normative commitments, I think we should adopt gender ascription practices that are dominant in trans-inclusive contexts. Even *if* these ascriptions turn out to be false in other contexts,⁵⁷ they work toward shifting the ontological boundaries around the gender groups -- that is, they push us toward talking about what we *should* be talking about when we talk about genders.

One might be worried that this sounds too much like humoring trans persons. That is, on one interpretation of this claim, it says that we should just pretend that (e.g.) trans women are women to not offend or harm anyone, even though *really* they aren't women. But this mischaracterizes my view for at least two reasons. First, I am not taking a position on the boundaries around gender groups. Rather, my point is that even *if* these categories unjustly exclude (or include)

⁵⁶ In this way, it may further the Haslangerian goal of eliminating the current categories of men and women, as she understands them.

⁵⁷ I assume here that one could be a woman in one context and not in another, just as one could be black in one context and not in other. This phenomenon is sometimes called 'traveling'. For further discussion of traveling, see Mallon (2004).

transgender or genderqueer individuals, this only means that we should work toward changing these categories (which might require changing our use of gender terms).

Second, the purpose behind these ascriptions is not avoiding offense, or even simply avoiding immediate harm, but changing gender structures. Since how we ascribe gender affects the meaning of our gender terms and, in turn, how we treat individuals, by revising when and why we make gender ascriptions, we can affect the referents of these terms as well as how gender functions in the social world. By respecting gender identities in our gender ascriptions, we can revise gender groups be more just. (And, if you think that social structures determine what gender is, revise gender itself in the process.)

Stepping back, I have provided two arguments for thinking that, for those aiming to justify claims about which persons we should acknowledge as (e.g.) women, to focus on the metaphysics of gender is to play the wrong game. The correct metaphysics of *woman* (if there is one) cannot answer what ascriptions we *ought* to make. For this reason, both conservatives and progressives have sidestepped the fundamental dispute about how we ought to *treat* each other – and relatedly, what gender structures we should pursue – with a debate about what (currently) makes someone a man or a woman. Both have come up with metaphysical theories of genders that they take to justify their own positions about how we ought to ascribe gender. But, if I am right, and the facts about who is a man or woman do not constrain the facts about when we should ascribe *man* or *woman*, then both the conservatives and progressives are missing the point. Metaphysical discussions can be set to the side: in most contexts, they mask what matters most by asking the wrong questions.

3.5 Methodological Implications

This brings me to my view's second implication -- that metaphysicians working on gender (and likely similar categories, such as race and disability) should not employ the alethic view. I've argued that the normative facts about just gender ascriptions might come apart from the metaphysical facts about gender categorization (if there are any). And this means that whether a metaphysical view implies that (e.g.) transgender women are women should not be taken to imply that we should not ascribe *woman* to these persons. Certainly, the view should not be dismissed for having this implication.

This is not to say that metaphysics cannot contribute to improving our gender categories. Let's distinguish between two types of metaphysical project: prescriptive and descriptive. The descriptive project seeks to illuminate the features of the gender categories we currently have. This methodology is, I think, nicely demonstrated in Ásta's work. On her conferralist theory of gender, Ásta acknowledges that there will be contexts where "being perceived as being of a certain sex may be an essential grounding property [of gender]," such that transgender persons are a gender other than the gender with which they identify.⁵⁸ The upshot for Ásta is, then, that the features that ground membership in gender categories can vary across contexts, and may be more or less conducive to social wellbeing and justice. In an unjust scenario, she says, one can try to resist their conferred gender, but this is very difficult -- indeed, it may be impossible until there is enough resistance to shift the boundaries around these gender groups. In other words, rather than taking the normative truth that *transgender woman should be acknowledged as women* as evidence for an underlying metaphysical truth about *woman*, Ásta maintains that the

⁵⁸ Ásta (2011), 61-2.

way our social ontology is and the way our social ontology ought to be can come apart, leaving us with the task of finding ameliorative revisions to our ontology.

The prescriptive path takes on the challenge of finding these ameliorative revisions. In particular, a metaphysician pursuing the revisionary path could use philosophical tools to articulate ontological categories or structures -- call them genders* if you wish -- that would make just gender ascriptions *true*. We might consider this metaphysical activism. The basic idea is that considering the ethics of gender ascriptions could provide the foundation for a prescriptive metaphysics -- i.e., a metaphysical theory that details what ontological categories would best serve our political ends, and prescribes using our gender terms to refer to *these* categories. Such theories would rest on the idea that revising our gender ideology, including out gender concepts, will construct (or reinforce) more *just* gender groups, and deconstruct unjust ones. And while this may require changing the semantics of gender terms (or even the metaphysical categories of gender) -- given the immediate importance of our normative goals, this doesn't much matter.

At this point, I should address a concern that this method of analysis is already employed in metaphysical theorizing about social categories. That concern goes something like this:

Haslanger's ameliorative project rests on exactly the point you are making here.

*This project, as she describes it, begins with questions about what concepts we ought to have for a certain social category, and then advocate for particular concepts based on our normative goals. How is your suggestion any different?*⁵⁹

⁵⁹ Thanks to Shamik Dasgupta for raising this objection.

There could be two worries here. The first is that the prescriptive project is what most philosophers have been doing all along in employing Haslanger's 'ameliorative project'. But in section one, I've already offered what I take to be good evidence that this is not a misrepresentation of metaphysicians' methodology. Taking them at face value, these philosophers take themselves to be pursuing projects that metaphysically analyze what gender (or race, etc.) *is*, not what it would be (if anything) in a more just society.

The second worry is that my proposal is not novel, because Haslanger's ameliorative project *just is* the prescriptive project. Here, it is worth separating two interpretations of the ameliorative project and clarifying how they relate to the prescriptive and descriptive projects.

Haslanger's ameliorative project begins by asking "What is the point of having the concept in question; for example, why do we have a concept of knowledge or a concept of belief? What concept (if any) would do the work best?"⁶⁰ It then moves to providing an account of the concept that will best do this work. But there are at least two ways of understanding this project, as it applies to some concept X:

Negative project: Develop a concept of X that best illuminates normatively problematic features of our everyday concept of X.

Positive project: Develop a normatively ideal concept of X.

For example, is the best concept of woman one that best illuminates the negative stereotypes embedded in our everyday concepts of women, or the concept that would operative in a gender

⁶⁰ Haslanger (2012), 367.

just society? Both are reasonable interpretations of the project, but can result in strikingly distinct concepts. Haslanger herself seems to adopt the negative project, and so proposes a concept of woman that highlights gender injustice.⁶¹ Similarly, on the reading of Haslanger advocated by Barnes and Jenkins, the ameliorative project identifies unjust social structures, and theorizes about the construction of that structure. This is a striking contrast to the positive project, which to the contrary, would lead one to develop gender concepts that, if adopted and employed in social interactions in lieu of our current concepts, would promote gender structures that manifest greater gender equality and trans* inclusion. That is, by making these new gender concepts socially operative, we would construct better gender structures (and change which gender ascriptions are true). And it is this positive project that corresponds to what I am calling the prescriptive project.

It is important to recognize that, while both projects are normatively motivated, they differ in their mechanism for achieving normative goals. Negative ameliorative analysis aims to eliminate problematic social categories by *describing* unjust social structures. Positive analysis aims to eliminate problematic social categories by *prescribing* improved social categories. That is, descriptive projects, such as negative projects, seek to *illuminate* what (e.g.) *woman* in fact is; revisionary projects are *positive* ameliorative projects that *prescribe* what *woman* (or something close to it) *ideally should be*.⁶²

Admittedly, this means that the revisionary project might change the subject, rather than doing metaphysical analysis in the traditional sense. But if the normative facts about gender ascriptions are not constrained by the metaphysical facts, and we care about having trans-

⁶¹ Haslanger (2000).

⁶² Haslanger goes to length to avoid the worry that the negative ameliorative project *changes* the subject matter, arguing that semantic externalism gives us good reason to think that -- even if the project's proposed concepts seems significantly different than the concept we began with -- we may be *illuminating* rather than changing the subject. Haslanger (2012), 386-8.

inclusive gender ascriptions come out true, then changing the subject may be exactly what we should be doing. Precising a path to just gender categories is, according to the normative view, at least one useful approach for metaphysicians who are motivated by social justice aims. The flexibility of social ontology affords us a unique place to assist legal and political theorists, feminists, and queer populations by carefully constructing concepts that match their normative commitments.⁶³

⁶³ Many thanks to audiences at the San Francisco Bay Area Feminism and Philosophy Workshop, the University of North Carolina (Asheville), Western Washington University, the Pacific APA (2016), and Wayne State University for feedback during the development of this chapter.

Chapter 4

What is Sexual Orientation?¹

4.1 Introduction

Ordinary discourse is filled with discussions about ‘sexual orientation’. Everyone seems to have opinions about it – whether it should be a legally protected class, whether it is apt for moral judgment, and whether Lady Gaga is right that, whatever our sexual orientations, we were ‘born that way’.²

This discourse suggests a common understanding of what ‘sexual orientation’ is. But even a cursory search turns up vastly differing, conflicting, and sometimes ethically troubling characterizations of sexual orientation. Consider the following, taken from (respectively) a professional scientific association, an LGBTQ advocacy organization, a neuroscientist and a philosopher:

- 1) Sexual orientation refers to an enduring pattern of emotional, romantic, and/or sexual attractions to men, women, or both sexes.³

¹ This chapter was previously published in *Philosopher's Imprint* 16 (3) (2016).

² Lady Gaga, (2011) “Born This Way.” *Born This Way*. Abbey Road Studios.

³ The American Psychological Association (2008).

- 2) "Sexual orientation" is the preferred term used when referring to an individual's physical and/or emotional attraction to the same and/or opposite gender.⁴
- 3) Sexual orientation ... is the trait that predisposes us to experience sexual attraction to people of the same sex as ourselves (homosexual, gay, or lesbian), to persons of the other sex (heterosexual or straight), or to both sexes (bisexual).⁵
- 4) A person's sexual orientation is based on his or her sexual desires and fantasies and the sexual behaviors he or she is disposed to engage in under ideal conditions.⁶

To name just a few of the worries that might be raised for these characterizations: (1)-(3) assume binary categories of sex or gender (i.e., male/female or men/women)⁷; (2) and (3) disagree on whether sexual orientation concerns gender-attraction (attraction to individuals with certain genders) or sex-attraction (attraction to individuals with certain sexes); and (4) appeals to the opaque notion of 'ideal conditions' for acting on one's sexual desires (more on this later).

⁴ Human Rights Campaign (2014).

⁵ LeVay (2011), 1.

⁶ Stein (1999), 45.

⁷ I here understand sex as a classification solely on the basis of human bodies' physical characteristics and gender as a classification (at least in part) on the basis of social situatedness. For more on this distinction, see section 3. Also, I acknowledge that gender-identity (the gender one self-attributes) and gender-expression (the external characteristics and behaviors that are socially interpreted as communicating that one belongs to a certain gender category) can come apart. In this paper, talk of gender-attraction is most easily understood as attraction to certain gender expressions, but I leave open that persons' gender identities can also play a role in gender-attraction.

Characterizations like these – assuming they are attempts to elucidate a shared, pre-existing concept of sexual orientation – reveal that we have an extremely poor grasp of this concept. And even if the characterizations are stipulative, we have good reason to resist adopting many of them. Inadequate understandings of sexual orientation can reinforce heteronormative assumptions (i.e., assumptions that heterosexuality should be privileged within society) by maintaining a majority/minority divide between heterosexuality and other sexual orientations that historically has been normatively loaded and policed. They also can reinforce cisnormative assumptions (i.e., assumptions that all persons are cisgender – that is, that all persons’ genders are the ones assigned to them at birth on the basis of their anatomy) by failing to provide recognition or clarity within the sexual orientation taxonomy for persons who are not cisgender or who are attracted to persons who are not cisgender.^{8,9}

The conceptual jumble surrounding sexual orientation suggests that the topic is overripe for analytical philosophical exploration.¹⁰ While this delay may be due to metaphysicians’ historical focus on discovering and articulating metaphysically necessary truths about reality, recent feminist critiques have brought topics related to contingent social realities into the subfield’s focus. Philosophers such as Charlotte Witt and Sally Haslanger have begun, for example, rich and growing literatures on metaphysical questions concerning gender and race. But sexual orientation has yet to receive due in-depth metaphysical exploration.

This paper lays the groundwork for one such in-depth exploration, and in so doing, encourages further analytic philosophical discussion of sexual orientation. Its target is twofold:

⁸ Often this cisnormative assumption is paired with the views that gender is biologically determined by one’s anatomy, and that gender is essentially a biological, rather than social category.

⁹ Throughout this paper, I will use ‘their’ as a gender-neutral singular pronoun. Some people may take grammatical issue with this. If you are one of those people, I encourage you with all good will to incorporate a gender-neutral singular pronoun of your own choosing into ordinary English discourse.

¹⁰ While sexual orientation has received little attention in the analytical tradition, the continental tradition has a rich history of thinking about sexual orientation – see, Foucault (1980) and Halperin (1990) & (2002), among others.

(i) the everyday concept of sexual orientation, and (ii) the corresponding concepts associated with the taxonomy of sexual orientation (e.g., gay, straight). These concepts are highly interwoven, since the concept of sexual orientation constrains the taxonomy. (For example, a concept of sexual orientation that centrally concerns a relation between a subject's own sex [or gender] and the sex [or gender] of the persons they are attracted to will imply a taxonomy containing correspondingly relational concepts.) My project sets out to engineer a revised concept of sexual orientation that implies a new taxonomical schema of sexual orientation. Both the revised concept and new taxonomical schema are intended to *elucidate* and *improve* our everyday concepts in light of particular theoretical and socio-political purposes. And importantly, this project is limited in scope: the proposed concept and taxonomy of sexual orientation are not meant to apply across any cultural context. Instead, my project constructs a concept that is both responsive to and critical of our everyday thinking in contemporary western society about sexual orientation.

