Summer 7-1-2022

Gender Norms and Gendered Traits

Rowan Bell
Syracuse University

ABSTRACT

Gender oppression is sustained in part through enforcement of and compliance with gender norms. Understanding how they work is therefore salient to the goal of gender liberation.

According to the *category-based view,* which is common in analytic feminist philosophy, gender norms are assigned to individuals based on their assigned gender category, such as *woman* or *man*. I argue that this is insufficient, because it ignores the experiences of those who are marginalized or excluded from those categories.

On a category-based view, individual responsiveness to gender norms will track gender category assignment; only individuals assigned the category *woman* will be responsive to and evaluated under feminine norms, and so forth. However, many trans and GNC people experience themselves as responsive to norms that were not assigned to them. For example, a person who hasn't been assigned the category *woman* may nevertheless feel that they *ought* to follow feminine norms. This *cross-category norm responsiveness* has considerable power over choice and behavior; but a category-based view does not explain this it. Moreover, many marginalized people are actively excluded from dominant, white-centric, cisnormative gender categories. However, the norms associated with these categories are nevertheless enforced on marginalized in particularly brutal ways, in part *because* they are not afforded full gender category membership. In neither of these cases does the category-based view in fact capture the way gender norms are enforced or experienced as normative.

I argue that gender norms primarily operate by attaching to *traits.* Traits are descriptive features of individuals or groups, which are coded as *masculine* or *feminine* in a context. Dominant social contexts mandate the coherence of a set of traits; if some individual or group exemplifies a feminine-coded trait, they are thereby expected to exemplify the rest of the set.

However, individuals who are disposed to express a trait which does not "match" other traits they are observed to express can thereby feel responsive to the norms associated with the trait in question, rather than with their assigned category. For example, a person who is disposed to express a masculine trait may therefore feel responsive to norms of masculinity. Similarly, those who are excluded from a gender category can nevertheless be punished for non-coherence. Gender norms may be enforced on marginalized based on the gendered traits they do express, even as they are excluded from category membership. I proceed over the course of four chapters.

Chapter 1 examines the literature on gender norms in analytic feminist metaphysics, and distinguishes between a commitment to *ascriptivism* and a commitment to the *category-based view* (CBV), which are often conflated in this tradition. I articulate the advantages of *ascriptivism,* but suggest that the CBV will face serious problems.

Chapter 2 motivates two major objections to the CBV, as outlined above; the *responsiveness objection* and the *evaluability objection.* I argue that the CBV fails to explain gender's normative power because it centers those with significant privilege.

Chapter 3 explains and defends my positive view, *traits ascriptivism* (TA). TA holds that gender norms are assigned on the basis of *traits*, rather than gender categories. TA has many advantages over existing views in the metaphysics of gender, while maintaining their valuable core commitments and insights. I explain how the view captures the important desiderata for a positive view of gender norms enforcement, and respond to objections.

Chapter 4 explores the normative power of *authenticity* over behavior. Gender norms are often experienced as authentic, despite being socially assigned and morally bad. Drawing on metaethical notions of normative authority and an existentialist tradition of socially embedded authenticity, I explain how gender norms can be authentic and therefore action-guiding.

GENDER NORMS AND GENDERED TRAITS

by

Rowan Bell

B.S., East Tennessee State University, 2014

Dissertation
Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Philosophy.

Syracuse University
July 1, 2022

ACKNOWLEDGMENTS

Haslanger for countless hours, a few moments of which you may actually have enjoyed; for supporting me even when the path seemed uncertain; for never letting me, even for a moment, fear that I would have to do this alone. I make a living out of words, but when I try to describe what you mean to me, they fall utterly short. I love you.

Finally, of course, thanks to the people I have undoubtedly forgotten to mention. I have not forgotten your contributions.

TABLE OF CONTENTS

> I am learning that adults react the same way to my interest in makeup as they do
> to my interest in matches and lighters. As if maybe, by being what I am,
> I might burn down something very important to them.
> Something that makes their life more comfortable and easy.
> Jennifer Coates
>
> If the body is always a sign being read, then not communicating is impossible.
> Riki Anne Wilchins

INTRODUCTION

Gender norms shape our lives in ways large and small. They tell us what kind of work we should do; who we should have relationships with, and of what kind; how and when we should speak, and to whom; even how we should comport, move, understand, present, decorate, and generally inhabit our own bodies. Very often, these prescriptions are deeply harmful, unequal, and oppressive. They afford power and privilege to some, while enabling the subordination and oppression of others. They create patterns of behavior which sustain unjust systems of power, and then render themselves invisible so that the behavior appears natural and unprompted. In brief, they differentially restrict us based on arbitrary features, thus creating an unequal social world where some people and groups are unjustly constrained and punished while others are unjustly enabled.

It is under these terms that many domains of gender norms, particularly the hegemonic norms enforced in dominant colonial contexts, have been the target of feminist critique. This is true of most or all features of gender. However, what is distinctive about gender norms in particular is that they seem to have a distinctive *normative power* over our behavior. They don't just tell us what the world is like, or what we are like, or even what we are permitted or able to do. They try to tell us what we *should* do. They affect our sense of how the world ought to be, and thus guide our behavior such that we aim to make it so. That is, gender norms have the power to affect our *normative deliberations*. What is most disturbing about this is that they seem

to be able to do this irrespective of our preferences. Rejecting the legitimacy of gender norms doesn't always mean shaking off the power they exert over our choices. If we want to understand how systems of oppressive power are perpetuated, even by those who are aware of and opposed to their normative structures, we need to understand how this works.

In her forthcoming book, *The Great Gender Divergence,* Alice Evans investigates the question of why, in the 10,000 years since the agricultural revolution, a majority of human societies have become patriarchal--in the literal sense that they have been ruled and organized by males (Evans forthcoming, Conover & Evans 2022). The answer is, unsurprisingly, complex. Evans' account describes a myriad of interlocking economic, social, and biological factors that, on the whole, tend to create a feedback loop where patriarchies accrue the wealth and privilege necessary to sustain themselves, while matriarchies or egalitarian societies tend to transform into, or be overthrown by, patriarchies. Moreover, patriarchies tend to involve *control of females,* while matriarchies tend to be more egalitarian. According to Evans, this is a major reason why patriarchies don't often transform into matriarchies; men control women, but women don't control men.

Evans describes the mechanisms which perpetuate this as surprisingly consistent. For example, as societies become stratified in terms of wealth--a phenomenon which tends to follow from the transition to agrarian life--it becomes important for the wealthy to preserve their lineage, in order to pass that wealth down to their children. To do this, male leaders of households need to be certain that their children are, in fact, *their children.* No such concern arises when females are the leaders of households. Powerful males therefore have motivations to control the women in their lives, in order to ensure that these women do not bear the children of other men. There are many ways of doing this. The men might, for example, control these

women *physically,* by enforcing limits on their movement. They might also control them *socially,* by forbidding them from interacting with others.

However, there are more effective and insidious methods of control. What stands out most to me about Evans' account is her description of the way that societies reproduce themselves by creating apparently self-sustaining patterns of behavior. That is, successful societies persist because they manage to consistently *motivate* groups and individuals to act in ways that sustain the social order. According to Evans, those ways re-occur in predictable patterns across societies. As societies become more patriarchal, they also tend to develop *normative structures* that constrain women's lives. It becomes, not just forbidden, but *wrong* for women to work or to socialize outside of the home. It becomes just as wrong for men to allow them to roam free. A good woman is one who stays home, out of public life, and away from men outside her family; a good man is one whose women follow these rules. These normative structures are often enshrined in complex social mores, such as honor codes or religious laws. For example, Evans points to the fact that over a third of the rules in Zoroastrianism, an early monotheistic religion, are aimed at controlling the behavior of women (Conover & Evans 2022).

Such normative structures are highly effective in maintaining structures of power. Moreover, they can constrain behavior long after the structures are gone. As societies become less patriarchal, the associated norms do not disappear. On the contrary; as the more coercive and obvious power systems fade, we see the norms evolving to compensate, often becoming more subtle and misleading in order to be effective (such as racist and sexist norms that pass themselves off as *colorblind* or *gender-neutral*; see Alexander 2010, Manne 2017).

But why do societies consistently develop *normative structures* to shape behavior? That is, what is it about *norms* that gives them so much power to perpetuate the social order? Evans'

answer is twofold: "People either internalize these ideas, [e.g.] that there's eternal damnation for stepping out of line; or, it's not that they really believe it--they might be privately critical--but the costs of noncompliance are so huge that you have to comply" (Conover & Evans 2022). That is, coercive social norms shape behavior in two ways. Either individuals *internalize* the norms and come to follow them as if they are *really* consequential, or they follow the norms because they fear *punishment* for failing to do so.

These two ways of reacting to a norm map nicely onto a phenomenon described by Charlotte Witt in her 2011 monograph *The Metaphysics of Gender.* Witt raises the following question: "Why are individuals *responsive to* and *evaluable under* social norms?" (2011, 42, emphasis added). That is, given the structural nature of social norms such as gender norms, and, crucially, the fact that we often *reject* or *disprefer* these norms, why do we feel we *should* follow them, and why do we evaluate others accordingly?

I argue that this is a crucial question for understanding how gender oppression sustains itself. Put broadly, norms operate as a functional bridge between the macroscopic structures of oppressive power that shape our world, and the practical agency of those who participate in, contribute to, and maintain those structures. Of course, oppressions such as racism, sexism, and coloniality are not solely, or even primarily, caused and perpetuated by individuals. (According to some poststructuralists, the opposite is actually true.) But I think we cannot understand this system without understanding how norms shape agency--with or without the knowledge, preference, or consent of agents. Iris Marion Young writes:

> Oppression in this sense is structural, rather than the result of a few people's choices or policies. Its causes are embedded in unquestioned *norms, habits,* and symbols, in the assumptions underlying institutional rules and the collective consequences of following those rules.... In this extended structural sense oppression refers to the vast and deep injustices some groups suffer as a consequence of *often unconscious assumptions and*

reactions of well-meaning people in ordinary interactions, media and cultural stereotypes, and structural features of bureaucratic hierarchies and market mechanisms— in short, *the normal processes of everyday life*. (Young 1990, 41, emphasis added)

Broadly speaking, I seek to understand what makes certain processes "normal" or "ordinary" while being, at the same time, constructed, institutional, "vast and deep." I want to know how these agential bridges are built.

This is, in some ways, the "biggest picture" version of the question. In this dissertation, I will address what I take to be a relatively narrow, but crucial, part of it. If we want to know how gender norms shape agency, we need to know how they *attach* to us, and how that attachment creates reasons in us. Recall Witt's question. Why do we respond to the call of gender norms, and why do others judge us according to them? Specifically, why does this happen *differently* to different people, in a way that generates *reasons* for many of those people, and often enables oppression? I take this to be a useful point of entry.

Witt has an answer for this. She argues that individuals are assigned to *social positions,* which come with collections of norms. That is, according to Witt, what explains gender norms' differential power over us is the *gender categories* that we occupy. This is in agreement with many other philosophers in recent analytic feminist philosophy, in particular Haslanger (2012), Ásta (2018), and Dembroff (2018b, 2020). On this family of views, gender categories both *ground* gender norms--that is, the categories are the reason the norms exist--and they explain how those gender norms are *assigned* to us.

However, I think that this answer, although ontologically tidy, is too narrow. It primarily explains the experiences of those who are already centered in hegemonic gender categories--i.e. those who have significant privilege. But many marginalized people--in particular transgender (trans) and gender-nonconforming (GNC) people and people of color--often have very different

experiences of gender's normative power. That is, those who are marginalized within, or altogether excluded from, dominant gender categorization often experience gender norms as exerting force over us in ways that do not align with gender categorization. For example, many trans and GNC people experience gender norms which have ostensibly *not* been assigned to us as exerting force over our behavior and giving us reasons to act. Similarly, many people who are excluded from particular categories on the basis of racialization, colonization, or general gender unintelligibility, are nevertheless subjected to the associated gender norms in particularly brutal ways. Moreover, these harms are not visited merely at the level of the individual, but are enforced against entire cultures, societies, and ways of life.

To ignore these experiences is a significant theoretical oversight. If we want to understand how gender norms exert normative power over the world, we need to consult those who have the best epistemic perspective on this power: those who have experienced it from multiple intersecting angles, and who are frequently in its crosshairs. This project therefore proceeds from the assumption that the experiences of people marginalized along dimensions of gender—who will, on an intersectional view, turn out to be far more than just those classed as *women*—ought to be centered in a study of gender norms.

How, then, should we understand gender norm assignment? I will argue that we should adopt a *traits-based view*. That is, I will argue that gender norms are assigned, and thus exert force, on the basis of the *traits* that some entity is perceived to have. "Entity" here is understood broadly; an individual can be an entity, but so can a group, a society, or an entire way of life. A traits-based view allows for a much more modular, malleable, context-dependent account of gender norm assignment, one that can capture the advantages of a category-based view without falling prey to its biases. With a traits-based view in hand, we can understand how gender norms

6

apply differently on the basis of various *features*--including those associated with sexed embodiment, such as genital structure and secondary sex characteristics, but *also* including other embodied features (such as racialized features), features of behavior, character traits, appearance and dress, and so on. This approach, I argue, can explain the normative experiences of marginalized people while also capturing what is attractive about a category-based view.

Before I say more about what the project is, let me say a little about what it is *not*. First, it is not an empirical inquiry about why individuals behave in ways that conform to gender norms. Sociologists have often described *internalization* in purely behaviorist terms, without reference to the normative deliberations, reasons, or other "internal" states of the individual. For example, sociologist John Finley Scott (1971) argues that norms become internalized as a response to social sanctions. This has come to be known as the *internalization hypothesis,* which is widely influential in the social sciences (Sripada & Stich 2006). According to Scott, *internalization* is best described in purely behavioral terms: "the actor has been sufficiently conditioned by sanctions that his behavior conforms to the norm at a spatial or temporal remove from sanctions" (Scott 1971, 92). Scott explicitly argues against the investigation of so-called "subjective aspects" of normative commitment, arguing that phenomenological states such as *obligation* are largely unknowable except by introspection (21). Since, according to Scott, introspection is unreliable at best, producing data that is "bewildering, contradictory, and often highly implausible" (25), Scott argues that the empirical scientist ought to focus only on that which is observable and quantifiable: *behavior.*

But this project is not primarily empirical. I do not aim, nor am I qualified, to give a scientific study of the causes of norm-compliant behavior. Rather, I am interested in what I am calling the *normative structures* of the world. I want to know how practical agency shapes and is

shaped by gender norms, and how agents (individual and collective) come to have and act on gendered *reasons* that weigh heavy in their normative deliberations.<*> There are, of course, a wealth of empirical issues in the vicinity: How do individuals come to be *inclined* towards gendered behavior? How does behavior become gendered in the first place? What causes individuals to consider different gender norms as candidates or non-candidates for possible reasons? How do particular norms affect the way individuals *behave*? These are all fascinating questions that I am neither qualified nor attempting to answer. Rather, I am interested in what I take to be a separate and distinctively philosophical question: *How does gender shape the landscape of agency in ways that justify and enable oppression?*

Second, I do not take myself to be giving an account of "what gender is." Similarly, I am not interested in questions about what *gender categories* are and what it takes to count as a member of a category, what gender *identity* is and what it takes to have one, and so on. I want to give these questions as wide a berth as possible. If one holds a view of gender on which it is partially or wholly constituted by gender norms, elements of my view will certainly have implications for what gender is. But I don't address these issues. On the contrary, following Mikkola (2016) and Antony (2020), I suggest that a focus on these kinds of questions has dominated gender metaphysics for too long, and distracted us from our more pressing goals. My aim is to focus on those features of the world which have *traveled under the banner of* gender, and to study the ways in which these features have enabled and justified various oppressions and harms. Again, following Mikkola (2016), I argue that I don't need a particular "metaphysics of gender" to do this. I only need to be able to *talk about it* in an intelligible and meaningful way-- which I am already doing.

Relatedly, and perhaps more radically, I do not try to give a clear answer to the question of what makes a gender norm a *gender* norm. I understand that gender norms are difficult to cleanly distinguish from other normative domains, such as etiquette or aesthetics. My aim in this project is not to defend gender norms as a distinctive and univocal domain. While I maintain that I want to avoid giving a metaphysics of gender, I will lay bare here at least one commitment. It is received wisdom that gender oppression is co-constructed with other social oppressions. I understand this as a strong metaphysical claim. In particular, the gender system which dominates the modern colonial world is inseparable from the *racial* and *economic* systems; it is, in the words of bell hooks, a "white supremacist capitalist patriarchy" (2000, 4). This is a phenomenon which is, I think, significantly under-theorized in analytic gender metaphysics. I understand this co-construction as entailing at least a partial *co-constitution*. In a very real sense, colonial, gender, racial, economic, and other oppressions *constitute* one another, and cannot be meaningfully disentangled without erasing those who are most targeted by this co-construction. In Chapter 4 I suggest that dominant domains of gender norms are likely unified by particular oppressive aims, such as *maintaining the colonial social order.* This is compatible with the thought that they are not wholly ontologically distinct from other domains. I want to leave it open whether such ontological distinction is possible for everything that we want to call "gender norms." However, I do not think it is necessary to do that here. My aim is to understand the dimensions of normative power that *have traveled under the banner* of gender, in order to pursue the political aims of ameliorating the oppressions and harms they enable. Again, following both Mikkola (2016) and Antony (2020), I hold that we don't need to complete the ontological project before proceeding with the political one. I'll say a bit more about this in Chapter 3.

Third, I am not trying to answer questions about *what agents are*. Certain ways of engaging with the questions I've raised thus far might be interpreted this way. For example, Foucault (1983) argues that human beings are "made subjects" by the normative structures of power which bound the social world. "Subject" here has two senses: first, it means a *conscious agent* (i.e. a subject as opposed to an *object*); second, it means something that is *subjected to* a greater force (i.e. the subject of a monarchy). For Foucault, the two are intertwined; subjects in the first sense exist because they are "subjected" in the second sense--they are made subjects. On this picture, then, normative power *creates* the agent. Similarly, specific to the *gender* case, Judith Butler (1990, 1993) argues that subjects are in part *constituted by* the linguistic and cultural gender norms that bound their behavior. For Butler, there is no stable gender identity *behind* or *before* the boundaries of gendered language and convention through which the subject acts and is interpreted. The performance of gender, which is always necessarily constrained by available language and normative convention, *constitutes* the subject. This is a radically anti-essentialist gender metaphysics--a picture on which it's not just that gender is socially constructed, but *bodies* and *individuals themselves* are constructed by gender.

My project, however, does not commit itself to a particular metaphysics of subjecthood. I take my claims here to be consistent with these approaches, as well as potential others. If agents are understood as analogous to "subjects" in the sense articulated by Foucault or Butler, then my central questions will certainly bear on how subjects come to be. But this is not specifically intended as an upshot of my view. Rather, I am interested in the mechanisms by which normative power shapes agency and choice itself. The question of where (or whether) the *subject* comes in is, I think, orthogonal. (To take a radical example, my view seems consistent with a view on which there is *no such things* as individual agents, but only *agency* arranged agent-wise, as it

were.) I do at times speak as if there is a clear line between agent and world; for example, I sometimes ask how norms "get into" the agent. However, I take this to be a handy way of describing the *apparent* distinction between subject and world as it is experienced by that (apparent) subject, rather than a claim that the distinction is really substantive. In chapter 4, I will discuss what it takes for a subject to construct themselves *authentically*; but again, my aim is not to discuss authentic *subjecthood* in itself, but rather how practical choice is guided by the normative demands of *authenticity.*

Fourth, and finally, I am not engaging in *ideal theory*. Broadly speaking, I want to study normativity, not in its robust, abstracted, and ideal form, but in terms of the actual, constructed, context-dependent, and imperfect normative standards and norms that in fact shape the agential landscape. The literature on agency in metaethics and philosophy of action has largely concerned itself with questions about relatively idealized normative standards, such as morality, prudence, or rationality. These things, it is sometimes thought, give us *real* reasons to act, in a way that is (explicitly or implicitly) contrasted with the merely apparent reasons given by other kinds of standards, such as *etiquette*. For example, Wodak (2019) argues that authoritative norms are the only *real* norms, and that other standards are "fictionally" normative--they imitate real normativity in the same way that a fake duck imitates a real duck, and therefore only *seem* to give us reasons to act. That is, we might act "as if" etiquette gives us a reason, but it doesn't *really.* What's missing from this discussion, I think, is the fact that many people seem to take ostensibly non-"real" normative standards, such as gender, to be giving them real reasons. But I think some of the tools used to study "traditional" idealized normativity can nevertheless be useful here, as I hope will become apparent.