On my proposed account of sexual orientation, which I call 'Bidimensional Dispositionalism', sexual orientation is based upon a person's sexual behavioral ^{dispositions} under the ordinary manifesting conditions for these dispositions (i.e., the conditions corresponding to applications of the term 'sexual orientation' and related terms), and having a particular sexual orientation is based upon what ^{sex[es]} and ^{gender[s]} of persons one is (or is not) disposed to sexual engage with under these conditions. Importantly, these particular categories of sexual orientation do *not* reference one's own sex or gender.¹¹

In what follows, I assume non-eliminativism about sex and gender. I use the terms 'male' and 'female' to refer to sex categories, though I do not assume that these terms exhaust or refer

¹¹ Acknowledging that some people wholly lack dispositions to sexually engage with other persons on the basis of sex- or gender-attractions will include asexuality with regard to sex and gender among the class of sexual orientations.

to discrete sex categories. Similarly, I use the terms ‘man’ and ‘woman’ to refer to gender categories, though I do not assume that these terms exhaust or refer to discrete gender categories.

4.2 Methodology and Framework

This section clarifies my project’s methodology and framework. I first discuss the project’s methodology, and then turn to the purposes guiding my analysis of the concept of sexual orientation. Finally, I distinguish between the central target of my analysis – the everyday concept of sexual orientation – and three distinct but closely related concepts.

4.2.1 The Engineering Project

In her work on gender, Haslanger points out the importance of distinguishing between three projects that ask a question of the form, What is x ? One project is conceptual: it asks only about the content of our ordinary concept of x . Another is naturalistic: it asks which natural kind (if any) our ordinary concept of x tracks. The last project, and the one that best categorizes the methodology of this paper, is what I will call the ‘engineering project’: it asks about the purposes of our concept of x , and (if necessary) improves or replaces the existing concept to better realize the purposes we want this concept to fulfill.^{12, 13} This project takes seriously that, as Alexis Burgess and David Plunkett put it, “our conceptual repertoire determines not only what we can

¹² Haslanger (2000), 32-33. Haslanger calls this the ‘analytic project’. In order to distinguish it from ‘conceptual analysis’, though, which is more akin to the ‘conceptual project’, I will refer to it throughout as the ‘engineering project’.

¹³ Though I do not want to take a strong stance on the nature of concepts, I am loosely understanding concepts here as ways of representing the world. I mean this, though, in a deflationary sense that remains neutral on the issue of whether concepts can carry non-descriptive, expressive content.

think and say but also, as a result, what we can do and who we can be.”¹⁴ Given this important feature of our conceptual repertoire, we can think of the engineering project as one that sets out to elucidate and possibly revise or replace our everyday concepts in light of the impact we would like them to have.

Importantly, the engineering project is not *required* to replace or even to revise an existing concept – what is important is that the final concept serves the proposed purposes. In some cases, these purposes may be best served by preserving (e.g.) the ordinary usage, connotation, or extension of the everyday concept in question. To quote Haslanger:

[I]f we allow that everyday vocabularies serve both cognitive and practical purposes, purposes that might also be served by our theorizing, then a theory offering an improved understanding of our (legitimate) purposes and/or improved conceptual resources for the tasks at hand might reasonably represent itself as providing a (possible revisionary) account of the everyday concepts. ... The responsibility is ours to define [these concepts] for our purposes. In doing so we will want to be responsive to some aspects of ordinary usage (and to aspects of both the connotation and extension of the terms).¹⁵

In this way, engineering projects may range in the descriptiveness and prescriptiveness of their conceptual construction. Given the purposes that projects assign to their target concepts, they might prescribe no revision to an everyday concept, or prescribe revisions that (among other things) preserve features of the everyday concept. These latter projects are descriptive insofar as

¹⁴ Burgess & Plunkett (2013), 1091.

¹⁵ Haslanger (2000), 33.

they elucidate and maintain certain features of the everyday concept, but prescriptive insofar as they propose revisions to the everyday concept in light of certain purposes.

Haslanger acknowledges that, because of this flexibility, an engineering project will confront issues regarding how conceptually conservative it intends to be – i.e., whether it intends to retain, revise, or eliminate the everyday concept. My project is somewhat conservative in one sense but not another. While – for reasons I will soon explain – it attempts to preserve the general extension of our everyday concept of sexual orientation, it does not attempt to preserve many of the connotations associated with the term ‘sexual orientation’ or the concepts associated with our current taxonomy of sexual orientation. Given this, my project will clarify and minimally revise our everyday concept of sexual orientation, but also eliminate and replace the everyday concepts associated with the taxonomy of sexual orientation.

Some might take my proposed concept of sexual orientation to constitute a new, distinct concept replacing the former concept, rather than revising it. I can see both sides of this issue. Insofar as my account attempts to be responsive to our ordinary usage of the concept of sexual orientation while revising the concept’s content, it may be thought merely revisionary. But insofar as this revised content implies a new taxonomy of sexual orientation and significantly changes the connotations of the term ‘sexual orientation’, it may be thought to entirely replace our former concept. This tension is fine; I’m not sure anything important hangs on whether my project is described as providing a revised or replacement concept of sexual orientation. Either description can point to what *does* matter – that as an engineering project, my project consists of two parts:

- (I) Elucidating purposes ideally served by our concept of sexual orientation.

- (II) Re-engineering our concept of sexual orientation (and the corresponding taxonomy of sexual orientation) in light of the purposes described in (I).¹⁶

Having now described the methodology of my project, I will turn to describing purposes that (I argue) are ideally served by our concept of sexual orientation and clarifying my central target concept. I will then spend the remainder of the paper re-constructing this concept and its corresponding taxonomic schema in a way that fulfills these purposes.

4.2.2 Framework I: Purposes

I propose that the following purposes are ideally served by our concept of sexual orientation:

- (i) Clarifies the criteria for ascribing sexual orientation, as well as how these criteria translate into a taxonomic schema of sexual orientation;
- (ii) Is consistent with relevant social scientific research – in particular, research concerning sex and gender;

¹⁶ These parts could be conceptually divided into two projects, one of which looks for suitable concepts in light of assigned purposes, and the other of which engineers revised (or new) concepts that meet these purposes, (should suitable ones not be found elsewhere). For simplicity, I am including both projects under the heading of the ‘engineering project’.

- (iii) Reduces or eliminates the presumption that cisheterosexuality¹⁷ is the normatively standard sexual orientation and all queer sexual orientations are normatively deviant;¹⁸ and
- (iv) Is conducive for establishing legal and social protections for persons who have queer sexual orientations.

These purposes are not merely stipulative; someone could disagree with me concerning whether these purposes should guide our concept of sexual orientation. I take each of them, though, to be rooted in everyday political and social realities.

My reasons for adopting (i)-(ii) are both theoretical and practical. As I've shown, sexual orientation is understood in a variety of conflicting ways – there is disagreement about how to articulate the criteria for ascribing sexual orientation (e.g., in terms of gender- or sex-attraction), as well as corresponding disagreement about the taxonomy of sexual orientation. There also are regular confusions between sex and gender, which suggests that these understandings are not informed by recent research concerning the distinction between sex and gender. This alone immediately reveals a need for an elucidation of the concept and taxonomy of sexual orientation, and possibly a revision ensuring their consistency with relevant research on sex and gender.

More practically, clarifying the criteria for ascribing sexual orientation (and how they translate into a taxonomic schema of sexual orientation) is a key ingredient in developing a

¹⁷ As will become clear in the subsequent section, because I understand sexual orientation as concerning both sex and gender, I reject the idea that heterosexuality picks out a specific sexual orientation. I believe that talk about 'heterosexuality' in ordinary discourse is usually talking about 'cisheterosexuality' – that is, the attraction of a cisgender woman to a cisgender man or vice versa.

¹⁸ I use the term 'queer' here to mean something like 'not cisheterosexual'. For reasons that hopefully become clear, I intentionally avoid terms such as 'same-sex', 'homosexual', etc.

concept that serves the social and political purposes stated in (iii)-(iv). Confusions between sex and gender – especially with regard to sexual orientation – regularly create difficulties for queer, gender non-conforming, and intersex persons, as well as their partners. How should a gender non-conforming, transgender, or intersex person (or their partners) describe their sexual orientations? How can or should non-discrimination laws address these sexual orientations? The current categories of sexual orientation offer little to no flexibility or clarity for these individuals. For these reasons, the current categories reinforce cisnormativity as well as heteronormativity. That is, because the current categories place queer orientations in a vast minority and have no place at all for many transgender or intersex individuals (or persons attracted to these individuals), they perpetuate prejudices that sexual orientations and gender identities that do not meet standard binaries of homosexual/heterosexual and cisgender man/cisgender woman are somehow deviant, dysfunctional, or even nonexistent.¹⁹ Ideally, our concept of sexual orientation would rid or at least diminish these harms by achieving purposes (iii)-(iv) above, and do to so (at least in part) by employing the tools articulated in purposes (i)-(ii).

In addition to guiding my project, purposes (iii)-(iv) also impose certain constraints. In particular, they constrain the project to construct a concept of sexual orientation that is feasible for public uptake. Only such a concept can move us toward achieving these social and political purposes. So, rather than rebuilding the concept of sexual orientation from scratch, I restrict myself to engineer a concept that clarifies and improves upon the *pre-existing* structure of our everyday concept and – on the basis of this clarification and improvement – rebuilds and expands the sexual orientation taxonomy.

¹⁹ Consider, for example, the well-recognized phenomenon known as ‘bisexual erasure’ (a tendency to explain away or simply deny evidence that persons are attracted to both men and women, or on alternative accounts, females and males). (See, e.g., Greenesmith (2010).) See also Stein’s (1999) critiques of the binary operationalization of sexual orientation in scientific studies.

To put this slightly differently, I am not out to develop a theoretically ideal or purely stipulative concept and then argue that the term ‘sexual orientation’ should be attached to my concept rather than the everyday concept. Instead, I am constraining my conceptual engineering so that it is responsive to our ordinary usage by generally preserving the extension of our everyday concept of sexual orientation.²⁰ I will say more about this soon, but in particular, this means that I limit my concept to one that primarily concerns sex-attraction and gender-attraction. I have no qualms if someone wants to describe this restricted project as building a ladder that we climb in order to eventually kick away and move on to a different concept of sexual orientation. It is a ladder that I think we must climb. And – I would suggest – this pragmatic approach to a conceptual project is well suited for any project that hopes to balance theoretical aims with a political and social agenda.²¹

4.2.3 Framework II: Target Concept

Given my project’s constrained scope, it is important to get a sense of the everyday concept’s extension. To this end, I will now argue that we should distinguish the everyday concept of sexual orientation from three other, closely related concepts.

The first is *sexual identity*, which I understand to refer to an individual’s self-identification with regard to sexual orientation. Because sexual identity concerns sexual orientation in this way, the concept of sexual identity is sensitive to the concept of sexual orientation. But we also

²⁰ I say ‘generally’ because it is unclear to me whether our everyday concept of sexual orientation extends to certain non-cisheteronormative pairings and simply fails to place them within its taxonomy, or whether it fails to extend to these pairings. My project secures this extension. It is also worth noting that this constraint is not an unusual move in metaphysics – e.g., Haslanger (2000) and Sider (2011) also constrain their conceptual engineering to a particular phenomenon in light of certain proposed purposes. It is also a move that is explicitly discussed in recent literature on conceptual ethics, such as Plunkett (2015) and Burgess & Plunkett (2013).

²¹ Thanks to an anonymous referee for pressing me to clarify this point.

acknowledge that someone can be self-deceived or in denial about their sexual orientation (or even lack the concepts necessary for self-identification), while still being truly said to have the sexual orientation that they fail to recognize. Given this, I will not address sexual identity in what follows.

The second is *romantic or emotional attraction*. Some characterizations of sexual orientation – for example, that of the American Psychological Association – understand sexual orientation in terms of “emotional, romantic, and/or sexual attractions.”²² I grant that romantic and emotional attractions are often reliable evidence of sexual orientation, and can be themselves the target of discrimination. But it seems that our concept of sexual orientation is distinct from the concepts of romantic and emotional attraction in that it primarily concerns sexual behavior. This is why, for example, I think we correctly call ‘asexual’ persons who are disposed to never engage in sexual behaviors, even though they may experience a range of romantic and emotional attractions. The fact that asexuality is considered (even by asexual persons) to be a single category within the taxonomy of sexual orientation, despite asexuals reporting a wide range of romantic and emotional attractions, suggests that these latter attractions are captured by concepts unique from that of sexual orientation.²³ (There is even a distinct taxonomy for these romantic and emotional attractions, e.g., ‘biromantic’, ‘panromantic’.)

Conversely, we can imagine that someone – or even every person – who has sexual attractions that lead us to ascribe a sexual orientation to them could completely lack romantic or emotional attractions. In short, it is not difficult to think of examples in which persons with seemingly the same sexual orientation have vastly differing romantic or emotional attractions, as

²² American Psychological Association (2008).

²³ See Emens (2014).

well as examples in which sexual orientation is unaccompanied by romantic or emotional attractions.