My overall strategy will be as follows. I will begin by discussing recent trends in the analytic metaphysics of gender on the matter of gender norms. I show that many prominent theorists in this tradition are committed to what I call the *category-based view,* or the view that gender norms apply on the basis of the gender categories that individual agents occupy. I then raise two major objections against the category-based view, both drawn from the experiences of marginalized people: I argue that it cannot explain how transgender (trans) and gender-nonconforming (GNC) people are *responsive to* gender norms, and that it cannot explain how many people and groups (such as colonized people and societies, enslaved people, and gender-nonconforming people) are *evaluable under* gender norms. In place of a category-based view, I argue that we should adopt a *traits-based view*, which can explain these phenomena while also doing the work we want a category-based view to do. Finally, I explore the question of why gender norms that are assigned on the basis of traits seem to be able to give us *reasons to act* that can weigh heavy against other considerations, such as morality or prudence. Let me discuss each of these moves in more detail.

In Chapter 1, I situate my discussion of gender norms within the existing literature in analytic feminist metaphysics. First, I give a sense of what prominent theorists in this tradition, in particular Witt (2011), Haslanger (2012), Ásta (2018), and Dembroff (2018b, 2020), have meant by saying that gender is *socially constructed.* I then draw out what these theorists have said about *gender norms.* I argue that theorists in this tradition have tended to conflate two commitments: *thin ascriptivism,* or the view that gender norms apply to agents without their participation or consent, and the *category-based view,* or the view that gender norms apply to agents on the basis of the gender categories that they occupy. I defend thin ascriptivism on the grounds that it is good fit for social constructionism and other feminist commitments, such as

structuralism. I then raise two major objections to the category-based view, which I discuss only briefly; I will elaborate at length in Chapter 2. I raise them here to argue that there are good reasons to distinguish thin ascriptivism from the category-based view.

In Chapter 2, as promised, I motivate the two major objections to the category-based view discussed briefly in Chapter 1. I begin by returning to Witt's question: Why are individuals *responsive to* and *evaluable under* gender norms? I frame these as two major components of gender's normative power. Accordingly, I raise two objections to the category-based view: the *responsiveness objection* and the *evaluability objection*. First, I argue that the category-based view cannot explain how agents are responsive to gender norms. I identify the phenomenon of *cross-category norm responsiveness*. Many trans and GNC people experience themselves as responsive to norms associated with gender categories that were not assigned to them. For example, many people who have not been assigned the category "woman" nevertheless experience themselves as responsive to feminine norms. Importantly, this often happens when the individual in question also does not understand *themselves* as belonging to the category associated with the norms to which they are responsive. I argue that a category-based view does not explain this. Second, I argue that the category-based view cannot explain how agents, and other social entities, are evaluable under gender norms. Many people who are actively excluded or ejected from dominant gender categories are nevertheless held as evaluable under the associated norms. For example, Lugones (2007, 2010) argues that membership in dominant gender categories of "man" and "woman" is historically reserved for white colonizers; Black and Brown colonized people are excluded from membership in these categories. Similarly, Spillers (1987) argues that Black enslaved bodies are *ungendered* through the legacy of chattel slavery. And Wilchins (1997) points to the ways in which trans and GNC people are treated as

"gendertrash," outside the boundaries of binary classifications. However, the gender norms associated with dominant gender categories are nevertheless enforced on those excluded from dominant categories in particularly brutal ways—in part *because* they are denied gender category membership. More broadly, gender norms do not only function at the level of the individual agent, as a category-based view suggests. For example, as Lugones has argued, macro-level social entities such as groups, societies, and ways of knowing are treated as evaluable under gender norms; this is a key feature of the construction of dominant gender norms through the legacy of coloniality. I conclude that a category-based view focuses on the experiences of those who are already privileged within dominant gender categories. We should therefore dispense with it.

However, a category-based view has one significant advantage: It gives us a clear and parsimonious explanation of the grounds on which gender norms are assigned, one that captures the way in which they apply differently to individuals based on how they are differently situated and interpreted within a gendered social system. Rejecting a category-based view also strips away some of this explanatory richness. We are left without an account of the mechanism of gender norm assignment. We need to provide a view that does not encounter the same problems that face the category-based view, but is still compatible with other feminist commitments, like *social construction*. What might fill this role?

In Chapter 3, I give an answer. I propose that we shift focus from gender *categories* to gendered *traits,* and adopt what I call a *traits-based view* about gender norm assignment. This shift, I contend, can provide a clear account of the mechanism by which gender norms are assigned and enforced, while maintaining the advantages of previous accounts in feminist philosophy, and avoiding the problems faced by a category-based view. I begin by identifying

six desiderata which I take to be crucial for explaining how gender norms attach to features of the world: *responsiveness, evaluability, social construction, nonconsensual assignment, differential assignment and embodiment,* and *oppression.* I explore how interactions between these desiderata can help prepare us to construct a positive view. I then explain and defend a traits-based view. Gender norms, I argue, are assigned to entities in the world based on the *traits* that they are perceived to display, where *traits* are understood as "bare descriptive facts" about entities. Particular traits are gender-coded, typically as either *masculine* or *feminine.* If an entity is perceived to exhibit a particular trait that is gender-coded in their locality, that person is thereby evaluated under the associated gender norm. In contexts where gender is assumed to function as a strict and exclusive binary, gender norms are expected to cluster together; this cluster can be enforced as normative. I suggest that a traits-based view is better suited to explain the way gender norms are experienced as normative, and enforced against, marginalized people, than a category-based view. I discuss the ways in which it meets the six desiderata I have identified, in the process further exploring its commitments. I then respond to three objections against the traits-based view, each from the perspective that the category-based view is better suited to explain gender norm assignment. I conclude that each objection fails, and that we should adopt a traits-based view in place of a category-based view.

In Chapter 4, I address the ways in which certain gender norms can be experienced as providing authoritative or "real" *reasons to act,* even when an individual might be subjected to multiple and conflicting standards. Trans and gender-nonconforming people sometimes say that certain gender norms--often those which conflict with their assigned categories--are *authentic* for them. For example, a trans man might say that abiding by norms of masculinity tracks *who he really is*. Authenticity as a standard is sometimes taken to appeal to an essential, pre-social

"inner self." It is also sometimes understood as a moral notion. Authenticity claims about gender

norms therefore appear inimical to two key commitments I have already articulated: *social*

construction, and *oppression* (which I extend to entail the claim that many domains of gender

norms are both morally and prudentially *bad*.) I argue, however, that that this apparent tension is

illusory. Concordant with existing trans narratives of authenticity, I articulate an existentialist

view that understands authenticity as a socially embedded, constructive project undertaken in a

non-ideal social world, rather than a reflective uncovering of a pre-given, essential self. I then

show that authenticity and morality can come apart; what is authentic for someone need not be

either morally good or good for them. I conclude that the authenticity of gender norms does not

cut against the feminist commitments that I identify. This conclusion enables a theoretical space

that is both respectful of trans experience and critical of dominant gender norms, an important

liberatory goal.

1.0 INTRODUCTION

On a prominent family of views in analytic feminist metaphysics, gender norms are understood as the normative standards which delimit appropriate behavior for individuals who are assigned to a given gender category. The story goes like this: An individual is sorted into a gender category by others in their social context on the basis of some observed or imagined feature or features, generally connected to their *sexed embodiment* (Witt 2011, Haslanger 2012, Ásta 2018). Each of these categories comes pre-loaded with a normative standard for "good" and "bad" behavior relative to the category—typically *masculinity* (for the category "men") or *femininity* (for the category "woman") (Haslanger 2012). Each standard is constituted by a group of interrelated *gender norms.* Those gender norms are then enforced on an individual via sanctions (rewards and punishments). Over time, the individual learns to calibrate their behavior relative to that standard, and become responsive to those norms (Witt 2011). On this picture, gender norms are understood relative to a gender category; their primary function is to guide the behavior of individuals in the relevant category so that it falls in line with the "appropriate" behavior assigned to that category.

A view like this has many advantages. Primarily, it gives a plausible social constructionist explanation of how individuals come to be responsive to gender norms. Another way of explaining this is to appeal to 'natural' or 'innate' gendered features; individuals feel moved to comply with gender norms because those individuals are 'naturally' or 'innately'

feminine or masculine. (For a plausible and trans-inclusive version of gendered innateness, see Serano 2007.) However, on a social constructionist view, gender is not natural or innate, but rather is socially constructed. Elements of gender come to appear natural or innate in order to obscure the mechanisms by which they are constructed, and thus to make gender hierarchy also appear natural or innate (see Haslanger 2012, Ásta 2018). Prominent views in analytic philosophy aim to capture both the socially constructed nature of gender norms and the way in which they affect our choices. Relatedly, they explain why individuals often find themselves to be responsive to and evaluable under gender norms even when they do not like, prefer, or identify with those norms or the associated category. Since people learn gender norms over time through conditioning, they come to follow those norms by habit, and find themselves responding to them even when they would prefer not to (Witt 2011).