Given cases like these, I assume in what follows that, while romantic and emotional attractions might fall under a concept of sexuality broadly construed, the concepts associated with these attractions are distinct from the concept of sexual orientation. For this reason, my project is not directly concerned with emotional or romantic attraction. That is, my analysis is not concerned with emotional or romantic attractions that have no effect upon one's dispositions toward sexual behavior, and only indirectly concerned with those that do. Should, for example, someone's romantic attractions significantly influence these dispositions, their romantic attractions will be part of what forms their sexual orientation under my account insofar as they have this influence. Any concern with attraction in the sequel will focus upon sexual (and I mean *sexual*!) attraction regardless of whether other forms of attraction accompany it. To this end, talk of attraction in what follows generally can be understood as shorthand for dispositions to engage in sexual behaviors.²⁴

The third concept to distinguish from sexual orientation is what I call *sexual druthers*, which refers to specific preferences of sexual partners within potential partners according to one's sexual orientation. This is often referred to as someone's 'type'.²⁵ Height, hair color, body structure and voice quality are all examples of traits about which people may have sexual druthers. In order to generally preserve the extension of our everyday concept of sexual

²⁴ Michael Rea raises the interesting question of what this distinction [between emotional/romantic attraction and sexual orientation] implies for someone who lacks dispositions to engage in sexual behaviors (perhaps, e.g., due to chronic deficiency of sex hormones), but who has higher-order desire for sexual intimacy. Does having only this higher-order desire preclude such persons from having a sexual orientation? I would answer 'no' — not so long as we consider asexuality a sexual orientation. Asexuality is generally understood as the lack of sexual attraction, or lack of first-order desire to have sexual contact with someone else. [See, e.g., The Asexual Visibility & Education Network (2012).] Asexuals can and often do experience romantic or emotional attractions, though. And they might have higher-order desire to experience first-order sexual desire or sexual intimacy. A person in the situation that Rea describes seems, for these reasons, to be best categorized as asexual.

²⁵ Of course, a single person may have multiple 'types'.

orientation, I do not include sexual druthers in my account of sexual orientation, and instead focus upon preferences of sexual partners with regard to sex and gender categories.

Admittedly, the cultural distinction we make between sexual orientation and sexual druthers seems somewhat arbitrary. It is not clear why attraction to certain sexes or genders is considered relevant to one's sexual orientation, but not attraction to a certain hair color, race or economic status. But sex and gender are, for better or worse, particularly salient social categories with respect to sexual orientation. As a result, we find ourselves in the position of classifying persons' sexual orientations on the basis of their sex- and gender-attractions, and not on the basis of other sexual attractions. And this makes persons with particular sex- and gender-attractions more vulnerable to discrimination than persons with attractions to persons with a certain hair color or economic status.

Again, I am not generally preserving the everyday concept's extension for its own sake, but in order to fulfill certain purposes. And the purposes that I've proposed are, I think, best served by maintaining the distinction between sexual druthers and sexual orientation (understood in terms of attraction to persons with certain sexes or genders). No one is interested in creating nondiscrimination laws to protect people attracted to blondes or baritones. We are, though, interested in creating legal and social protections for queer, transgender, gender non-conforming, and intersex persons. And, as mentioned before, I am here assuming that an account of sexual orientation should be aimed at better realizing these political and social purposes. So, given the pragmatic interests guiding the boundaries of 'sexual orientation', I think that I can best fulfill purposes (iii)-(iv) by retaining these boundaries and separating sexual orientation from sexual druthers. For this reason, I will hold fixed that sexual orientation primarily concerns sex- and gender-attractions, and not other sexual attractions.

That said, one might worry that even once we take on board the distinction between sexual orientation and sexual druthers, it remains vague because many of the traits that are objects of druthers also (at least in part) construct gender. In other words, the worry goes, if sexual orientation concerns attraction to persons of a certain gender, and gender is a social construction that concerns (e.g.) performativity (behaviorisms, dress, etc.) or social status, then the distinction between sexual orientation and sexual druthers is vague. I agree with this point, but still insist that there are cases where sexual orientation and sexual druthers come apart – that is, cases where someone is attracted to a particular feature that lacks gendered connotations.²⁶

In fact, it may be that some persons lack any gender- or sex-attractions, and are *solely* sexually attracted to persons with non-gendered features such as wealth or red hair. That is, there may be persons whose sexual attractions are only based on what I've termed sexual druthers.²⁷ This, of course, puts pressure on the distinction between sexual orientation and sexual druthers – why don't we think that being (e.g.) solely attracted to redheads is a sexual orientation?²⁸ And if it is not, what sexual orientation do such persons have, since they are not asexual (i.e., having no sexual attractions to anyone), but experience no sex- or gender-attractions? These questions press

²⁶ Whether or not certain druthers have gendered connotations should be assessed from a subjective perspective. It could be, for example, that someone has a druther (or, if you prefer, fetish) for a particular shoe color because *for that person* this shoe color is associated with aspects of sexual engagement that express their sexual orientation. While such druthers might appear to others to have no gendered connotations, they have these connotations for the persons who have them. For this reason, I think that an account of sexual orientation should concern such gender-laden druthers, albeit indirectly, and as expressions of persons' underlying sexual orientations.

²⁷ Thanks to David Black for bringing this possibility to my attention.

²⁸ A more common example of a similar phenomenon is pedophilia, or exclusive attraction to pre-pubescent children regardless of their sex or gender. Is pedophilia a sexual orientation? Those inclined to think that it is *not* a sexual orientation might think that it is instead a sexual disorder, which (unlike sexual orientation) is something that should be subject to psychiatric and medical treatment. Those inclined to think that it *is* a sexual orientation, though, can maintain that sexual orientation concerns sex- and gender-attractions. On a plausible view of sex-categories, these categories are based on continuous and fluid groupings of primary and secondary sex characteristics. Because these characteristics undergo significant change during puberty, it is arguable that humans shift sex-categories during puberty. These possibilities are surely not exhaustive, and I acknowledge that this is a difficult (as well as a morally and emotionally loaded) case. Undeniably, there are a host of similar difficult cases that beg for further discussion. Length limitations and the freshness of this topic to philosophical debate constrain me from providing more than an extremely general discussion of sexual orientation in this paper.

on our ordinary concept, and its embedded assumption that there is a clear line between sexual druthers and sexual orientation. Happily, I take no such position. Borderline cases like these make it clear that our social categories have fuzzy edges and do not cut at deep joints. But we would be mistaken to expect that they would.

For my purposes, what is important is that – though the separating line can be fuzzy – the majority of cases show that there clearly *is* a distinction between sexual orientation and sexual druthers. I leave it to future papers to further discuss borderline cases, and whether these cases ought to affect the extension of our concept of sexual orientation; the sequel will aim toward a general account of sexual orientation that preserves the central bases of sexual orientation as it is ordinarily understood – namely, sex- and gender-attraction. For this reason, I will assume that – just as sexual orientation concerns romantic and emotional attractions *only insofar* as they affect sexual attraction – sexual orientation concerns sexual druthers *only insofar* as features that are the object of druthers go into constructions of gender (or insofar as primary or secondary sex characteristics that are objects of sexual druthers affect sexual behavioral dispositions).

4.3 Bidimensional Dispositionalism

In what follows, I propose a concept of sexual orientation that is designed to satisfy purposes (i)-(iv). However, I first address two issues that constrain and shape my concept of sexual orientation: the distinctions between sex and gender and between behaviorism and

dispositionalism. I then state my proposal and discuss its implications, as well as additional philosophical questions pointing to further expansion on my proposal.²⁹

4.3.1 Preliminary Issues

Sex and Gender

As mentioned, previous characterizations of sexual orientation typically and without argument equate and assume binary categories of sex and gender. They often also analyze sexual orientation in terms of either sex-attraction to the exclusion of gender-attraction, or vice versa. These assumptions lead to understandings of sexual orientation according to which sexual orientation is unidimensional – tracking either sex- or gender-attraction, but never both independently of each other – and limited to a small number of discrete sub-categories.

The position that there is no distinction between sex and gender might be understood in two ways: as the claim that sex (taken as anatomical) wholly determines gender, or that gender (taken as socio-political or psychological) wholly determines sex. The former – call it the ‘cisnormative view’ – is much more prevalent than the latter – call it the ‘unification view’ – which has a small presence within feminist theory and other academic literature.³⁰ In this paper, I am primarily concerned with rejecting the cisnormative view. While I find the unification view provocative, I

²⁹ Specifically, I will address the application of discrete vs. continuous categories of sex or gender and essentialism vs. constructionism to accounts of sexual orientation.

³⁰ See, e.g. Butler (1990), Wittig (1992) or Halpern (2000). Butler and Wittig (both feminist theorists) argue that sex classifications follow labor- and politically-driven gender classifications. Wittig, for example, writes, “No biological, psychological, or economic fate determines the figure that the human female presents in society: it is civilization as a whole that produces this creature.” (For dissenting discussion of these views, see Alcoff (2005).) In contrast, Halpern (a psychologist) emphasizes sex differences in cognitive capacities, meaning that one who understands gender as a primarily psychological feature might argue based on Halpern’s research that sex differences follow gender identity.

also think it incorrect. Much of the motivation for the view, it seems, comes from the conviction that we should not posit sex/gender along a nature/culture binary because this binary has historically provided justification for women's oppression or because – as Linda Alcoff rightly identifies – “in an important sense, everything is natural.”³¹ But (as Alcoff also points out) it is confused to think that understanding sex as an anatomical category entails understanding it as purely ‘natural’ or culture-independent. Our sex categories seem to (in some way) track anatomical features associated with reproductive functions, but these very anatomical categories of reproductive roles could be (at least partially) the result of cultural practice. As I see it, one can maintain that both sex and gender categories are (at least partially) culturally constructed while distinguishing between them on the grounds that they are constructed upon different physical and social features and aimed at fulfilling different purposes. Given this, I see little to be gained by adopting the unification view.

Of course much more can be said, but for the sake of brevity I will now turn to the cisnormative view. A pathway to challenging this view was famously laid by Simone de Beauvoir, who marked sex as a biological category and gender as a category concerning the social position (e.g. exploitation and oppression) experienced by those exemplifying femininity.³² While the details of de Beauvoir's ideas have been challenged in contemporary discussions, her sex/gender distinction is the standard view in psychology, sociology, and queer and women's studies, as well as in feminist philosophy.³³

³¹ Alcoff (2005).

³² Åsta, (2011), 48. As de Beauvoir famously said, “One is not born, but rather becomes, a woman.” I will adopt Haslanger's view that ‘sex’ refers to a classification on the basis of anatomy, (e.g. primary sex characteristics), though this is compatible with these classifications having vague boundaries, or boundaries heavily influenced by socio-political interests. I am also sympathetic with Haslanger's view that gender is not merely a social construct, but is defined in terms of social relations. [Haslanger (2012), 39.]

³³ For an overview of recent philosophical approaches to sex and gender, as well as the interaction between philosophical approaches and other (e.g. psychology, political) approaches to sex and gender, see Mikkola (2012) and Haslanger (2012).

Given this, it would be fairly uncontroversial for me to simply assume this distinction moving forward in my account of sexual orientation. It is worth saying explicitly, though, that not only is the distinction theoretically useful, dividing what seem to be distinct phenomena, but it is also politically and socially advantageous. For one, it provides a helpful framework through which to understand the gender identity or anatomical transition of (e.g.) gender nonconforming, androgynous and transgender individuals. That is, because it separates sex as an anatomical category from gender as a category of social situatedness, it creates the possibility for understanding how the two can be combined in a variety of ways. It also creates an avenue for addressing the ways in which gender categories can be altered to combat patriarchal social structures. If, for example, gender is defined in terms of social situatedness such that (as Haslanger argues) to be a woman is (in part) to be in a position of systematic social oppression, then acknowledging this clarifies the changes that should and can be made to our gender categories if we are to establish gender equality.³⁴

For all of (but certainly not only) these reasons, I hold that the following constraint on an account of sexual orientation will move us closer to achieving what were earlier established as the purposes ideally served by a concept of sexual orientation:

- a)** The account must be compatible with the distinction between sex and gender.

If the cisnormative assumption and unification view of sex and gender are dismissed – and I think they should be – then the unidimensional view of sexual orientation also should be. The

³⁴ See, e.g. Haslanger (2000).

distinction between sex and gender allows for various combinations of sex and gender across individuals, making it clear that an account of sexual orientation should be sensitive to the fact that individuals may be sexually attracted to persons with various sex/gender combinations.³⁵ For example, someone may be attracted only to transgender men who have not had genital or top surgery, or only to cisgender men and women. And given that we recognize that conferrals of sexual orientation tracks both gender- and sex-attraction, we also should recognize that it tracks various combinations of these attractions. For this reason, I place a further constraint on my account of sexual orientation:

- b) The account must permit individuals' sexual orientations to be based on both gender-attraction and sex-attraction.

Someone perfectly happy with (a) may still resist (b), and argue that sex-attraction (or gender-attraction) should be taken as a mere sexual druther, allowing sexual orientation to remain unidimensional. Consider, for example, someone who is attracted to women and not men, but is only attracted to cisgender women. Why think that this latter attraction is anything more than a sexual druther? That is, why should we think that someone attracted only to cisgender women has a different *sexual orientation* than someone attracted to both transgender and cisgender women? This line of argument might suggest that, while we preserve the extension of our ordinary sexual orientation concept, we should make the relevant criteria for ascribing sexual

³⁵ I take it to be a fairly uncontroversial assumption that we *can* (though an individual need not) experience sexual attraction to purely anatomical features as well as gendered features. (I acknowledge that the boundary between these features is slippery.) Given this, we can already begin to see how sexual orientation is significantly dependent upon both biological traits and particular social contexts. For example, if we hold that one's sexual orientation concerns (at least in part) gender-attraction, and that gender is merely a social construct, this sexual orientation will be dependent on placement in a context that has gender-constructs.

orientation only gender-attraction (or only sex-attraction), and also categorize sexual orientations along only one of these dimensions.

There are a number of considerations against thinking of sex-attractions (or gender-attractions) as mere sexual druthers. The two most weighty (and related) considerations are: (i) the frequency with which people experience sexual attraction to not only to individuals with particular gendered-features, but also to individuals with particular primary and secondary sex characteristics, and (ii) the need to recognize the community of persons who are exclusively (or strongly) attracted to transgender individuals, or who are themselves transgender and seeking persons with these attractions.

The first consideration is fairly straightforward: it is simply the observation that, in addition to attraction to particular genders, persons can also be exclusively interested in partners with particular sex characteristics. It is not unusual, for example, for persons to become disinterested in pursuing a relationship with someone upon discovering that they are transgender or intersex.