I argue that views in this tradition often conflate two commitments; *thin ascriptivism*, or the view that gender norms are socially assigned to individuals without their participation or consent, and the *category-based view*, or the view that gender norms apply to individuals on the basis of the gender category of which they are a member. Thin ascriptivism is an important view for feminist purposes, and I argue that we should keep it. However, the category-based view faces serious objections, in particular those drawn from the experiences of people at the margins of dominant gender categories. I briefly discuss two such objections. *First*, the category-based view does not explain the way that individuals actually experience gender norms as applying to them. For example, many transgender (trans) and gender-nonconforming (GNC) people, experience gender norms which do not match their assigned gender category as exerting normative force over them and guiding their actions. *Second*, the category-based view does not explain the way gender norms are enforced. Often, those who are not included in a particular

gender category are nevertheless treated as evaluable under associated norms. These examples suggest that gender norms do not apply solely on the basis of gender categorization.

My primary aim in this chapter will be twofold. First, I will outline the commitments of prominent existing views in the literature on feminist analytic metaphysics with respect to gender norms. In particular, I will discuss the work of Witt (2011), Haslanger (2012), Ásta (2018), and (to a lesser extent) Dembroff (2018b, 2020). Second, based on these exegetical conclusions, I will argue that this tradition tends to conflate thin ascriptivism with the category-based view. Specifically, each of these views either implicitly or explicitly adopts a version of thin ascriptivism that either defends or presupposes a category-based view. In successive chapters, I will argue at length that a category-based view is an insufficient framework to explain the normative power of gender norms and their role in enabling and justifying oppression. However, thin ascriptivism is an important and valuable commitment for feminist purposes. My goal here, then, is to distinguish between the two commitments.

In 1.1, I identify *social constructionism* about gender norms as a central commitment of feminist theory. I define this term with reference to the literature in feminist analytic philosophy. I then discuss the connection between social constructionism and gender norms, primarily drawing from the work of Witt (2011). I show that Witt's arguments in favor of her view, which she calls *ascriptivism,* do not support the substantive view she proposes; however, they do support a much thinner view, which I accordantly call *thin ascriptivism*. In 1.2, I defend thin ascriptivism as a theoretically and politically powerful commitment for feminist purposes. Concordantly, I show how many prominent analytic feminist philosophers either defend or assume this commitment. In 1.3, I argue that, within the tradition I am discussing, thin ascriptivism is often conflated with what I call the *category-based view*. I briefly raise two

objections to the category-based view, in order to demonstrate that it is problematic; I note that I will defend these objections at length in successive chapters. In 1.4, I conclude with a brief summary and preview of the next step.

1.1. SOCIAL CONSTRUCTIONISM IN FEMINIST ANALYTIC PHILOSOPHY

It is an important baseline commitment of feminist theory that gender is socially constructed. However, "social construction" is used in many different ways, and is often misunderstood, which can cause miscommunication. Of particular relevance here is such miscommunication between feminist theory and trans theory (Salamon 2010). Among the clearest and most theoretically elaborate articulations of the social construction of gender comes from the "debunking" tradition in analytic feminist philosophy. In this section, I will articulate some of the core commitments of this family of views. This literature has typically treated gender categories as theoretically primary, and discussions of other gender phenomena, such as gender norms, have been secondary. Since gender norms are the target of this inquiry, I will investigate the relationship between gender categories and gender norms as it is theorized in the target tradition. I examine prominent views in analytic feminist metaphysics, as defended by Witt (2011), Haslanger (2012), Ásta (2018), and Dembroff (2018b, 2020). This family of views shares three commitments about gender norms: (1) social constructionism, (2) thin ascriptivism, and (3) the category-based view. (1) is taken to lead to (2), and (2) is conflated with (3). (1) is well-established and plausible; I touch on it only briefly here, primarily to articulate its core commitments and defend its relationship to (2). In successive sections, I will defend the plausibility of (2) and argue that it can come apart from (3), which is less plausible.

Broadly, social constructionism in feminist philosophy holds that gender is not a natural or fixed part of human life. Rather, gender arises out of contingent historical practices of assigning social significance to certain features. Those features themselves may be natural or fixed parts of human life, such as reproductive capacities (Witt 2011, 27-29; Ásta 2018, 3); however, the associated social significance is contingent, and varies across cultures and contexts. Over time, socially constructed gender phenomena take on a life of their own. A gender category's historical roots in, say, a particular role in sexual reproduction might in principle bear remarkably little resemblance to its modern iterations (Haslanger 2012, 44; Briggs & George forthcoming). Social constructionism is positioned in opposition to *biological essentialism* about gender. According to biological essentialism, gender categories have an intrinsic essence determined by biological features such as chromosomes and hormones. On this view, most phenotypic gendered features, such as appearance and behavior, are causally determined by this intrinsic essence (Bach 2012, 233). The social constructionist argues that biological essentialists "mistake what is social and variable for what is natural and fixed" (Witt 2011, 7-8); that is, biological essentialism mistakes the contingent social significance of certain human features for an inevitable result of natural features, and thus mistakes gender for an inevitable, natural, and thus "normal" part of human existence (Ásta 2018, 46).

Feminist social constructionism is often engaged in the anti-biological-essentialist *debunking project* about gender. A debunking project undertakes to expose hidden mechanisms of social construction in practices that are widely believed to be naturalistic (Haslanger 2012, Ásta 2018). Debunking is deployed against social phenomena which are constructed and unjust, but which, in masquerading as natural, falsely appear natural and thus inevitable (Haslanger 2012, 127). Social constructionist feminist philosophers dating back to Beauvoir (1948) have

made the normative claim that gender is unjust, coupled with the descriptive claim that it is socially constructed and contingent. This contingency is sometimes understood to open up possibilities for gender liberation; what is constructed and contingent may be resisted where what is natural and inevitable may not (Haslanger 2012, 132; Ásta 2018, 56).

Gender norms are the spoken and unspoken standards of behavior, appearance, and comportment which apply to those who exist in gendered social space in virtue of how they are positioned and interpreted in that gendered social space. That is, gender norms are social norms which vary according to who you are and how others understand you within a system of gendering. Moreover, because gender norms are created and enforced by contingent social practices, they are culturally specific and contextual; the norms for a person in one context are different than they would be for that same person in another context, even if we hold fixed their body, behavior, clothing, and so forth. In dominant colonial gendering practices, gender norms typically come in two varieties: *masculinity,* or the standards typically associated with maleness and men, and *femininity,* or the standards typically associated with femaleness and women. In what follows, I discuss how analytic philosophers have theorized the nature of this association between category and normative standard.

Ascriptivism, as defined by Witt (2011), is a metaphysical account of how social normativity works. According to Witt's view, social norms apply to a person because of the social position the person occupies. For example, the norms of femininity apply to a person because she is socially positioned as a woman (Witt 2011, 44-45). On this view, a person is positioned as a woman because she is socially *recognized* as a woman. "Social recognition" here is "a complex, holistic status comprised of both public, institutional recognition and interpersonal acknowledgment" (45). If someone is socially recognized as a woman, she is then "ascribed" the

corresponding normative role and evaluated according to its norms. We can define Witt's ascriptivism about gender norms as follows:

> **Witt's Ascriptivism:** Some gender norms apply to subject S because S is socially recognized as belonging to the gendered social position G, rather than because of S's voluntary self-identification, choice, or preference.

According to Witt, ascriptivism explains why people feel the "normative pull" of social norms; that is, it explains why individuals are psychologically *responsive to* particular social norms, such that they calibrate their behavior accordingly. According to Witt, a person who has occupied the social position of "woman" for a significant amount of time will become used to responding as if she were under evaluation as a woman. That is, she will be habituated, in an Aristotelean sense, to respond to feminine norms, and will therefore experience those norms as relevant to her decisions—even if she rejects those norms and refuses to follow them (Witt 47). Witt's position here mirrors the *internalization hypothesis* in social science, where individuals come to behave according to social norms as a result of social learning and training (Scott 1971, Sripada & Stich 2007). Witt also notes that because of habituation, most people who are ascribed a particular gendered social position also identify with it. Despite this correlation, Witt argues that we shouldn't be misled into thinking that this identification grounds gender's normative power. The ascription of a gendered social position is what "brings [a person] under its normative umbrella"—that is, what leads to their being evaluable by others under the norms associated with that position, as well as feeling the pull of its norms (44).

Witt contrasts ascriptivism with *voluntarism* about social norms. On voluntarism, according to Witt, "the normative pull originates in the agent's decision to adopt a given [social position], thereby accepting that her behavior is subject to certain norms or reasons for acting" (43). On this view, social norms apply to someone because she prefers, endorses, or identifies

with those norms (Witt 2019, 5). The norms are, in this sense, *voluntary*; they apply based on the individual's choice. Witt think this is clearly wrong, because it cannot explain why people are evaluated according to and moved by norms with which they do not identify. For example, if a mother ceases to identify as a mother, that may not change whether she feels that the norms of motherhood are relevant to her conduct, and it certainly won't change whether others hold her accountable to them (Witt 2011, 44). Moreover, we are evaluated under and responsive to most social norms, including gender norms, long before we are capable of identifying with them (45). According to Witt, ascriptivism can capture these facts and voluntarism can't.

Witt attributes the voluntarist view primarily to Korsgaard (1996, 2009). On Korsgaard's voluntarism, norms can genuinely guide us only if and because we willingly adopt a practical identity which requires us to follow them; the norms of femininity will bind one's behavior *only if* one voluntarily adopts or takes up that practical identity of "woman". But I think Witt's critique doesn't capture Korsgaard's aim. Korsgaard is concerned with *robust* or *authoritative* normativity; that is, she is interested in the genuinely authoritative reasons we have for acting. She argues that adoption of a practical identity is what determines whether a person is genuinely *bound by* a norm—whether one *really* ought, in a deep sense, to follow it. Witt, by contrast, is interested in whether a person is "responsive to and evaluable under" a norm (42). This is relatively thin, normatively speaking; it only captures the way norms operate at the social and psychological level, and not any real normative authority they might have. Korsgaard's target concept is therefore more substantive than Witt makes it out to be.

However, this divergence can help us understand Witt's view more clearly. Witt is not invested in understanding robust normativity; that is, she isn't describing gender norms as a part of what anyone *really* ought to do, in the sense that we *really* ought to save a child from

drowning, or believe things for which we have a preponderance of evidence. Rather, she is describing the non-ideal norms that in fact operate in the social world; what people are in fact expected to do, and the expectations to which they in fact respond. Our metanormative discussions about this can be purely descriptive. That is, we can theorize about these norms without ever discussing substantive normative facts about what the people involved *really* ought to do. Ascriptivism is a view about how, descriptively speaking, gender as a general social phenomenon can exert power over people's behavior, not about whether that power is normatively robust. I discuss robust gender norms in Chapter 4.

Witt sets up *voluntarism* and *ascriptivism* as the only possible views for how gender norms apply to individuals, although she doesn't say why. We might read her here as trying to map the available logical space with respect to the kinds of views that a social constructionist might have about gender norms. If the norms in question depend on social practices, then facts about how they apply must also depend on those social practices. Norms that do not depend on social practices may have other conditions for application. A naturalist moral realist, for example, will hold that moral norms are just natural properties of the world, and therefore apply to people because of some natural facts about the kind of beings that they are. However, this type of explanation is not available for norms that are social in origin, such as gender norms. A social constructionist is therefore already committed to the claim that the application of gender norms must be explained at the social level. With respect to the question of how social norms apply to a person, we can divide social practices into two rough categories: (1) the practices of others with respect to that person (e.g. assignment, training), (2) the practices of that person with respect to themselves (e.g. choice, preference, identification). Broadly, Witt uses *voluntarism* here to describe any view which holds that gender norms apply in virtue of (2). She rejects all such views on the

grounds that self-directed practices do not seem to affect whether gender norms apply to us. Ascriptivism, on the other hand, holds that gender norms apply in virtue of (1)—specifically, according to the assignment of a social position and the associated norms.

Notice, however, that accepting Witt's argument against voluntarism merely commits us to the claim that gender norm application depends *in some way* on the social practices of others with respect to the individual. That is: If we are social constructionists, and we reject voluntarism, we are only committed to a relatively thin position which can be cashed out in a variety of ways. Witt's version of ascriptivism is more substantive than this. She argues that gender norms apply to an individual *because the individual is socially recognized as belonging to a gendered social position*. This part of Witt's ascriptivism may be evaluated independently of her rejection of voluntarism.

I suggest, then, that there are two broad kinds of views that a social constructionist might have about gender norm assignment: *voluntarism,* on which gender norms apply to an individual based on their own self-regarding practices; and an alternative view on which gender norms apply to an individual based on the practices of others. To distinguish the latter from Witt's original substantive or "thick" view while still paying homage to her important work, I will call it *thin ascriptivism.* That is:

> **Thin ascriptivism:** Some gender norms apply to subject S because of the practices of others with respect to S, rather than because of S's voluntary self-identification, choice, or preference.

Together, voluntarism and thin ascriptivism exhaust the theoretical possibilities for a social constructionist with respect to gender norm assignment. Both are relatively basic; either may be cashed out in a variety of ways. For example, either voluntarism or ascriptivism will end up being compatible with Witt's more substantive commitment about gendered social positions. It

can be true that gender norms apply to individuals on the basis of the gendered social position that they occupy, *even if* that social position is voluntaristically adopted.

Let me bookmark this discussion for now; I will return to it in 1.3. In the next section, drawing from the work of Witt and others in the analytic tradition, I will defend thin ascriptivism as an important view for feminist purposes.

1.2. THE IMPORTANCE OF BEING (THINLY) ASCRIPTIVIST

In this section, I will argue that thin ascriptivism is an important theoretical commitment for a feminist explanation of the harms of gender norms in dominant contexts. I show that several prominent feminist metaphysicians are (thick or thin) ascriptivists about gender norms. As noted above, I am here primarily discussing the harms of *dominant* gender norms. This is not to say that there are not important, rich, and legitimate domains of gender norms that operate otherwise; for example, we might think that gender norms in subaltern queer and trans contexts work quite differently. Perhaps such norms are not oppressive or harmful; perhaps one reason for this is that these domains are voluntarist, rather than ascriptivist. However, since the dominant context is *dominant,* and tends to affect most or all of us, it is important to capture the way in which those gender norms work in order to explain how oppression is perpetuated.

In her discussion of ascriptivism and voluntarism, Witt gives two reasons for thinking that ascriptivism is particularly important for feminist political purposes (2011, 47). I argued in 1.1 that we should understand Witt's arguments in favor of ascriptivism as supporting thin ascriptivism, rather than the more substantive version Witt proposes; I therefore take the following arguments to support that view. First, (thin) ascriptivism adds to the richness of our understanding of the grip of oppressive social norms by explaining why an individual might feel

drawn under the normative umbrella of a social role of which she is also critical. Second, a (thin) ascriptivist explanation of gender norms suggests that feminist politics should focus on how the social world is normatively structured and criticize those norms that are oppressive, rather than getting bogged down in explanations that appeal to individual psychology.

This case is compelling. I will here elaborate on both points, and strengthen the second. To the first point: Feminist theory has at times been divided on the matter of self-described feminists who embrace feminine aesthetic norms (Serano 2007, 320; hooks et al. 2014). For example, in conversation about this phenomenon, a feminist friend said this to me: "I hate gender norms, but I love high heels." On a voluntarist account, it is difficult to make sense of this position, at least without casting my friend as deeply confused. Wearing high heels is a paradigm of feminine expectation in my friend's cultural context. Moreover, high heels have at times been a focal point of feminist criticism for the physical stress they place on the body. If my friend hates gender norms—that is, if she does not endorse, prefer, or identify with them—her preference for this earmark of normative femininity seems, on a voluntarist account, mystifying. One might make a case about adaptive preferences and the shaping of feminine socialization; perhaps some women are so warped by their societies that they *believe* they do not endorse gender norms, but really, unbeknownst to them, they do. However, a simpler explanation is available that also has the advantage of respecting my friend's full agency. She has become habituated to following feminine norms, and is responsive to them in practice, despite a theoretical distaste for them.

This is a familiar struggle for many of us, and thus points to the intuitive appeal of thin ascriptivism. Simply refusing to identify with some gender norm is not enough to shake its power over us, either externally (as social enforcement) or internally (as felt normative pull).

Importantly, understanding the assignment of gender norms as involuntary or nonconsensual is crucial to understanding the particularly pernicious way in which they harm and oppress. Gender norms license censure, punishment, and violence against those who do not conform to them (Butler 1990); moreover, feminine norms create psychological trauma and a loss of autonomy for women who are subject to them (Bartky 1979, Young 1980). If gender norms attached to us because of voluntary adoption, avoiding these consequences would be simple; we could just refuse to adopt the norms as our own. But this is not sufficient. On the contrary, those who refuse to adopt gender norms are often at greater risk of harm and punishment as a result. As Witt points out, this is key to our understanding of various gendered harms.

The second point is also strong. Witt argues that thin ascriptivism can help feminist politics focus on "changing restrictive social positions and oppressive social norms"—which, she claims, is more important than individual psychology or personal choice. Witt is drawing here from a *structuralist approach* to explaining oppression. The idea here is that we should look to social structures, not individual bad actors, to explain the causes and perpetuation of oppression. As Dembroff puts it, "social kinds are not in the head. If we want to analyze the metaphysics of a social kind, or see whether a certain kind operates in a social context, we must look to the relevant structures and practices in that context" (2018b, 4). For example, according to a structuralist approach, sexism is thought to be perpetuated by norms, habits, symbols, and institutional forces, rather than by individual people who have sexist attitudes and behaviors; thus, it would be possible to weed out all of the active sexists in the world and still grapple with sexism (Young 1990, 56). We therefore ought to undermine gender norm oppression at the structural level. Put differently, it should not be a feminist goal to try to change my friend's mind about high heels. In fact this would seem both unforgivably paternalistic and somewhat fruitless.

What we ought to change is the normative structure of the social world which makes it the case that certain women in certain contexts are *expected* or *required* to wear high heels. A thin ascriptivist perspective can help us do this, as it places our focus on the norms themselves and the social practices involved in their assignment and enforcement, rather than on the individual attitudes of people responsive to them.

Importantly, however, individual responses to gender norms are still relevant to a structuralist position, insofar as they tell us how individuals come to perpetuate structural oppression habitually or subconsciously as a response to normative motivation. That is, norms play a key role in our cognition and behavior. They tell us what we *ought to do,* which in turn shapes what we do. Gender norms therefore shape what we (individually and collectively) do; importantly, they do this in a way that perpetuates gender, and thus gendered oppression. It's a key insight of Witt's view that individuals are shaped by and calibrate their behavior according to gender norms, even when they may prefer not to. Understanding how gender norms "get inside" us in this way will be important to understanding the individual *element* in a structuralist explanation. Nevertheless, a structuralist explanation will take the structure as explanatorily primary to the individual's cognition and behavior; that is, it will hold that the individual's behavior is explained by the norms as they exist at the level of social structure, rather than the other way around. This will be important to remember in successive chapters.

Feminist metaphysicians have tended to treat gender norms as evidence for views about gender classifications or kinds, rather than as objects of study in their own right. As a result, there are few enough discussions of gender norms in their own right within this literature. However, in the discussions that do exist, recent prominent theorists in analytic feminist philosophy have adopted ascriptivist commitments, although not always under that name. Witt's

view gives the most theoretical elaboration here (hence the naming rights). However, Haslanger (2012), Ásta (2018), and Dembroff (2018b, 2020)--three of the most prominent gender metaphysicians in the recent literature--share remarkably similar commitments.