At first, one might be tempted to chalk up all scenarios like this to transphobia or other prejudices, and not these persons' sexual orientations. However, there also are numerous cases within the growing 'trans-oriented' community of persons who experience strong or exclusive sexual attraction to transgender persons.³⁶ These persons report feeling misplaced among the current categories of sexual orientation, identifying neither as straight nor as gay. Some, for example, consider themselves a 'different kind of gay' – indicating that the current taxonomy of sexual orientation simply fails to capture their sexual orientation, since they experience strong or exclusive attraction specifically to persons who are not cisgender. This failure is also reflected

³⁶ See www.transoriented.com or the most recent work of British journalist and transgender rights activist Paris Lees on the question, 'Is trans-oriented an emerging sexual orientation?'

within academic literature, where a variety of terms have been suggested for these individuals (e.g., MSTW [men sexually interested in transwomen], gynemimetophilia/andromimetophilia).³⁷

While – unsurprisingly – none of these terms have caught on, the community of trans-oriented persons (and the research concerning this community) suggests that individuals can have exclusive sexual preference for transgender persons that is not caused by social prejudice. Similarly, one would expect, individuals can have unprejudiced exclusive preference for cisgender persons. Recognition of these possibilities is not only important for trans-oriented persons; it is also important for transgender persons who may experience rejection by both ‘straight’ and ‘gay’ potential partners, and who are seeking someone who is (though perhaps are not exclusively) trans-oriented. It is also important for ensuring legal and social protections for persons with these attractions, insofar as we can reasonably expect that (e.g.) many partners of transgender persons will encounter discrimination that they would not if partnered with cisgender persons. Without concepts that capture these attractions as part of individuals’ sexual orientation, it becomes difficult if not impossible to guarantee the protections that such individuals deserve. For all of these reasons, I maintain (ii), and disagree with the position that either sex- or gender-attraction should be classified as mere sexual druthers.

Of course, adopting (a) and (b) does not resolve the issue of whether sex or gender (and therefore sexual orientation) should be understood in terms of discrete or continuous categories. I revisit this issue in discussing my account’s implications. But it is worth noting here that this neutrality is, I think, appropriate for a general account of sexual orientation. Understanding sexual orientation categories as discrete or continuous should piggyback on, and not decide,

³⁷ See, e.g., Weinberg & Williams (2010); Money & Lamacz (1984).

understanding sex and gender categories as discrete or continuous. And the debate over this issue has not reached a clear consensus.

For similar reasons, my project does not take a precise stance on which features are the basis of sex and gender categories. As seen from my discussion so far, I do assume that sex and gender are real (i.e., non-eliminativism), that sex and gender are distinct, and that sex categories are related to anatomical features while gender categories are related to relational and social features. (Of course, there may be overlap in the features that provide the basis for sex and gender ascriptions – what’s important is that they are not identical.) More specific theories of sex and gender can be filled into the forthcoming schematic understanding of sexual orientation (and its taxonomy). I purposively build this flexibility into my account in order to construct a concept of sexual orientation (and of its taxonomy) that can be structurally preserved even when the number or understanding of recognized sex and gender categories undergoes shift.³⁸

Conditions (a) and (b) are also intended to be neutral with regard to whether we can in the future adopt *further* dimensions of sexual orientation, and subsequently expand our concept of sexual orientation. As discussed previously, my current project is limited to these two dimensions because it aims to construct a readily accessible, but politically and socially beneficial concept of sexual orientation.

Behaviorism and Dispositionalism

The previous subsection argued that we should take sexual orientation to involve both sex- and gender-attraction. But it is not clear how to assess these attractions in order to determine

³⁸ Thanks to an anonymous referee for pressing me to clarify this feature of my project.

someone's sexual orientation. The task of clarifying the criteria for ascribing sexual orientation and how these criteria translate into a taxonomic schema of sexual orientation, then, is not complete.

The following part of my project continues this task. It is primarily descriptive in nature, though it will also contain a prescriptive element. It is centrally aimed at further elucidating criteria for ascribing sexual orientation as well as articulating (in light of social and political motivations) criteria for placing individuals within a taxonomy of sexual orientation.

The following discussion will compare two main approaches to this task: *behaviorism* and *ideal dispositionalism*. Both of these approaches, I will argue, fail to provide an acceptable analysis of sexual orientation because both insist on overly rigid conditions for ascribing sexual orientation – behaviorism insists on rigid actual conditions, and ideal dispositionalism insists on rigid ideal conditions. I conclude by demonstrating how a different form of dispositionalism – call it *ordinary dispositionalism* – captures an intuitive balance between actual and ideal conditions for ascribing sexual orientation.

Behaviorism

One way of understanding sexual orientation is as nothing over and above (i.e. reducible to) one's *observable behaviors* – that is, as something solely concerning behavior and not at all concerning psychological states, except perhaps states that can be in turn reduced to behavior.

We can understand this view – behaviorism – as the following claim:

A person's sexual orientation is determined solely by their observable sexual behavior.³⁹

In other words, under a behaviorist account, an individual's sexual orientation is decided simply by looking at their sexual behaviors, and seeing what sex[es] and gender[s] of persons they sexual engage with. For example, if they only sexually engage with cisgender men, their sexual orientation is ascribed accordingly.

An immediate difficulty for behaviorism is determining what behaviors and span of time are relevant to someone's sexual orientation. Even setting this aside, though, three more egregious problems remain.⁴⁰ First, behaviorism doesn't allow that individuals can behaviorally repress their sexual orientation. Consider, for example, the case of Episcopal Bishop Gene Robinson, a cisgender man who, after privately identifying himself as gay during seminary, was married and faithful to a cisgender woman for fifteen years.⁴¹ Cases like these are extremely common within the LGBTQ community – under extreme social pressure to conform to cisheteronormativity, many individuals enter so-called 'straight' relationships and so behaviorally (if not also psychologically) repress their sexual desires. Additionally, homeless LGBTQ persons are often forced into prostitution, thereby sexually engaging with individuals of sexes and genders that these persons do not find sexually desirable.⁴²

But by behaviorist lights, it is correct to categorize these individuals' sexual orientations according to their coerced behaviors, rather than according to their (freely or forcibly) behaviorally repressed desires. It seems obvious to me that this is a bad result. Because sexual

³⁹ Alternatively, Edward Stein describes this as the view that "a person's sexual orientation is indexed to his or her sexual behavior." [Stein (1999), 42.]

⁴⁰ For further discussion of the merits and demerits of behaviorism, see Stein (1999).

⁴¹ Robinson (2012).

⁴² Ray (2006).

behavior can be – and for LGBTQ persons, frequently is – coerced by societal pressures, we must understand sexual orientation as something ‘deeper’ than observable behavior. Even if influenced by social pressures, sexual orientation cannot be explicitly forced upon someone by these pressures.⁴³ To deny this is to do an injustice to a large number of LGBTQ persons, and especially in countries where queer sexual behavior can result in prison or even death.

Two other, related problems for behaviorism regard its implications for voluntary celibates and persons who are not sexually active, as well as sexually active persons in situations lacking a variety of potential sexual partners, (e.g. prisons). Behaviorism wrongly dictates that persons in the first situation either lack a sexual orientation or ought to be classified as asexual, and that the sexual orientation of persons in the second situation should be determined with no regard to the extremity of their circumstances. These too are bad results, and ones that blatantly conflict with the general extension of our everyday concept of sexual orientation.

Ideal Dispositionalism

A plausible account of sexual orientation should account for situationally specific sexual behaviors. Behaviorism fails to do this. And yet behaviorism admittedly captures something important about sexual orientation: our concept of sexual orientation tracks (with qualifications) sexual behavior, and not self- or other-identification, emotions, or *purely* psychological states. But insisting that it concerns only *actual* behavior is, as we have seen, riddled with problems. For this reason, I propose that analyses of sexual orientation should move toward dispositional accounts – that is, accounts that define sexual orientation in terms of individuals’ *dispositions* to

⁴³ See, for example, the near-universal recognition of the total failure of so-called ‘reparative therapy’.

engage in sexual behaviors with a certain class of persons (rather than their actual sexual behaviors), and that concern psychological states only insofar as they influence individuals' behavioral dispositions.

After assuming this shift to talk of dispositions, though, significant and difficult questions remain. A standard account of dispositions tells us that:

Something x has the disposition to exhibit manifestation m in response to being situated in stimulating circumstance c iff, if x were to be situated in c , it would exhibit m .⁴⁴

That is, for example, a match is disposed to light (i.e. is flammable) in response to being in a certain circumstance if and only if, were the match in that circumstance, it would light. Applying this to sexual orientation, we can let x range across the domain of human persons, and let m be engagement in sexual behavior (broadly construed) with persons of a certain sex and gender. But determining what c should be is a much more complicated task. And without specifying c , dispositionalism gives us:

A person's sexual orientation is determined solely by what sex[es] and gender[s] of persons S is disposed to sexually engage under certain stimulating circumstances.

⁴⁴ Choi (2008), 796. For simplicity, I have removed the variable ranging across times.

This claim is enough to get us to the position that there is *some* particular scrutability basis of sexual orientation – namely, relevant sexual behavioral dispositions. Without specifying the conditions under which these relevant dispositions manifest, though, we have not made much headway beyond behaviorism. To assign *actual* conditions to *c* would make the view indistinguishable from behaviorism – if the manifesting conditions are actual conditions, then the relevant dispositions should be those dispositions manifested in actual conditions – that is, actual behaviors. And this is precisely what we wanted to avoid. To capture the general extension of our everyday concept, we will need a different theory of what circumstances manifest these dispositions – one less narrow than ‘actual conditions’, and more informative than, say, ‘all the physical facts’.

In Edward Stein’s *The Mismeasure of Desire*, he proposes that the dispositions relevant to determining sexual orientation manifest “*under ideal conditions*.”⁴⁵ He goes on to say that, “Conditions are ideal if there are no forces to prevent or discourage a person from acting on his or her [sexual] desires, that is, when there is sexual freedom and a variety of appealing sexual partners available.”⁴⁶ According to Stein, then, we can understand sexual orientation in terms of the sexual behaviors someone would engage in if nothing – *nothing at all* – were stopping them. While it is not clear that Stein is attempting to capture the general extension of our ordinary concept of sexual orientation, his proposal suggests that one way to fill out the dispositional schema above is as follows:

A person’s sexual orientation is determined solely by what sex[es] and gender[s] of persons S is disposed to sexually engage under *ideal conditions*.

⁴⁵ Stein (1999), 45. My emphasis.

⁴⁶ Stein (1999), 45.

In other words, Stein suggests that sexual orientation is determined by how someone would sexually behave if we held fixed their sexual desires and ensured that *nothing* is stopping them from acting on those desires. We can then consider whether Stein's suggestion can be used to capture the extension of our *ordinary* concept of sexual orientation by specifying that the relevant features of these sexual behaviors are the sex[es] and gender[s] of the persons that they involve.

While I think that this type of proposal rightly pushes against behaviorism's rigid focus on *actual* conditions, I also think that the extreme shift to 'ideal' conditions would create two different problems.

First, this view legitimizes a worrisome cross-cultural projection of our concept of sexual orientation. The ideal conditions for (e.g.) ancient Greeks to realize their sexual desires might vastly differ from the ideal conditions for (e.g.) a contemporary European or North American.⁴⁷ If, for example, political power-dynamics were significantly built into their sexual desires, then the ideal conditions for ancient Greeks acting on sexual desires might be unlike the conditions we think of as ideal for acting on our sexual desires – perhaps they would include a specific political scenario. Likewise for any culture in which sexual desires largely concern (e.g.) social status, wealth, or particular survival skills, not to mention ones where the recognized sexes or genders differ from those built into our concept of sexual desire.⁴⁸ And yet the 'ideal conditions' view suggests that our concept of sexual orientation can be aptly applied within all of these

⁴⁷ For example, Miriam Reumann's (2005) *American Sexual Character* develops a compelling case for the existence of uniquely American sexual desires and patterns, shaped by uniquely American politics, social life, gender roles, and culture, as well as racial and economic divides. The main takeaway for the purpose of this paper is that it would be naïve – and in fact, simply incorrect – to assume that sexual desire has a universal and cross-culturally consistent character.

⁴⁸ Interestingly, Stein addresses this possibility later in his book amidst a discussion of essentialism and constructionism.

widely ranging ‘ideal conditions’. In other words, the view implies that we can ascribe sexual orientation (*as we understand it*) to all human beings across cultures by holding fixed their sexual desires and projecting them into corresponding ‘ideal conditions’, regardless of how foreign these conditions are to the conditions that correspond to contemporary concepts of sexual desire. But given how tightly our concept of sexual desire is entwined in our concept of sexual orientation, I am doubtful that sexual orientation can be cross-culturally applied to this extent. For this reason, this approach would seem to get the extension of our concept wrong, extending it beyond its appropriate reach.⁴⁹

Second, there are some empirical reasons to think that sexual desires cannot be ‘held fixed’ independently of someone’s actual social context, and that these desires would not remain constant when projected into ideal circumstances. In particular, when surrounded by a vast variety of sexual partners and lacking any inhibitions, there is evidence suggesting one’s sexual desires – and so, one’s sexual behaviors – will undergo significant alteration from what they were under ordinary circumstances. In particular, social psychologists have discovered that sexual desires frequently increase or decrease (depending on other characteristics of the individual) in situations with high sexual opportunity. The corresponding principles, known as the “satisfaction principle” (high opportunity decreases desire) and the “adaptation principle” (high opportunity increases desire) are perfect examples of why we should doubt that individuals’ sexual desires would remain constant when they are placed in a situation with

⁴⁹ One might respond on Stein’s behalf that our concept of sexual orientation (and other cultures’ related concepts) should be understood as socially-historically constrained, not allowed of cross-cultural application. While this is not clear from Stein’s text, and I would still have concerns about the view (see the following argument), I do think that this would improve the account. In my own proposal – and as we will soon see – I attempt to develop a way of ensuring this social-historical sensitivity by looking to ordinary language use as a guide to the relevant manifesting conditions.

complete sexual freedom and availability.⁵⁰ But if these desires undergo significant shift, then we should expect that an ‘ideal conditions’ account of sexual orientation will frequently dictate ascriptions of sexual orientation that conflict with our everyday understanding of sexual orientation. To put this in slightly stronger terms, there is some reason to think that it would significantly *change the subject* from what we were originally talking about when we were talking about sexual orientation.