Haslanger (2012) talks of gender norms in terms of *prescriptions*. According to this view, "gendered ideals [depend] on social arrangements" (46). She writes that gender norms "function prescriptively: not only do they serve as the basis for judgments about how people ought to be (act, and so on), but also we decide how to act, what to strive for, what to resist, in light of such norms" (44). The "prescriptive force" of gender norms is "backed by social sanctions" (44) such that "in internalizing the relevant gender-norms, we develop "gender identities"; these gender identities represent reality—self and world—in a form that motivates our participation in the assigned gender role" (45, fn21). She also notes that "although I don't aspire to satisfying this ideal [of femininity], this doesn't prevent others from judging me in its terms" (43). I take it as clear from this discussion that Haslanger is committed to at least thin ascriptivism.

Ásta's (2018) account of gender norm assignment is a bit more complicated. On her view, *conferralism*, observers assign a conferred property of "being of gender G" to a subject on the basis of some observed or imagined *base property*, such as "role in biological reproduction…role in societal organization of various kinds, sexual engagement, bodily presentation, preparation of food at family gatherings, self-identification, and so on" (74-75). The conferred property has social significance over and above the significance of the base property alone. In particular, it comes with certain *constraints and enablements* (75). For example, someone might have the property *being a woman* conferred on them in virtue of being believed to possess the base property *having a uterus*. Having a uterus comes with certain physical constraints and enablements; one may be able to bear children, but is generally

incapable of begetting them, for example. But those upon whom the property *being a woman* is conferred are thereby subject to further constraints and enablements that are not conceptually connected to the possession of a uterus. For example, in the context of my childhood church, women were uniquely permitted to wear dresses, but forbidden from dating other women; neither of these things is importantly related to one's uterine status.

Ásta does not here say much about the relationship between gender conferral and gender norms. Elsewhere, however, in a discussion of *social identity,* she defines a social identity as a "location on the social map" that "consists in constraints and enablements to one's behavior, and accompanying it are social norms for behavior befitting that social location" (122). Let me try to draw out the "accompaniment" relation here, as it appears to be how norms are assigned to an identity on this view. In her discussion of gender conferral, Ásta argues that individuals confer gender in contexts, and enforce associated constraints and enablements, based on their understanding of social arrangements that "exist outside of the context" and are brought in by means of individuals' "gender maps" which develop throughout situations; contextual gender kinds are therefore interconnected across contexts in a "systematic, structural" way (75). Assuming a gender map is a kind of social map, a gendered social identity must be a more stable version of the contextual gender conferrals that come and go. Moreover, drawing from Jenkins' (2016) account of a "gender map," I assume that we need some relatively stable heuristics to help us navigate the gendered world--particularly if, as Ásta argues, it is constituted by a near-infinity of deeply contextual and constantly shifting constraints and enablements.

I suggest, then, that for Ásta, a gendered social identity is a broad *accrual, aggregation,* or *average* of the various and fleeting gender conferrals we experience across contexts--specifically, one that can tell us what to do when gender is conferred on us in a context. Because

identity is relatively stable, where gender conferrals are always contextual, identity can provide us with a normative conception of guidelines for good behavior that can actually shape that behavior. That is, where a contextual gender conferral passes with the context, a social identity can persist across contexts, and therefore can stick around long enough that we can start to become habituated to notions of *what we ought to do* based on commonalities among the constraints and enablements assigned to us. Gender norms are the operative bridge between contextually conferred gendered rules and our behavior. Constraints and enablements say what we can and can't do in a context, but they don't shape our normative deliberations; they are too fleeting for that. Only *norms*, which are connected to our more stable and consistent social identities, can do that.

Ásta further distinguishes between *objective* and *subjective* social identities. An objective identity is "the location on a social map that we occupy stably", and which we can have without being aware of it; a subjective identity is "the location on the social map in the context that we identify with" and for which we "take [the associated] norms as applying to us" (122). Notice, however, that both kinds of social identities are defined with respect to "locations" on a social map. Since locations on a social map "consist in constraints on and enablements to one's behavior" (122)--which has consistently been the description of a *socially conferred property*-- that suggests the locations themselves, and thus the identities, are assigned to us whether we want them or not. What changes between objective and subjective social identity is whether or not we happen to identify with our map location and take up the associated norms. This, as it turns out, is remarkably similar to Witt's view. We can voluntarily identify with our assigned gender norms and try to live by them, or not; that doesn't change the fact that they are already there, packed into the gender classifications that others are always already conferring on us.

Ásta does explicitly allow that individuals can be involved in their own gender ascriptions. In some contexts, for example, the "base property" that others are trying to track in gender conferral might be *self-identification* (76). However, gender conferral and the assignment of constraints and enablements still follows the same pattern in these contexts. It seems clear, then, that Ásta is at the very least a thin ascriptivist about gender norms.

Finally, Dembroff (2018b, 2020) has perhaps the least to say about how gender norms are assigned. However, what they do say clearly assumes an ascriptivist position. For example, they write that "[s]ocial roles, expectations, *norms*, and practices, not to mention self-conceptions, are *imposed on people* based on their gender classification" (2018, 3, emphasis added); and, similarly, that "trans persons *unwillingly* are understood by others in terms of dominant gender kinds, and the unwanted roles, *norms, and expectations* that accompany them" (2018, 19, emphasis added). This language of the *imposition* of norms and of *unwillingness* with respect to that imposition points to an underlying ascriptivist metaphysics; individuals do not choose or prefer the gender norms that apply to them, they are assigned those norms by others.

As we can see, prominent gender metaphysicians have explicitly or implicitly adopted an ascriptivist view about gender norms—even those who hold that elements of gender identity can be voluntary or subjective. This is for good reason. Thin ascriptivism is a theoretically and politically powerful view for explaining how gender norms can oppress. Moreover, in many ways it seems intuitively right. The gender standard does feel mandatory and externally applied. Identification with a set of norms does not affect whether one is held accountable to them by others or whether one feels their normative pull. For example, as both Witt (2011) and Jenkins (2016) point out, a person who rejects feminine norms may still feel that they ought to, e.g., shave their legs, even if they don't endorse that feeling. As Witt puts it, "individuals who do not

practically identify with their socially ascribed gender are nonetheless responsive to those norms and evaluable under them" (2011, 45). Ascriptivism can capture this where voluntarism cannot.

However, all of the views outlined so far come with baggage. In particular, they share a (weaker or stronger) commitment to what I call the *category-based view* about gender norms. In the next section, I define the category-based view and argue that it can come apart from thin ascriptivism. This is fortunate, as I will also briefly raise serious objections to the category-based view. I'll conclude that it is not only possible for us to distinguish between these commitments, it is desirable to do so.

1.3. THE CATEGORY-BASED VIEW

In 1.1 and 1.2, I gave various defenses for a thin ascriptivist view, drawn in part from the work of Witt (2011). I will now return to Witt's original version of ascriptivism, to discuss the way in which her *thickly* ascriptivist view is shared to a greater or lesser extent by other theorists in this tradition. In this section, I will articulate a common commitment of these views; the idea that gender norms apply to individuals on the basis of their *gender categories*. I call this the *category-based view*, and argue that it is specifically *not* licensed by the arguments in favor of thin ascriptivism.

Recall that Witt defines ascriptivism as the view that "[a] social role is normative for an individual if she or he occupies a given social position…whether or not that individual consciously identifies with or chooses that social position" (43). That is, for Witt, ascriptivism is about the connection between some *social position* and the associated norms of conduct. Witt's "thick" ascriptivism is true if the norms associated with a particular category apply *only to* those

who occupy the social position, *because* they occupy the social position. Put plainly, "the normativity attaches to the social position occupancy itself" (Witt 43).

I have argued that this latter claim is a substantive commitment in its own right, one that comes apart from, and is not supported by the arguments in favor of, thin ascriptivism. I will now draw out this commitment and demonstrate its role in the views of other prominent theorists, in order to investigate it on its own merits. Let me begin by articulating what I take to be the "substantive commitment" in question:

> **The category-based view (CBV):** Gender norms apply to individuals only if and because those individuals belong to the associated gender categories.

I understand 'gender categories' to be quite broad, encompassing a variety of phenomena which might be understood differently. For example, Witt (2011) describes "gendered social positions"; Dembroff (2020) is alternately interested in what they call "gender kinds" or "gender classifications"; Haslanger (2012) talks about "social classifications"; and Ásta (2018) talks about "identities". I do not mean to blur important distinctions between these views, or argue that there are not important metaphysical differences between kinds, classifications, positions, identities, and so on. Rather, I am targeting what I take to be the following commonalities among these views: (1) the idea that individuals are divided, classified, or categorized according to gender in some way; (2) the idea that the resultant categories have associated norms; and (3) the idea that this categorization explains why the individuals in question are subject to the relevant gender norms. In its broadest formulation, the category-based view is a commitment to the claim that, *however* we divide, classify, or categorize individuals according to gender, *that division, classification, or* categorization—that is, the ontological status of being sorted, classified, or

categorized—is what explains the norms to which those individuals are responsive and under which they are evaluable.

As discussed, Witt explicitly states that gender norms attach to social position occupancy, and thus only apply to those who are ascribed a gender category (2011, 43). This is clearly a category-based view. Haslanger (2012) takes a similar position: "Particular traits, norms, and identities, considered in abstraction from social context, have no claim to be classified as masculine or feminine. The classification of features as masculine or feminine is *derivative*, and in particular, depends on prior *social* classifications" (46, emphasis in original). For Haslanger, gender categories are inherently normative, in that it is partly constitutive of those categories that they give rise to a normative standard for members of the category. Since the function of these normative standards is to delimit appropriate or "ideal" behaviors for category members, only category members can be held to those standards (45-46).

Ásta holds that social norms are associated with a *social identity.* The norms "accompany" one's location on a social map (2018, 122). She notes that one can identify with this location or not, and thus one's *objective* and *subjective* social identities can come apart. In either case, however, the location itself is constituted by the constraints and enablements that are 'conferred' by others. I argued in 1.2 that constraints and enablements in particular contexts have broader, normative correlates, or simply *norms,* which serve as a "bridge" between what is allowed to do and not do in and across contexts, and one's normative cognition or motivation. For Ásta, then, social norms are dependent on social identity, and one is subject to social norms *because* one occupies the relevant location on the social map and thus has the associated identity.