Ordinary Dispositionalism

The problems facing behaviorism and ideal dispositionalism may be avoided by appealing to conditions somewhere between ‘actual’ and ‘ideal’ conditions. One possibility – which I here tentatively adopt – incorporates an adaptation of Sungho Choi’s notion of ‘ordinary’ manifestation conditions for a given disposition. Choi argues that it is possible to articulate manifesting conditions for dispositions that will not be subject to the standard counterexamples to analyses of dispositions (e.g. ‘masked’ or ‘finkish’ dispositions),⁵¹ but which also will not be vacuous (e.g. ‘if it were struck then, unless the match didn’t light, it would light’).⁵² He proposes that this challenge is met by examining the purpose behind our concept of a given disposition, which in turn reveals what manifesting conditions are conceptually connected to that disposition. To put this slightly differently: Choi suggests that whatever conditions those having the concept

⁵⁰ See, for example: Gebauer, Baumeister, Sedikides and Neberich (2014).

⁵¹ A ‘finkish’ disposition is one whose stimulus conditions also removes the disposition – for example, a glass’s disposition to break is finkish if every time it is struck, God hardens the glass so that it is no longer disposed to break. A ‘masked’ disposition is one that is simply prevented from manifesting under the stimulus conditions – for example, a glass’s disposition to break is masked if it is bubble wrapped when it is struck. See Choi and Fara (2014).

⁵² Choi (2008).

consider ‘ordinary’ manifesting conditions for that disposition are the relevant manifesting conditions for that disposition.

In a slight divergence from Choi, I suggest instead that the relevant manifesting conditions for the dispositions determining sexual orientation are not the conditions that those who possess the concept *consider* ‘ordinary’, but the conditions under which people *in fact apply* the term ‘sexual orientation’ (and relevantly associated terms).⁵³ Or, to use the language of purpose (i), the relevant manifesting conditions just are the conditions under which ascriptions of sexual orientation typically take place.

Here I follow Haslanger, who makes a useful distinction between ‘operative’ and ‘manifest’ concepts: The operative concept of ‘cool’, for example, is “the concept that actually determines how we apply the term to cases, i.e. (roughly) being such as to conform to the standards of the in-group.”⁵⁴ In contrast, the manifest concept of cool is “the concept that users of the term typically take themselves to be applying, i.e., being intrinsically or objectively cool.”⁵⁵ By defining ‘ordinary’ manifesting conditions in terms of the concept those applying the term take themselves to have, Choi restricts our search for these conditions to the conditions attached to manifest concepts. But these conditions may be nonexistent (i.e. ‘intrinsic coolness’) or severely mistaken. Better, I think, is to identify the ‘ordinary’ conditions as those corresponding to the everyday operative concept – that is, the conditions corresponding to applications of the relevant terms.

⁵³ I remain neutral on the question of whether this sort of account should be applied widely across all types of dispositions. Perhaps some of the manifesting conditions for some dispositions can be given a purely physics based explanation, for example. Here, I only commit to an account for determining the ordinary manifesting conditions for the dispositions relevant for determining sexual orientation. (Thanks to Michael Rauschenbach for raising this issue.)

⁵⁴ Haslanger (1995), 102.

⁵⁵ Haslanger (1995), 102.

Consider, for example, a match's disposition of flammability. Using this adaption of Choi, we look at the operative concept of 'flammability' and find that the purpose of it is to determine whether a match will light when it is struck in normal temperatures, when dry, etc. These conditions, that is, determine how we apply the term 'flammable' to a match. Because of this, they are the relevant manifesting conditions *c* in the statement, 'A match is disposed to light in response to being situated in stimulating circumstance *c* iff, if the match were to be situated in *c*, it would light.'⁵⁶

Central to this proposal is the idea that finding the ordinary manifesting conditions for a disposition means looking to the (often pragmatic) goals determining application of the term referring to that disposition. In the case of sexual orientation, then, the manifesting conditions for the behavioral dispositions relevant to determining sexual orientation must be understood within the framework of the purposes behind the everyday operative concept of sexual orientation – finding potential partners, establishing laws (be they protective or discriminatory), predicting behavior, enabling scientific research of sexual attraction, and so on.⁵⁷ These purposes determine the 'ordinary' conditions under which the term is applied – that is, they are the conditions corresponding to the operative concept.

⁵⁶ One might worry that this account is circular – that it relies on the 'ordinary manifesting conditions' that identify a disposition by appealing to the concept of that very disposition. Choi (2008) argues that this objection fails. As he points out, because the ordinary conditions for a disposition are understood as "extrinsic conditions that are ordinary to those who possess the dispositional concept" those persons need no knowledge of a conceptual account of the disposition in question.

⁵⁷ While these purposes at bottom will result in the same behaviors as the purposes behind the manifest concept of sexual orientation, they importantly differ in the interpretation of those behaviors. Whereas those applying the term 'sexual orientation' may take themselves to be (e.g.) identifying moral failing or categorizing psychological defects, this is simply using fictions as a mask for what Haslanger calls the "explicitly social content of the operative concept." So too, those applying the term may take the manifesting conditions relevant to sexual orientation to be anything from 'having certain genetics' to 'being cursed by god', but these cannot be the conditions we are concerned with. We are instead concerned with the conditions that *actually* determine application of the term 'sexual orientation', regardless of what someone thinks they are doing when applying it.

Importantly, using these conditions as the relevant manifesting conditions for a particular disposition does not mean forfeiting any revision to the everyday operative concept. It simply means that these conditions are built into the revised concept, guaranteeing that the everyday concept's extension is generally preserved. These conditions only provide constraints on the eligible criteria for ascribing sexual orientation – they do not determine these criteria, much less determine the taxonomy resulting from them.

More needs to be said about what these ordinary conditions are – that is, what *are* the conditions corresponding to the everyday operative concept of sexual orientation? Or, in other words, what conditions lie behind our ascriptions of sexual orientation?

My primary goal in moving away from actual or ideal conditions, and toward the conditions corresponding to the everyday operative concept of sexual orientation, is to escape the rigidity of both behaviorism and ideal dispositionalism. I want to avoid a view that ascribes sexual orientation on the basis of only observable behaviors, or only behaviors within unattainable, potentially culturally distant ideal conditions. This is not to say that the operative concept is without substance. But I will not pretend to articulate necessary and sufficient conditions corresponding to the operative concept of sexual orientation – nor do I maintain that such conditions exist. The conditions corresponding to our ascriptions of sexual orientation admit, no doubt, of borderline and vague cases. My primary concern is to capture the core elements of these conditions in order to generally preserve the extension of our everyday concept of sexual orientation.

And I do think that a number of things can be said to elucidate the conditions under which we typically confer sexual orientation. In particular, I propose the following as conditions constraining our ascriptions of sexual orientation – that is, as conditions corresponding to the

operative concept – reminding the reader to think of these as generalities that admit of exception and vagueness, rather than as strict rules of use.

- (I) The operative concept assumes attraction to persons of a certain sex or gender (at least partially) *because* they are that sex and/or gender.

For example: Say that Elijah has strong sexual druthers for persons with long hair, but has no preference between men or women as sexual partners. Elijah lives in a town where the only people with long hair happen to be women. As a result, it is true to say that Elijah is attracted only to women. But because he is not attracted to them *because* they are women, we would not say that Elijah has the sexual orientation corresponding to exclusive attraction to women.⁵⁸ This sort of case illustrates that we expect an explanatory relation to hold between one's sexual orientation and the sex or gender of the persons they are attracted to.

- (II) The operative concept assumes attraction to certain persons while having a reasonable diversity of potential sexual partners.

This generalization is far from Stein's suggestion that sexual orientation is based on attractions with *no* restriction on sexual partners. But it captures why we do not consider behaviors in outlying circumstances where potential partners are extremely limited or homogenous (e.g. prisons, boarding schools, deserted islands) as reliable indicators of one's sexual orientation.

⁵⁸ Or because Elijah *thinks* they are women. This would leave room for cases in which, e.g., someone attracted to cisgender men is attracted to someone they take to be a cisgender man, but who is anatomically female.

- (III) The operative concept assumes that one is willing and able to sexually engage with other persons.

We refuse to ascribe sexual orientations to someone on the basis of their actual sexual behaviors if (e.g.) they are voluntarily celibate, subject to sexual contact without consent, or possess a prohibitive medical condition. These scenarios indicate that it is also important to the operative concept of sexual orientation that the behaviors relevant to ascribing sexual orientation are ones that are engaged in willingly and with the physical and psychological ability to engage or not engage in the behavior.⁵⁹ It might also explain why we judge abnormal sexual behavior under the influence of alcohol or narcotics (and therefore, nonconsensual) to be an unreliable indicator of sexual orientation.

Again, (I)-(III) are generalizations of the conditions that I think are built into the operative concept of sexual orientation, and they therefore will admit of occasional exceptions or borderline cases. They remain, though, useful guidelines explaining why we consider extreme circumstances poor guides to determining sexual orientation, and how we can reliably ascribe sexual orientation to persons without appealing to ‘ideal’ conditions.

Someone may here object that, in appealing to the conditions underlying our operative concept of sexual orientation in order to construct a revised concept of sexual orientation, I appear to be doing mere conceptual-analysis. To this, I would again emphasize that my project is necessarily in part descriptive because it aims to generally preserve the extension of our everyday operative concept of sexual orientation. But it is prescriptive insofar as I am out to

⁵⁹ Of course, some (and perhaps all) asexual persons will never be in a situation in which they are willing to engage in sexual behavior. In that case, we can determine that, because it is impossible for them to meet condition III, they do not have any sexual behavioral dispositions that would be manifested under the ordinary conditions – that is, they are asexual. This distinguishes asexuals from (e.g.) voluntary celibates.

precisify and revise this concept in order for it to more efficiently and ethically serve the purposes assigned to it. Noticeable revisions concern decisively including both sex- and gender-attraction as criteria for ascribing sexual orientation and – as we will now see – distancing the concept from concerning the relation between a subject’s own sex or gender and the sex or gender of the persons they are attracted to. These revisionary aims are importantly distinct from projects that intend to radically revise the conditions determining concept deployment.⁶⁰

4.3.2 Conceptualizing Bidimensional Dispositionalism

Putting together the previous discussions of gender/sex and behaviorism/dispositionalism, we arrive at my positive proposal:

Bidimensional Dispositionalism [BD]: A person S’s sexual orientation is grounded⁶¹ in S’s dispositions to engage in sexual behaviors under the ordinary condition[s] for these dispositions,⁶² and which sexual orientation S has is grounded in what sex[es] and gender[s] of persons S is disposed to sexually engage under these conditions.^{63,64}

⁶⁰ See Haslanger (2005), 114.

⁶¹ I use the term ‘grounding’ here in the loose sense of ‘is dependent on’ or ‘is explained by’. One may also be able to understand it in terms of ‘is built on’, which (roughly) means ‘is less fundamental than’ or ‘is accounted for in terms of’. [See Bennett (forthcoming)].

⁶² I remain neutral on the debate over whether properties can have dispositional essences or if all dispositions reduce to categorical properties. For my purposes here, I don’t have a dog in that fight.

⁶³ As Shamik Dasgupta pointed out, someone might be concerned that BD, as stated, does not ensure that the dispositions relevant to sexual orientation are particularly ‘deep’ or ‘self-disclosing’. Sexual orientation, one might think, deserves protection *because* it is deep and in this way outside (or mostly outside) a person’s control. While I acknowledge this worry, I disagree with the idea that sexual orientation must be particularly ‘deep’ to merit special protections. Whether or not sexual orientation has these features is orthogonal to its merit for protection. Even if we shifted sexual orientation every week, (and even if we could do so by choice), I would insist that sexual orientation deserves protections. However, one might worry that even apart from questions of protections, sexual orientation is

In other words, I propose that – whatever the categories we place within ‘sexual orientation’ – their ascription should be based on the sex and gender of the persons someone is disposed to sexually engage with under ordinary conditions for ascribing sexual orientation.⁶⁵

This analysis recasts sexual orientation as pertaining to bidimensional attraction – that is, as pertaining to both sex- *and* gender-attraction. But, importantly, BD does not require that, in order to be ascribed a sexual orientation, someone must have a certain sex-attraction or gender-attraction. One could be neutral as to one or both, or be attracted to neither (i.e. be asexual with regard to sex and gender). All of this would be revealed by their dispositions to engage (which could be dispositions to never engage) in sexual behavior with certain persons (at least partially) on the basis of their sex and gender.⁶⁶

By emphasizing only these dispositions, BD understands sexual orientation *solely in terms of the sex and gender of the persons one is disposed to sexually engage, without reference to the sex or gender of the person so disposed*. Under this framework, for example, a cisgender man and transgender woman disposed to sexually engage only with cisgender women have the same sexual orientation, and so too for a cisgender man and gender-nonconforming female disposed to engage only with men. In emphasizing this shift in our categories of sexual orientation, BD

a ‘deep’ and unchangeable (or nearly unchangeable) feature of who someone is. I want to remain neutral on this question, and so the formulation of BD allows but does not require someone’s sexual orientation to undergo frequent shifts.

⁶⁴ By ‘sex[es] and gender[s] of persons...’ I do not mean to imply that there must be any particular persons of this sex and gender, or particular persons with whom S is disposed to sexually engage. That is, S can be disposed to engage with persons who are cisgender women even if there were no cisgender women, or even if there were no particular cisgender women with whom S is disposed to engage.

⁶⁵ While I will not address this issue here, we arguably should also acknowledge that these dispositions themselves come in a range of strengths, which would add an additional dimension to sexual orientation. (Thanks to Justin Christy for this suggestion.)

⁶⁶ I expect that we are often attracted to certain persons because they have characteristic that are associated with particular genders, and not because of the totality of their gender expression. For my purposes, this sort of connection is sufficient to allow for the explanatory connection between gender and attraction, though it leaves many open questions regarding what (if anything) is essential to particular gender expressions, and more generally, how we should think about the constitution of gender expressions. I leave these questions to persons working in the metaphysics of gender.

rejects the idea that sexual orientation can be classified in terms of a relation between persons of the ‘same’ or ‘opposite’ sex or gender.

This taxonomical shift is important to the fulfillment of purposes (iii)-(iv). Recall that these purposes stated that an analysis of sexual orientation should “reduce or eliminate the presumption that cisheterosexuality is the normatively standard sexual orientation and all queer sexual orientations are normatively deviant,” and be “conducive for establishing legal and social protections for non-cisheterosexual persons.” The categorization shift proposed by BD moves us closer to accomplishing both of these tasks.

First, BD promotes the aims of purpose (iii) because BD *eliminates the distinction between cisheterosexuality and queer sexual orientations* and *provides a taxonomic schema capable of recognizing persons outside the gender or sex binary*. On the former point, on BD, there are no such sexual orientations as (e.g.) ‘homosexual’ or ‘heterosexual’. And there is no distinction in the sexual orientations of (e.g.) a cisgender man and a transgender woman who both are exclusively attracted to women. The statistical divide between cisheterosexuality and queer sexual orientations simply disappears because these categories disappear, and their members are reorganized into new categories. While this will not of itself eliminate discriminatory attitudes, it does change the concept of sexual orientation such that it does not simply fall out of the concept that cisheterosexuality is statistically standard, and all else is deviant. It also removes the connotation that ‘sexual orientation’ is what distinguishes (e.g.) the so-called ‘straight’ and ‘queer’ communities. I believe that this is a socially and politically beneficial result, encouraging dismantling the divide between these communities.

On the latter point, BD does not build in either discrete or binary gender or sex categories, and so has the flexibility to adopt a variety of sex and gender taxonomies. With this flexibility, it

is capable of providing taxonomic recognition for persons out the sex or gender binaries (e.g., gender queer or intersex persons), as well as their sexual partners.

Second, BD achieves (or at least moves toward achieving) purpose (iv) by providing the conceptual tools for lawmakers to secure protections for sexual orientation under pre-existing protections against gender- and sex-discrimination. That is, because sexual orientation makes no reference to one's *own* sex and gender on BD, any discrimination against someone in response to their sexual orientation can be re-described as discrimination on the basis of their gender or sex.

This conceptual shift is, in fact, ripe for public uptake. Supreme Court Chief Justice John Roberts recently articulated a similar shift in thought during oral argument in *Henry v. Hodges*, a case concerning the legalization of same-sex marriage. Justice Roberts re-described the same-sex marriage question in terms of sex discrimination, and (perhaps rhetorically) asked why the issue could not be decided on the basis of pre-existing protections against sex discrimination:

I'm not sure it's necessary to get into sexual orientation to resolve this case... I mean, if Sue loves Joe and Tom loves Joe, Sue can marry him and Tom can't. And the difference is based on their different sex. Why isn't that a straightforward question of sexual discrimination?⁶⁷

As Justice Roberts here notes, cases of sexual orientation discrimination can be easily re-described in terms of gender or sex discrimination by holding fixed that multiple individuals share the same sex or gender attractions, and yet some are discriminated against simply because they have a particular sex or gender in addition to those attractions. BD goes a step further by

⁶⁷ Liptak (2015). This argument in favor of legalizing same-sex marriage on the basis of pre-existing laws against sex-discrimination also was the central argument of an *amicus curiae* brief filed by a number of legal scholars in *Henry v. Hodges*, Supreme Court Case No. 14-556.

saying that the sex and gender attractions – again, understanding these attractions in terms of behavioral dispositions under ordinary conditions – are all that matter for sexual orientation. My own sex and gender, for example, do not matter for my sexual orientation. And so, if I am discriminated against for having the attractions constituting sexual orientation X and a man who has sexual orientation X is not discriminated against, I can recast this discrimination as gender discrimination and appeal to pre-existing laws prohibiting this discrimination as the basis for my legal protection.

One worry about BD is that it does not give us enough epistemic access to our own sexual orientations. How will we know our sexual orientation if it would require being places under circumstances that we are not actually in? How could, say, a lifelong celibate priest *know* that they would take certain actions if they were under these ‘ordinary circumstances’? It might seem as though any compelling account of sexual orientation will make it possible for someone to know their own sexual orientation, and BD does not do this.⁶⁸

This objection, though, makes a substantive assumption: that the correct metaphysical analysis of sexual orientation must bend to a demand for epistemological transparency (or something close to transparency). And I see no reason to think this. In fact, we have good reason to deny it, given the many examples of repression and self-deception of sexual orientation under (e.g.) social, religious or familial expectations.⁶⁹ And this does not mean that we have no idea what our sexual orientations are – in general, people seem to have a ‘good enough’ idea of their sexual attractions and how they do or would act under certain circumstances that they also have a ‘good enough’ idea of their sexual orientation to seek out specific (or no) sexual partners.

⁶⁸ Thanks to Peter Finnochiaro for raising this objection.

⁶⁹ Indeed, the testimony of many queer persons suggests that discovery of one’s own sexual orientation can be a long and difficult process.

Insofar, too, as we think that persons have some manner of epistemic privilege in self-assessments of desire, attraction, and so on – features that inform and direct their behavioral dispositions – we can maintain that persons also have some manner of epistemic privilege in self-ascriptions of sexual orientation.

Another worry for BD concerns the relation between sexual dispositions and sexual desires. Why, someone might ask, should we go to the trouble of analyzing sexual orientation in terms of dispositions and all their metaphysical baggage, when we can much more simply analyze it in terms of sexual desire, understood as an occurrent mental state?

The first and most important response is, I think, to emphasize that in order to achieve the pragmatic goals discussed earlier, it is important to avoid an account that wholly psychologizes sexual orientation. While I leave open that the behavioral dispositions for which persons need political and social protections have categorical psychological bases, these bases will not be the focus of a socio-politically oriented account of sexual orientation. For these purposes, someone with the psychological features of a ‘heterosexual’ but queer behavioral dispositions can and should be protected from anti-queer prejudice. In other words, given the pragmatic goals of my account, the questions of whether or which psychological states ground the behavioral dispositions at issue are interesting questions for neuroscientists, but not one that should guide a politically motivated account of sexual orientation.

Second, given the current main contending theories of desire, an account in terms of desire either amounts to a problematically restricted dispositional view or else creates new (and worse) problems. Suppose, for example, that one is partial to an action-based theory of desire, articulated in terms of dispositions. In this case, sexual desires just are the categorical basis of the kinds of behavioral dispositions that I have been talking about. At first, this might seem like this

view would be co-extensive to my own, but preferable because it is articulated in familiar terms (desire) rather than in the technical language of dispositions. This thought would be mistaken. A sexual desire view of sexual orientation would not be co-extensive with my own, because ordinary dispositionalism allows sexual behavioral dispositions to have a range of categorical psychological bases (or no categorical basis), and certainly does not restrict the relevant dispositions to ones grounded in the mental states that we would categorize as ‘sexual desire’. For example, if someone is attracted to women on the basis of, say, a constant curiosity about what it is like to have sex with women, but not because of desire-like attitudes typically considered sexual attraction, my account does not rule out that this person can be classified as sexually women-oriented. A view of sexual orientation restrict to sexual desire (and the corresponding behavioral dispositions) would not be able to accommodate this case.

If instead one prefers, for example, a pleasure-based theory of desire, then the formulation will be too narrow to capture the concept of sexual orientation. Not everyone receives pleasure from sexual behavior. Still other theories of desire (e.g. attention-based or holistic) are too broad to capture the concept.⁷⁰ And so on, I would argue, for the other main candidate theories of desire. Of course, one could simply appeal to a ‘common understanding’ of desire, but I am skeptical that there is any such thing.

Third, one might worry that desires are too context-sensitive to capture the general (though perhaps not necessary) stability of sexual orientation. Earlier, I argued that we should not use Stein’s “ideal conditions” as the relevant manifesting conditions for sexual behavioral dispositions because we have good reason to think that someone’s desires would be significantly altered in a scenario with wholly unrestricted access to a huge variety of sexual partners.

⁷⁰ For an overview of these (and other) theories of desire, see Schroeder (2014).

Whatever these desires (and the corresponding behaviors) are, I argued, they are not reliable indicators of someone's actual sexual orientation. But, similarly, I think there are cases where someone's actual desires are not reliable indicators of their sexual orientation. For example, it is reasonable to expect someone in a context lacking reasonable variety of potential partners, such as a prison, may undergo shifts in sexual desire. And yet we would, I think, still deny that these shifted desires are reliable indicators of their sexual orientation, or that these shifts in desire constitute a shift in their sexual orientation. Perhaps, to avoid this result, one could insist that the desires relevant to sexual orientation are those that one *would* have in the ordinary conditions that I've described. But, in that case, we've only moved from behavioral to psychological dispositions; we haven't gotten rid of dispositions, or pinpointed particular mental states determining sexual orientation.

My respondent might be fine with accounting for sexual orientation in terms of dispositions to desire rather than dispositions to behavior, strange as it may seem. But framing sexual orientation in this way would not avoid yet another problem for any desire-based account: dispositions to desire would underdetermine sexual orientation because *desire* underdetermines sexual orientation. Consider someone who is behaviorally disposed to sexually engage with certain persons, but does not possess the emotional or cognitive features of sexual desire. (Again, we could imagine that they are motivated to sexually engage with persons on the basis of curiosity, free of desire.) That is, they don't (e.g.) feel sexual yearnings, spend time thinking about sexual behavior or receive particular pleasure from sexual behavior. Does this person have a sexual orientation? It seems to be that they do, suggesting that desires are not *necessary* for sexual orientation.

But are they sufficient? Consider too the unlikely but imaginable case of someone who feels desire for say, cisgender men, but is disposed only to sexually engage with women. In this case, and particularly for the socio-politically motivations discussed above, I would argue that this person's sexual orientation is one of orientation toward women, and not cisgender men. But I admit that intuitions about our concept's extension may get fuzzy with regard to both of these hypotheticals – I can only report my own. I suspect that one's response may come down to whether they tend to think about sexual orientation as something predominately action-oriented or predominately internal. But more importantly (given that this is an engineering project), I support the former view as better equipped to achieve the social and political purposes behind the concept of sexual orientation, and as not clearly in conflict with the general extension of our everyday concept.

Even with this advantage, BD is only the beginning of a full analytic account of sexual orientation. It remains neutral on a number of important and closely related philosophical questions. I turn now to these questions, which will highlight where further research can expand philosophical discussion of sexual orientation.

Discrete/Continuous Categories

BD requires that we acknowledge that sexual orientation is bidimensional, pertaining to both sex- and gender-attraction. Within this bidimensionality, though, we can and should ask whether our categories of sexual orientations will be discrete or continuous. The most common current categories of sexual orientation are uniformly discrete. In ordinary discourse, we typically hear two, or at best three discrete categories: 'heterosexual', 'bisexual' and 'homosexual'. Expanding

our concept of sexual orientation to include both sex and gender will increase the number of sexual orientation categories. But it will not of itself revise these categories such that they are no longer discrete, since it will not of itself revise our categories of sex and gender to no longer be discrete (much less binary).

Suppose someone accepts, for example, a sex binary (male and female) and a gender binary (men/women). In that case, under BD, they would have four categories for attraction qua sex and four for attraction qua gender, arriving at sixteen discrete categories of sexual orientation. Consider the following example of how one might retain discrete categories of sexual orientation under BD, where ‘attraction’ can be taken as a useful shorthand signaling dispositions to engage in sexual behavior:

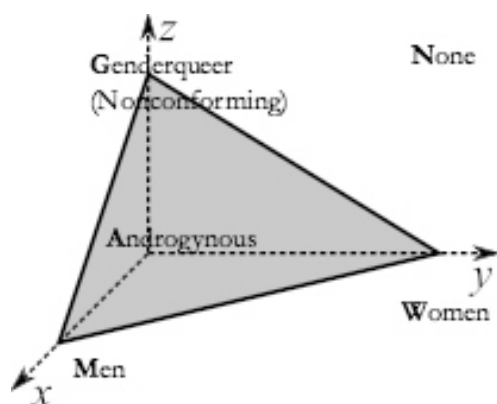
Illustration 1: Discrete Categories

Sex-Attraction (Assumes two sexes – male/female)	Gender-Attraction (Assumes two genders – men/women)
A: Not attracted to either sex qua sex	1: Not attracted to either gender qua gender
B: Attracted to males	2: Attracted to men
C: Attracted to females	3: Attracted to women
D: Attracted to males and females	4: Attracted to men and women

Under a view such as this, a category within sexual orientation might be ‘C2’, which refer to someone who – under ordinary conditions – is disposed to engage in sexual behaviors only with transgender men who have not had genital surgery. One could also easily add ‘intersex’ to the left-hand column, ‘genderqueer’ to the right-hand column, and so on. The element I mean to emphasize in such a view is the insistence upon discrete categories for sex and gender, which leads to discrete categories of sexual orientation.

If instead (as I prefer), one understands gender (and perhaps sex) to refer to a continuous spectrum, then sexual orientation will also refer to a continuous spectrum. Of course, within ordinary discourse, we typically have heuristic markers along continuous spectrums for pragmatic purposes, (e.g. ‘tall’ or ‘hot’).⁷¹ But it is generally understood that these markers are merely heuristic, and do not refer to neatly closed categories. This could easily be applied within discourse about sexual orientations. The following illustrate a view under which sex categories are discrete, but gender categories are continuous, resulting in continuous categories of sexual orientation:

⁷¹ In other words, I prefer an account that rejects epistemicism about gender categories, just as most of us, I think, would reject epistemicism about ‘tall’ or ‘hot’.