Finally, Dembroff (2018b, 2020) again says relatively little about the assignment of gender norms to individuals. Even in their brevity, however, we can find clear commitments to a category-based view. Recall: "norms...are imposed on people *based on their gender classification*" (2018b, 3, emphasis added). We could not ask for a more straightforward endorsement of category-based ascriptivism.

In short, although opinions differ on the exact relationship between gender categories and gender norms, we can see in each of these views, implicitly or explicitly, a shared commitment to what I have called the category-based view: gender norms apply to individuals only if and because those individuals belong to the associated gender categories.

I will begin by raising two points about the category-based view. First, there is relatively little explicit defense given of it in any of these views. Witt (2011) argues in favor of *thin ascriptivism*, but, as discussed, those arguments leave it open as to whether the further commitment is justified. Haslanger (2012) comes the closest to defending a category-based view. She argues that gender norms are causally related to gender categories, because without gender categories we would not have gender norms. However, this in itself is not a defense of the view that gender norms only *apply to* those who occupy the relevant categories. In fact, Haslanger herself notes that "norms and roles can also fall desperately out of sync when the norms remain rigid while social roles change; gender-norms 'often take on complex lives and histories of their own, which often bear little resemblance to their functional roots'" (44, quoting DiStefano 70). I will investigate Haslanger's interesting and rich functionalist picture of gender norms at length in Chapter 2. For now, let me just say that the category-based view as I have defined it here is far more often *assumed to be true* as a feature of ascriptivism than it is defended on its own merits.

This leads me to the second point: There is no principled reason to think that the category-based view is a necessary feature of ascriptivism. The two quite clearly come apart. Why, then, have they been treated as coextensive within this literature? I suggest a potential *explanatory* reason: Analytic gender metaphysics has, on the whole, adopted category-based analysis about gender. That is, it has tended to treat *all* gender phenomena as if they proceed from gender categories. When analytic philosophers ask what gender is like, they often end up asking what gender *categories* are like: What makes something a gender category? What is it to be a woman, a man, or neither? Who counts as a member of this kind or that kind, and why? The Stanford Encyclopedia of Philosophy explains this thusly:

> The aim of feminism is, in the most general terms, to end the oppression of women. The goal of feminist theory is, therefore, to theorize how women are oppressed and how we can work towards ending it. But what is this group *women*? Whose oppression is the movement aiming to end? For articulating the various ways in which women are oppressed, there is a need for a working definition of what it is to be a woman" (Haslanger and Ásta 2017: https://plato.stanford.edu/entries/feminism-metaphysics/]

In this context, as noted above, gender norms are often treated as data points providing evidence for a further view, rather than an object of inquiry in their own right. A category-based view allows gender norms to do this theoretical work.

However, recent work in analytic philosophy has begun to question the assumption that gender categories are explanatorily primary in understanding gender as a social structure. Rather, some philosophers have begun to argue that we should understand *gender* as a complex and multifaceted phenomenon involving many different kinds of practices and features; *gender categories* are merely one iteration of this beast. For example, Mikkola (2016) has argued that feminist philosophy ought to move away from its laser-focus on the "woman question." Such a focus, she argues, has bogged us down in complex semantic and ontological puzzles. However,

such puzzles are not necessary to doing good feminist theory. We are perfectly capable of talking about "women" as a group without elucidating a thick concept of "woman." Mikkola suggests we move towards a *humanist feminism* on which the harms we seek to ameliorate are harms to the *humanity* of women, rather than their womanhood.

Similarly, Antony (2020) defends a *deflationary* concept of "woman", one that will allow us to assert generic claims about women without giving a substantive metaphysics of gender categories. She writes that "whatever the inherent philosophical interest of such questions as 'what is it to be a woman,' there is no practical or political need to answer them" because our "most pressing political needs, as feminists, are to challenge injustices and harms" (Antony 2020, 531). Barnes (2020) argues that we ought to separate our analyses of *gender* as a social system from our analyses of *gender terms* and their extensions. And Dembroff (forthcoming) argues that patriarchy does not depend on gender kinds alone, but rather is a kind of self-sustaining social ecosystem which serves to privilege "real men." This is not a gender category but is rather a purely normative distinction: "real men" are men who perform masculinity "correctly," a task which involves having the right kind of body, sexuality, economic position, etc.—and which in fact leaves out most members of the category "man."

These theorists are moving away from treating gender kinds, categories, and roles as the focal point of inquiry. The thought is that, while the infamous "woman question" may be important for questions about identity and inclusion in one's preferred category (cf. Watson 2014, Jenkins 2016), it can come apart from, and moreover, is not more theoretically important than, other questions about gender—such as who is socially dominant and subordinate, who is subject to which sanctions, how we go about gendering and policing one another, and how all of this is *naturalized* and thereby justified.

This insight is far from unique to feminist analytic philosophy. Feminists in other traditions, in particular feminists of color and intersectional feminists, have for a long time understood gender as a multifaceted and complex phenomenon that goes beyond univocal gender categories (see, for example, Crenshaw 1989; Lugones 2007, 2010). This is, I suggest, all the more reason for feminist metaphysics to develop views which focus on other phenomena, such as gender norms.

It is perhaps unsurprising, then, that the question of how gender norms apply to individuals is distinct from the question of how individuals come to be members of gender categories. Notice also that a category-based view about gender norms is compatible with various views about the latter, including: an ascriptivist view, on which gender categories are assigned to individuals by others; a voluntarist view, on which gender categories are adopted by the individual; or even various essentialist views, such as one on which gender categories are essential properties of individuals which partially constitute their fundamental being.

In 1.2 I argued that it is important for feminist purposes to maintain commitment to *thin ascriptivism*, or the view that gender norms are assigned to individuals by others rather than adopted by those individuals. However, this alone does not justify the thicker view on which gender norms are assigned *on the basis of gender category*. We should evaluate the *category-based view* on its own merits, rather than as a necessary commitment of ascriptivism.

Those merits, as it turns out, are questionable. The category-based view faces important objections from the experiences of marginalized people. First, the category-based view does not capture the way individuals experience themselves as *responsive to* gender norms. Second, the category-based view does not capture the way individuals and other entities are *evaluable under* gender norms. Although these objections are key to my overall argument in this dissertation, I

will sketch them here only in brief, in order to provide a preview for how that overall argument will go. Chapter 2 will be devoted to motivating objections against the category-based view.

My first objection to the category-based view is drawn from the experiences of trans and GNC people. As noted in 1.2, one advantage of Witt's "thick" ascriptivism is that it can explain involuntary gender norm *responsiveness*, or why individuals feel the normative pull of gender norms which are assigned to them but with which they don't identify. However, a category-based view does not explain what I call *cross-category norm responsiveness*. Some individuals, in particular some trans and GNC people, are responsive to gender norms which are associated with a gender category that was *not* assigned to them. For example, some people have *not* been assigned the category "woman," but nevertheless feel the pull of *feminine* norms—which they have *not* been expected to follow, and, moreover, which they may be punished for following. In a culture where it is a feminine norm to shave one's legs, for example, some transfeminine people and gender-nonconforming cis men feel that they ought to shave their legs. This is a problem for all category-based views, including voluntaristic ones. Cross-category normative force often occurs at a time when the person in question also does not *identify with* the gender category associated with the norms, such as in childhood. Rather the opposite; cross-category normative experiences are often the impetus for adopting a gender identity, rather than the result of this adoption (see Jenkins 2016). Cross-category norm responsiveness may also be felt by people whose gender categorization *never* matches the norms to which they are responsive; consider, for example, the rich history of butch women in queer culture, who identify and are categorized by others as women, but nevertheless may feel responsive to norms of masculinity.

My second objection concerns the way that marginalized people, groups, and societies are *evaluated* under gender norms. Often, those who are not included in a particular gender

category are nevertheless treated as evaluable under associated norms. For example, María Lugones (2010) has pointed out that Black and Brown colonized peoples have historically been denied assignment to a binary gender category. Colonialist efforts to control and 'civilize' Black and Brown people "used the hierarchical gender dichotomy as a judgment, though the attainment of dichotomous gendering for the colonized was not the point of the normative judgment" (744). Moreover, Spillers (1987) argues that Black enslaved people were "ungendered" and thereby dehumanized and brutalized. And Wilchins (1997) chronicles the way that many trans and GNC people are treated as "gendertrash" relative to rigid binary gender classifications; they are "thrown out" of those categories, as it were.

Not being afforded a particular gender category, however, does not excuse marginalized people from evaluability under the associated gender norms. Rather the opposite. In many such cases, the people excluded from the gender categories are marked as inferior on the associated gender standard *because* they are not afforded gender classification; they are "set up to fail" on standards of masculinity and femininity. In the case of colonized people, for example, this justifies "unimaginable exploitation, violent sexual violation, control of reproduction, and systematic terror" against their bodies (Lugones 2010, 744). And Spillers (1987) writes of how, in the case of Black enslaved people, "the captive body becomes the source of an irresistible, destructive sensuality" while "at the same time—in stunning contradiction—the captive body reduces to a thing, becoming *being for* the captor" (67). A feminist account of gender norms ought to be able to explain the ways in which they enable harm and oppression. If the category-based view cannot explain these distinctive harms, then it fails to do the work we need it to do.

Again, I say more about these objections in Chapter 2. I include them here to make this preliminary point: There are good reasons, drawn from the experiences of marginalized people,

to think that gender categorization does not tell the whole story about gender norm assignment. In particular, a category-based view does not capture either the way gender norms exert normative force over those who are most directly harmed by them. This is an unacceptable theoretical oversight. Insofar as feminists understand ascriptivism to be an important theoretical commitment, we should question its association with the category-based view.

1.4. CONCLUSION

To recap: I have distinguished the arguments for (thin) ascriptivism from the arguments for the category-based view, and argued that (thin) ascriptivism is important to a feminist perspective, while the category-based view faces serious problems. Under these conditions, we should consider how a feminist account of gender norm assignment might proceed without presupposing a category-based view.