*Illustration 2: Continuous Categories***Gender-Attraction.**⁷²

Sex-Attraction
(Assumes two sexes – male/female)
1: Not attracted to either sex qua sex
2: Attracted to males
3: Attracted to females
4: Attracted to males and females

Under this view, a category within sexual orientation might be ‘MA-1’, heuristically referring to someone who – under ordinary conditions – is disposed to engage in sexual behaviors with persons who present as (roughly) androgynous or masculine, regardless of their sex.

BD remains neutral on the issue of discrete vs. continuous categories of sexual orientation. But the questions surrounding the issue require much closer attention, and my hope is that further

⁷² Attraction may be represented at one, no or multiple locations on the diagram

research on sex and gender will allow an expansion of BD that specifies sexual orientation as continuous. It would thereby become more equipped to recognize sexual diversity and fulfill the need for an account of sexual orientation that eliminates the idea that cisheterosexuality is the ‘standard’ sexual orientation and all else is ‘deviant’.

Essentialism/Constructionism

As stated, BD is also neutral on the question of essentialism vs. constructionism about sexual orientation. Roughly, essentialism is the view that sexual orientation is something necessary or unifying about humans as a kind or as particular humans (thereby applying cross-culturally), whereas constructionism is the view that sexual orientation is socially constructed (thereby culturally-specific), and may not (depending on the type of construction) in any sense ‘carve at the joints’ of reality.⁷³

One common form of essentialism is biological essentialism – often heard in the sentiment ‘Born This Way’ – which claims that sexual orientation is a biologically determined feature of a person.⁷⁴ While BD is compatible with this view, to hold both one must hold to biological determination of gender-attraction. That is, if sexual orientation is to be genetically determined *and* pertain to gender-attraction, biological essentialism faces the difficult challenge of explaining how gender could be (at least partially) socially constructed, and yet gender-attraction

⁷³ See Haslanger (1995) for a wonderful look at how different levels and kinds of social constructions correspond to what we admit into our ontology.

⁷⁴ See, for example, LeVay’s (2011) analysis of sexual orientation as “*the trait that predisposes us to experience sexual attraction to people of the same sex as ourselves (homosexual, gay, or lesbian), to persons of the other sex (heterosexual or straight), or to both sexes (bisexual).*” [Emphasis added.] By identifying sexual orientation with the [according to him, biological] trait that predisposes us to have certain attractions, LeVay adopts a biological essentialism about sexual orientation.

be biologically determined.⁷⁵ For this reason, it seems that BD – though compatible with biological essentialism – is unlikely to be paired with it.

I do not have space here to discuss alternative forms of essentialism about sexual orientation, other than to say that I think it will be difficult to find a form of essentialism that agrees with contemporary theories of gender, which almost always incorporate some degree of constructionism.⁷⁶ For this reason, I tentatively lean toward a moderately constructionist view according to which our social context (which may be self-selected to some extent) directs the manifestation of and categories for the manifestation of biological tendencies toward certain sexual attractions. On this view, the concepts and interpretive framework related to sexual orientation are socially constructed, but they refer to real properties (or traits, if you prefer) of individuals that are in part socially formed and in part biologically determined.⁷⁷

Moderate constructionism's main benefit is allowing the possibility of cross-cultural behaviors and biological traits loosely translating onto a contemporary framework – namely, the framework of the 'ordinary conditions for ascribing sexual orientation' – while maintaining that the contemporary understanding of sexual orientation should not be applied to cultures lacking the concept of sexual orientation *within that culture*. But I also worry that moderate constructionism gives too prominent of a place to biology, and underestimates the degree to

⁷⁵ Of course, one could also insist (against the prevalent view) that gender is biologically determined.

⁷⁶ One interesting line of inquiry would be whether Charlotte Witt's 'uniessentialist' account of gender, which combines both individual essentialism and social construction, could be applied to sexual orientation. On this view, certain properties of someone, such as their gender, can be essential to that person as a social individual, which Witt views as a one of the three parts in the trinitarian ontology of 'selves' (i.e. human organism, person, social individual). If one is willing to adopt Witt's ontology of 'selves', perhaps a case could be made for sexual orientation as an essential property of a social individual.

⁷⁷ I predict that whether you consider this a realist or fictionalist account of sexual orientation will largely depend on whether you are a realist or fictionalist about certain social kinds. Within the distinction of idea- and object-construction, this view would incorporate elements of each. While sexual orientation classifications would be considered idea-constructions, the way that we are socially and politically impacted by these classifications (whether our self-identity aligns with our sexual orientation or not) incorporates a large dose of object-construction into our understanding of persons as sexual orientated beings. [For more on this issue, see Haslanger and Ásta (2011).] I am partial to the idea that there is a balance between these constructions and biological influence in determining sexual orientation.

which sexual orientation is socially constructed. In any case, I currently have only tentative opinions about these issues, and so I have chosen to here state BD as neutral on issues of essentialism and constructionism.⁷⁸

4.4 Conclusion and Implications

Although I've gone to lengths to clarify what issues I do not take a firm position on, I do not mean to leave the impression that BD is an uncontroversial thesis. To clarify this, I will briefly state its central implications.

First, to adopt BD is to reject our current taxonomy of sexual orientation. The assumptions that sexual orientation is always one-dimensional – concerning either sex-attraction or gender-attraction, but never a combination of the two – and that sexual orientation concerns the sex or gender of *both* potential partners are deeply embedded within the concepts associated with these categories. For this reason, the current concepts of 'homosexual' and 'heterosexual' do not pick out sexual orientations under BD. These terms inherently refer to a relation between the sexes (or genders) of sexual partners, whereas BD focuses exclusively on the sex and gender of the persons one is attracted to.

Second, BD implies – but does not necessitate – that we should reject biological essentialism about sexual orientation. This again comes on the heels of BD's incorporation of both sex- and gender-attraction in its analysis of sexual orientation. Third, and also for this reason, BD implies that the categories of sexual orientation ought to be continuous, rather than discrete.

Contemporary research suggests that the categories of gender (and perhaps sex) are continuous,

⁷⁸ For an interesting perspective on the need for recognizing individuals' agency in determining their sexual orientations, see Behrens (2013).

and so any dispositions related to gender (including the ones at issue in BD) must be sensitive to this continuous scale. This has political and social implications, as it raises questions about how to understand sexual orientation as a protected class or legal decisions concerning sexual orientation, and it puts pressure on the idea that cisheterosexuality (or any sexual orientation) is normative or non-deviant.

Fourth, because BD appeals to the ‘ordinary conditions’ for ascribing sexual orientation, it requires that we hesitate in ascribing any category of sexual orientation to an individual on the basis of their behavior without first carefully considering their social context. This hesitation is particularly required when considering persons under (e.g.) religious, situational or familial pressures to partner with someone of a certain sex or gender, as well as any cross-cultural applications of our categories of sexual orientation.

Though not directly tied to BD, I also hope to have shown that the issues surrounding sexual orientation need further philosophical analysis. A vast number of questions about essentialism, dispositions, choice, reduction, social kinds and properties – not to mention the political and ethical implications of our answers – remain unexplored. BD only scratches the surface of this promising philosophical landscape.⁷⁹

⁷⁹ Many thanks to audiences at the San Francisco Bay Area Feminism and Philosophy Workshop, the Central APA (2015), the University of Notre Dame, New York SWIPshop, the Berkeley Social Ontology Group, and UC Irvine’s Perspectives on Gender conference, as well as two anonymous referees at *Philosopher’s Imprint* for feedback during the development of this chapter.

Chapter 5

Conclusion

The most important theme unifying “Categories We (Aim to) Live By” is best captured by Sally Haslanger’s comment that “the task [of social critique] is to situate ourselves differently in the world, not just to describe it more accurately.”¹ My goal here has been to engage in social critique as it applies to social categories: to not simply analyze the categories we live by, but also to articulate mechanisms for building the categories we *aim* to live by. This approach is not new – it is well grounded in the work of previous feminist ontologists, along with LGBTQ philosophers and critical race theorists (among others). Those within these traditions already are well aware of the distinction and relationship between revealing and revising unjust social structures. There is an important sense, then, in which I am proud to say that I am doing nothing new.

That said, taking seriously Marx’s adage that critical theory (including social critique) is “the self-clarification of the struggles and wishes of the age,”² I hope that the previous chapters have captured either new struggles against the categories we live by, or else captured old struggles in a new way. One of the most important struggles, for my purposes, is the struggle against a so-called value neutral conception of analytic metaphysics, and especially the delusion that to do analytic metaphysics is only to *describe* reality (and perhaps only ‘fundamental’ reality), but

¹ Haslanger (2012), 29.

² In Haslanger (2012), 22, citing Fraser (1989, 113), citing Marx (1843).

never to dabble in *prescribing* it.³ My goal is to articulate, reveal, and apply the conviction that when metaphysical inquiry enters the social domain, such a distinction is not only impossible but also undesirable.⁴

Why impossible? As I argued in “Oppressive Truths”, how we articulate social categories is not insulated from the social mechanisms forming and sustaining those categories. There are no value-neutral descriptions of social categories, because there are no descriptions of social categories that do not cue – for some purpose – background ideologies underlying these categories. To purport to describe such a category – say, *black* or *women* – in a value-neutral way is to simply to further entrench the dominant ideology sustaining that category; to describe it critically is to advocate revision to that category.

This point also speaks to why the distinction is undesirable. Ontological categories in the social domain – and perhaps in other domains as well – guide the distribution of social power.⁵ Return to Ron Mallon’s point – also mentioned in the introduction – that social categories are used for “explaining and predicting the behaviors of other individuals and groups; signaling to and coordinating with others; representing the world to ourselves; and stigmatizing, valorizing, and regulating the behavior of ourselves and others.”⁶ These categories are tools for social interaction at an individual level, as well as structural and institutional levels. As such, it seems fitting if not necessary that the questions we ask about these categories are guided not only by the usual epistemic standards, (e.g., adequacy, coherence, and simplicity), but also by the goal of

³ Both such views of metaphysics are nicely encapsulated in Schaffer (2009), which begins, “On the now dominant Quinean view, metaphysics is about what there is. ... I will argue for the revival of a more traditional Aristotelian view, on which metaphysics is about what...[is] *fundamental*.” See Barnes (2014) for a critique of Schaffer’s view.

⁴ This point is also forcefully made in Haslanger (2012) and with respect to scientific inquiry in Anderson (1995).

⁵ I here follow Fricker (2007) in understanding social power as “a socially situated capacity to control others’ actions.” For more on the notions of being ‘socially situated’, see Donna Haraway (‘Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective’, *Feminist Studies*, 14, no. 3 (1988), 575–99).

⁶ Mallon (2016), 1.

reorienting social understanding so as to create categorical tools that encourage just distribution of social power. This method requires embracing the contingency and situatedness of social ontology: metaphysical analysis of social categories ought to be concerned (and openly so – I’m not advocating hidden agendas) with issues of justice that are salient in one’s context.

This outlines some reasons *for* recognizing and embracing the value-laden nature of social ontological inquiry. “Oppressive Truths” and “Real Talk on the Metaphysics of Gender” also reveal a danger in *not* recognizing it: distraction from the political questions that matter most by a fixation on discovering metaphysical truth, or ‘getting it (ontologically) right’. As discussed at length in “Real Talk on the Metaphysics of Gender”, this is particularly clear in the discourses surrounding gender ascriptions. Rather than focus on what gender categories should be operative, these conversations frequently get bogged down in extremely difficult metaphysical questions about the features of our current gender categories, and what qualifies someone for membership in these categories. These questions often lead to endless exchanges of counter-examples, or to people talking past each other. This could be avoided if we used political goals to direct gender categorization, shifting the focus of disagreement from *what makes someone a woman?* to the more tractable question *what should be our political goals with respect to gender categories?*

“What is Sexual Orientation?” applies this approach to social ontology. Rather than simply describe current sexual orientation categories, I criticize and propose revisions to them in light of explicit political aims that are salient to me not only because of my social-historical context, but also because of my queer identity. Of course, it is important to be explicit about the potential advantages and limits of a socially situated project. As discussed at length in standpoint and especially feminist standpoint epistemology, there is a potential evidential advantage in drawing from one’s individual perspective, and especially while exploring the oppressive effects of social

categories.⁷ But situatedness also brings temporality. As I just mentioned, my view is formed in response to the political concerns of my time and place – hopefully, political concerns regarding sexual orientation will continue to change and progress. My project is far from timeless. But importantly, this is part of what feminist ontology requires. Rather than hide behind a veneer of value neutrality, feminist ontologists openly proceed with skin in the game, engaging in metaphysical inquiry that begins and ends with concerns of justice in our own historical moment.

⁷ For more on this point, see (e.g.) Anderson (1995a) & (1995b) and Hartsock (1983). For more on its application to feminist ontology, see Haslanger (2012), 24.

Bibliography

ADA National Network (2013) “An Employee View of the Changes from the ADA Amendments Act.” <https://adata.org/factsheet/employee-view-ada>. Accessed April 16, 2017.

Alcoff, L. (2010) “Epistemic Identities”. *Episteme* 7:(2): 128 – 137.

Alcoff, L. (2005). “The metaphysics of gender and sexual difference.” In Barbara S. Andrew, Jean Clare Keller & Lisa H. Schwartzman (eds.), *Feminist Interventions in Ethics and Politics: Feminist Ethics and Social Theory*. Rowman & Littlefield Publishers.

Allen, S. (2017) “It’s Time to Debunk Rachel Dolezal’s Transgender Lie.” *The Daily Beast*. <http://www.thedailybeast.com/articles/2017/04/06/it-s-time-to-debunk-rachel-dolezal-s-big-transgender-lie.html> Accessed April 20, 2017.

American Psychological Association. (2008). *Answers to your questions: For a better understanding of sexual orientation and homosexuality*. Washington, DC. www.apa.org/topics/orientation.pdf. Accessed April 16, 2017.

Anderson, E. (1995a.) “Feminist Epistemology: An Interpretation and a Defense,” *Hypatia: A Journal of Feminist Philosophy*, 10(3): 50–84.

Anderson, E. (1995b) “Knowledge, Human Interests, and Objectivity in Feminist Epistemology,” *Philosophical Topics*, 23(2): 27–58.

Andler, M. (forthcoming) “Gender Identity and Exclusion: A Reply to Jenkins.” *Ethics*.

The Asexual Visibility & Education Network (2012). *General FAQ*. URL = <http://www.asexuality.org/home/overview.html>.

Ásta (2011) “The Metaphysics of Sex and Gender.” In Charlotte Witt (ed.), *Feminist Metaphysics*. Springer.

Ásta (2013). “The Social Construction of Human Kinds.” *Hypatia* 28 (4):716-732.

Bach, T. (2012). “Gender Is a Natural Kind with a Historical Essence.” *Ethics*, Vol. 122, No. 2, pp. 231-272.

Barnes, E. (manuscript). “Gender Realism and Gender Terms.”

Barnes, E. (2014) "Going Beyond the Fundamental: Feminism in Contemporary Metaphysics." *Proceedings of the Aristotelian Society* 114 (3pt3):335-351.

Barnes, E. (2016). "Realism and Social Structure." *Philosophical Studies*.

Basu, R. (ms) "Wronging Beliefs."

Bauer, N. (2015) *How to Do Things with Pornography*. Harvard University Press.

Behrens, M. (2013). "Born That Way? The Metaphysics of Queer Liberation." *APA Newsletter on Philosophy and Lesbian, Gay, Bisexual and Transgender Issues* 12 (2):2-7.

Bennet, K. (Forthcoming). *Making Things Up*. Oxford University Press.

Bettcher, T. (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. pp. 233-250.

Brubaker, R. (2016) "The Dolezal affair: race, gender, and the micropolitics of identity." *Ethnic and Racial Studies* 39(3): 414-449.

Brubaker, R. (2006) *Ethnicity Without Groups*. Harvard University Press.

Burkett, E. (2015) "What Makes a Woman?" The New York Times.
https://www.nytimes.com/2015/06/07/opinion/sunday/what-makes-a-woman.html?_r=0
Accessed April 16, 2017.

Choi, S. (2008) "Dispositional Properties and Counterfactual Conditionals," *Mind* 117:468: 795-841.

Choi, S. and Fara, M. "Dispositions." (2014). *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/dispositions/>. Accessed April 16, 2017.

Clemente, R. (2013) "Who is Black?" The Huffington Post
http://www.huffingtonpost.com/2013/10/24/rosa-clemente-who-is-black_n_4155793.html
Accessed April 16, 2017.

Cox, D. & Jones, R. (2015) "Most Americans Believe Protests Make the Country Better; Support Decreases Dramatically Among Whites if Protesters are Identified as Black." PRRI.
<https://www.prii.org/research/survey-americans-believe-protests-make-country-better-support-decreases-dramatically-protesters-identified-black/#.VYlfpfIVikp> Accessed April 24, 2017.

Crompton, L. (2003). *Homosexuality & civilization*. Cambridge, Mass.: Belknap Press of Harvard University Press.

- Darwall, S. (ms) "Contempt as an Other-Characterizing, 'Hierarchizing' Attitude."
- Dembroff, R. (2016) "What is Sexual Orientation?" *Philosopher's Imprint* 16 (3).
- Diamant, L. (1995) "Sexual Orientation: Some Historical Perspective." In *The Psychology of sexual orientation, behavior, and identity a handbook*, Louis Diamant and Richard D. McAnulty (eds.). Westport, Conn.: Greenwood Press.
- Diaz-Leon, E. (2015) No title. In "Philosophers on Rachel Dolezal." *The Daily Nous*. June 15. <http://dailynous.com/2015/06/15/philosophers-on-rachel-dolezal/>. Accessed March 16, 2017.
- Diaz-Leon, E. (2016) "Woman as a Politically Significant Term: A Solution to the Puzzle." *Hypatia* 31 (2): 245-258.
- Emens, E. (2014). "Compulsory Sexuality." *Stanford Law Review* 66:(303).
- Epstein, B. (2017) "What are Social Groups? Their Metaphysics and How to Classify Them." *Synthese*, published online April 6, 2017.
- Foucault, M. (1980). *The History of Sexuality. Volume One: An Introduction*. Vintage Books.
- Fricker, M. (2007) *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press.
- Gebauer, J., Baumeister, R., Sedikides, C. and Neberich, W. (2014). "Satisfaction-Adaptation Principles in Sexual Desire: Exploring Gender Differences Across the Life Span." *Social Psychological and Personality Science* 5:176-184.
- Glasgow, J. (2009). *A Theory of Race*. Routledge.
- Greenesmith, H. (2010). "Drawing Bisexuality Back into the Picture: How Bisexuality Fits Into the LGBT Strategy Ten Years After Bisexual Erasure." *Cardozo Journal of Law and Gender* 17: 65–80.
- Hacking, I. (2000) *The Social Construction of What?* Harvard University Press.
- Halperin, D. (1990). *One Hundred Years of Homosexuality: and other essays on Greek love*. Routledge.
- Halperin, D. (2002). *How to do the history of homosexuality*. Chicago: University of Chicago Press.

Hartsock, Nancy, 1983. "The Feminist Standpoint: Developing the Ground for a Specifically Feminist Historical Materialism," *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and the Philosophy of Science*, Sandra Harding and Merrill Hintikka (eds.), Dordrecht: D. Reidel, 283–310.

Haslanger, S. (2000). "Gender and race: (What) are they? (What) do we want them to be?" *Noûs* 34 (1):31–55.

Haslanger, S. (1995). "Ontology and Social Construction." *Philosophical Topics* 23 (2):95-125.

Haslanger, S. (2017) "Racism, Ideology, and Social Movements." *Res Philosophica*, 94 (1):1-22.

Haslanger, S. (2012) *Resisting Reality: Social Construction and Social Critique*. Oxford University Press.

Haslanger, S. and Ásta (2011). "Feminist Metaphysics", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/entries/feminism-metaphysics/>. Accessed April 16, 2017.

Human Rights Campaign. (2014). *Sexual Orientation and Gender Identity: Terminology and Definitions*. URL = <http://www.hrc.org/resources/entry/sexual-orientation-and-gender-identity-terminology-and-definitions>. Accessed April 16, 2017.

James, O. (2015) "What Rachel Dolezal Doesn't Understand." *The Washington Post*. https://www.washingtonpost.com/posteverything/wp/2015/06/12/what-rachel-dolezal-doesnt-understand-being-black-is-about-more-than-just-how-you-look/?utm_term=.25256d29d552 Accessed April 16, 2017.

Jenkins, K. (2015). "Amelioration and Inclusion: Gender Identity and the Concept of Woman." *Ethics* 126 (2):394-421.

Kagan, S. (2011). "Do I Make a Difference?" *Philosophy and Public Affairs* 29 (2):105-141

Kapusta, S. (2016) "Misgendering and Its Moral Contestability." *Hypatia: A Journal of Feminist Philosophy* 31 (3):512-519.

Krauss, Robert M.; Chiu, Chi-Yue (1998) "Language and social behavior." In *The handbook of social psychology*. Gilbert, Daniel T. (Ed); Fiske, Susan T. (Ed); Lindzey, Gardner (Ed). Vols. 1-2, 4th ed., (pp. 41-88). New York, NY.

Lamont, M. & Mointer, V. (2002) 'The Study of Boundaries in the Social Sciences'. *Annual Review of Sociology*. 28:167–95.

Liptak, A. (2015). "Gender Bias Issue Could Tip Chief Justice Roberts Into Ruling for Gay Marriage." *The New York Times*. <http://nyti.ms/1OD08RL>

- Mallon, R. (2016) *The Construction of Human Kinds*. Oxford University Press.
- Mallon, R. (2014) "Passing, Traveling and Reality: Social Constructionism and the Metaphysics of Race." *Nous* 38:4, pp. 644-673.
- McKittrick, J. (2015). "A dispositional account of gender." *Philosophical Studies* 172 (10):2575-2589.
- Mikkola, M. (2012). "Feminist Perspectives on Sex and Gender", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2012/entries/feminism-gender/>. Accessed April 16, 2017.
- Mikkola, M. (2011). "Ontological Commitments, Sex, and Gender." In *Feminist Metaphysics*. Ed. Charlotte Witt. Springer.
- Money, J. & Lamacz, M. (1984). "Gynemimesis and gynemimetophilia: Individual and cross cultural manifestations of a gender-coping strategy hitherto unnamed." *Comprehensive Psychiatry* 25 (4): 392–403.
- Mortimer, D. (2016) "Can Straight People be Queer?" *Vice*.
https://www.vice.com/en_us/article/can-straight-people-be-queer-435 Accessed April 16, 2017.
- Nazworth, N. (2015) "What Liberals Think About Transgenderism." *The Christian Post*.
<http://www.christianpost.com/news/analysis-what-liberals-think-about-transgenderism-140028/page1.html> Accessed April 16, 2017.
- North Carolina. Legislature. Assembly. House Bill DRH40005-TC-1B. Second Extra Session 2016. (March 23, 2016). *North Carolina State Assembly*. Web. Accessed April 26, 2017.
- On Century Avenue. (2015) "Harm is harm, hello: An interview with Catherine McKinnon." <http://oncenturyavenue.org/2015/03/harm-is-harm-hello/> Accessed April 26, 2017.
- Pickett, B. (2011). "Homosexuality", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/homosexuality/> Accessed April 16, 2017.
- Plunkett, D. (2015) "Which Concepts Should We Use?: Metalinguistic Negotiations and the Methodology of Philosophy." *Inquiry* 58 (7-8):828-874.
- Plunkett, D. & Burgess, A. (2013a) "Conceptual Ethics I." *Philosophy Compass* 8 (12):1091-1101.
- Plunkett, D. & Burgess, A. (2013b). "Conceptual Ethics II." *Philosophy Compass* 8 (12):1102-1110.

- Ramirez, G. (2016) "On Realizing I am Black" *Noir Image*.
<https://thenoirimage.wordpress.com/2016/02/09/on-realizing-i-am-black-by-gabriel-ramirez>.
 Accessed April 16, 2017.
- Ray, N. (2006). "Lesbian, gay, bisexual and transgender: An epidemic of homelessness." New York: National Gay and Lesbian Task Force Policy Institute and the National Coalition for the Homeless. URL = <http://www.thetaskforce.org/downloads/HomelessYouth.pdf> Accessed April 16, 2017.
- Reumann, M. (2005). *American Sexual Character: Sex, Gender, and National Identity in the Kinsey Reports*. University of California Press.
- Robinson, G. (2012). *God believes in love: straight talk about gay marriage*. New York: Knopf.
- Saul, J. (2013). "Politically Significant Terms and Philosophy of Language". In S. Crasnow and A. Superson (eds), *Out from the Shadows: Analytical Feminist Contributions to Traditional Philosophy*. Oxford University Press.
- Schaffer, J. (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. pp. 347-383.
- Schroeder, T. (2014). "Desire". *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/desire>. Accessed April 16, 2017.
- Sedgwick, E. (2008). *Epistemology of the Closet*. Berkeley: University of California Press.
- Shelby, T. (2003) "Ideology, Racism, and Critical Social Theory." *Philosophical Forum* 34 (2):153–188.
- Sider, T. (2011). *Writing the Book of the World*. Oxford University Press.
- Steensma, Thomas, Baudewijntje Kreukels, Annelou de Vries, Peggy Cohen-Kettenis. (2013). "Gender identity development in adolescence", *Hormones and Behavior*, Volume 64, Issue 2: 288-297.
- Stein, E. (1999). *The Mismeasure of Desire: The Science, Theory, and Ethics of Sexual Orientation*. Oxford University Press.
- Stevenson, N. (2015) "Autism Doesn't Have to be Viewed as a Disability or Disorder" The Guardian. <https://www.theguardian.com/science/blog/2015/jul/16/autism-doesnt-have-to-be-viewed-as-a-disability-or-disorder>. Accessed April 16, 2017.
- Swanson, E. (ms) "Slurs and Ideologies."

Talusan, M. (2015) "There is no comparison between transgender people and Rachel Dolezal." *The Guardian*. <https://www.theguardian.com/commentisfree/2015/jun/12/comparison-transgender-people-rachel-dolezal> Accessed April 20, 2017.

Thomasson, A. (2017) "Metaphysical Disputes and Metalinguistic Negotiation." *Analytic Philosophy* 58 (1):1-28 (2017).

Urquhart, E. (2015) "It Isn't Crazy to Compare Rachel Dolezal with Caitlyn Jenner." *Slate* http://www.slate.com/blogs/outward/2015/06/15/rachel_dolezal_caitlyn_jenner_how_transgender_is_different_from_transracial.html Accessed April 20, 2017.

Weinberg, M. & Williams, C. (2010). "Men Sexually Interested in Transwomen (MSTW): Gendered Embodiment and the Construction of Sexual Desire". *Journal of Sex Research* 47 (4): 374–383.

Witt, C. (2011a). *The Metaphysics of Gender*. Oxford University Press.

Witt, C. (2011b). "What Is Gender Essentialism." In *Feminist metaphysics explorations in the ontology of sex, gender and the self*, Charlotte Witt (ed.). Dordrecht: Springer.

Wittig, M. (1992). "One Is Not Born a Woman." In *The straight mind and other essays*. Boston: Beacon Press.

Young, I. (1990) "The Five Faces of Oppression." In *Justice and the Politics of Difference*. Princeton University Press.