I have identified a persistent conflation in predominant views in analytic feminist philosophy: a commitment to *thin ascriptivism,* or the view that gender norms are assigned to individuals without their participation or consent, is run together with a commitment to the *category-based view,* or the view that gender norms apply to individuals on the basis of the gender categories that they occupy. Analytic feminist philosophers have often focused on the importance of gender categories as explanatory for various kinds of gender phenomena. The category-based view reflects this focus; it prioritizes gender *categories* in explaining how gender norms attach to individuals. However, this view leaves out the experiences of those who are marginalized or excluded according to dominant gender categories. Often, these are the people who are most harmed by gender norms, *because* they are marginalized within or excluded from those dominant categories. A feminist account of gender norms ought to capture the experiences of those who are most harmed by them.

The theoretical work here is just beginning. Without a category-based view, we lack an account of the way that gender norms are assigned, and thereby come to hold sway over individuals. This is the project I will undertake in the remainder of the dissertation.

2.0. INTRODUCTION

Many prominent accounts within analytic gender metaphysics proceed from the assumption that gender norms apply on the basis of *individual gender categories*. I have called this the *category-based view*. On a view like this, norms of femininity function to constrain the behavior of those who occupy the category *woman,* and thus apply to individuals because they occupy that category. In Chapter 1, I articulated the category-based view as a substantive commitment shared by many such theorists, in particular Witt (2011), Haslanger (2012), Ásta (2018), and Dembroff (2018b, 2020). We can see the impact of the category-based view throughout the recent literature in analytic gender metaphysics.

In her defense of the category-based view, Witt points to a central question about social norms more generally: "Why are individuals responsive to and evaluable under social norms?" (2011, 44). Witt's question is aimed at understanding the *normative power* of social norms. From this way of framing the question, we can understand this normativity to have two related prongs, corresponding to "responsive to" and "evaluable under." Roughly speaking, *responsiveness* is the way that agents calibrate their own behavior and normative deliberations according to the norms, and *evaluability* is the way that entities are candidates for criticism and enforcement (that is, reward and punishment) on the basis of their compliance or noncompliance with a norm. On this view, then, an account of the normative power of gender norms should be able to explain two things: 1) Why do agents respond to certain gender norms (but not others)? 2) Why are agents evaluable under certain gender norms (but not others)? Both of these are key elements of gender's normativity. If we want to understand how gender norms function to reify oppression,

we need to understand their role in sustaining *inequality,* which in part requires understanding their *differential impact.*

At first blush, the category-based view has plausible answers. Casual observation of the gender norms at work in many, or most, dominant contexts will return the following analysis: People are sorted into gender categories according to how others interpret their sexed bodies. People are also held to different normative standards according to how others interpret their sexed bodies. The normative standards applied and the gender categories assigned tend to match up. We might reasonably conclude that the two are importantly related. Those who have certain kinds of sexed bodies are assigned particular gender categories, and the associated normative standards are enforced on them accordingly. Moreover, people tend to act in ways that match these normative standards; those classed as *women* tend to follow feminine norms, while those classed as *men* tend to follow masculine norms. As Simone de Beauvoir famously observed, "anyone can clearly see that humanity is split into two categories of individuals, with manifestly different clothes, faces, bodies, smiles, movements, interests, and occupations" (2011, 4). Like Beauvoir, we should reject gender essentialism as an explanation of these phenomena. It's not the case that humans have deep and fundamental gendered parts of themselves that make them masculine or feminine. Why, then, is gendered behavior reproduced by individuals who may or may not endorse, or even be consciously aware of, the norms that guide it? Gender categories seem capable of doing a lot of explanatory work here. They are, by and large, conferred on individuals without consent, and they come with a set of norms, which are then enforced so ubiquitously that individuals may find themselves behaving in accordance with those norms as a matter of lifelong habit--whether they would prefer this or not.

But this analysis is incomplete. It presumes, first, that the individuals in question are permanently and unambiguously placed in fixed gender categories; that membership within those categories explains the way gender norms are enforced; and that, as a result, the norms associated with those categories straightforwardly guide behavior. These can be useful heuristics for understanding the experience of those who have significant embodied privilege--i.e. whose experience of gender norms is mediated by the fact that they are also white, cisgender, non-disabled, heterosexual, thin, and so on. Those without this privilege, however, have quite a different experience.

In this chapter, I will draw from the experience of marginalized people to show how a category-based view faces two major objections in terms of explaining gender's normative power. First, it cannot explain *responsiveness* to gender norms in the case of what I call *cross-category norm responsiveness.* Many people, particularly transgender (trans) and gender-nonconforming (GNC) people, experience themselves as responsive to gender norms which do not match any gender category that they appear to occupy. A category-based view cannot explain this. Call this the *responsiveness objection.* Second, a category-based view cannot explain *evaluability* under gender norms, specifically in cases where marginalized people are actively excluded, ejected, or otherwise constructed as *outside* of particular gender categories, but are nevertheless evaluated (and often brutally punished) according to the associated norms. Relatedly, the view cannot explain how groups, cultures, and other big-picture social entities--what I'm calling *macro-level social entities*--are treated as evaluable under gender norms. Call this the *evaluability objection*. There are many versions of the evaluability objection. In this chapter I focus on three: Lugones' account of the enforcement of colonial gender norms on Black and Brown colonized people and societies, Spillers' account of the "ungendering" of

Black enslaved women through the legacy of chattel slavery, and Wilchins' account of trans and gender-nonconforming people as "gendertrash."

All together, this suggests that a category-based analysis is only useful for explaining a narrow phenomenon: the individual experiences of those who have unambiguous membership in dominant gender categories, in virtue of their significant privilege along various dimensions. The experiences of those who are marginalized within or excluded from those categories, and the ways in which gender norms operate at the macro-level to shape groups and societies, are ignored by a category-based view. This suggests that the view is not sensitive to the deep systemic co-construction of gendered and racialized norms through the legacy of coloniality, chattel slavery, and white supremacy. I argue that this is reason to reject it.

I proceed as follows. In 2.1, I describe the phenomenon of *cross-category norm responsiveness* and distinguish it from other closely related phenomena. In 2.2, I draw from the previous section to motivate the *responsiveness objection* to a category-based view. I argue that cross-category norm responsiveness demonstrates that gender categories can't explain how agents become responsive to gender norms. In 2.3, I discuss the ways in which gender categories as a frame of analysis have failed to capture the experiences of those who are excluded from membership in those categories. In 2.4, I draw from the previous section to motivate the *evaluability objection* to a category-based view. I argue that these examples show that gender categories can't explain how entities are evaluable under gender norms. Rather the opposite: gender norm enforcement is particularly brutal for those who are excluded from dominant categories. Moreover, I argue that a category-based view is fundamentally ill-prepared to explain the effects of gender norms at what I call the *macro-level*. In 2.5, I conclude by drawing together previous suggestions to argue that a category-based view focuses on the experiences of gender-

conforming white people, and as such, is insufficient to explain how gender norms exert
normative force in the world.

2.1: CROSS-CATEGORY NORM RESPONSIVENESS

In this section, I will motivate the *objection from cross-category norm responsiveness*.
This objection will draw from the experience of trans and GNC people to argue that the
category-based view does not explain the way agents actually calibrate their behavior relative to
gender norms. I will begin by articulating the phenomenon of norm responsiveness. I draw here
from Witt's work on Aristotelean habituation (2011, 2020) and Jenkins' work on gender identity
(2016, 2018). I then articulate and describe a common (though in no way universal) experience
among trans and GNC people: a sense of responsiveness to some gender norm or norms which
have *not* been assigned and which do *not* correspond with one's assigned gender category. I
argue the category-based view cannot explain this phenomenon. As such, it misses a key feature
of the way that gender norms are experienced as normative.

According to Witt's ascriptivism, individuals experience themselves as "responsive to
and evaluable under" various kinds of social norms (2011, 42). One is *evaluable under* a social
norm when "the individual is a candidate for evaluation by others in relation to that norm" (33)--
that is, others expect one to adhere to it, and perhaps engage in punishment or censure when one
does not (or, at the very least, would consider themselves justified if they did). One is *responsive
to* a social norm when one experiences the norm as guiding one's behavior. This could mean that
one experiences any of a "full range of possible reactions to a norm....from compliance to
critique" (33). One might, for example, unthinkingly accept the norm as binding and comply
with it; actively reject it and rebel against it; or carefully consider its relationship to one's values

and motivations, and ultimately decide to comply. Each of these behaviors is a way of being responsive to a norm. That is, responsiveness to a norm is not merely behavioral. There is a phenomenological component; one *feels* evaluable under or *experiences oneself as* guided by the norm, and one has a sense of doing something *wrong,* on the relevant normative standard, when one does not comply.

Norm responsiveness plays an important role in Jenkins' (2016, 2018) influential account of gender identity. She compares the following two cases:

> Consider a woman who feels that having visible body hair on her legs is unattractive, embarrassing, and unacceptable. In a visceral way, having hairy legs feels wrong for her.... Contrast this with the experience of another woman who does not remove hair from her legs. Her awareness of her body includes the awareness that in having hairy legs she is contravening dominant norms of feminine appearance--on some level she knows that people like her are not meant to look like that, according to dominant ideology. This may be so despite the fact that she is perfectly content to have hairy legs and for them to be seen by others. Her experience of social and material reality includes navigating the norm that women should have hairless legs, even though she is not complying with it. (2016, 411)

Jenkins draws from these examples to argue that one's gender identity is grounded in a sense of experiencing some gender norm or sent of norms as *relevant to you*. On this account, these women share a gender identity because they share this sense that the feminine norm is relevant to their conduct; as Witt puts it, their "behavior is calibrated in relation to the norm" (2011, 32). Where Jenkins' two women differ is in how much they care about that fact.

This passage evokes precisely the phenomenon I am trying to capture. Both women are described by Jenkins as being intimately connected to "a norm of feminine appearance" (412) which recommends having hairless legs. However, their subsequent relationship to this norm is not merely cognitive. The women do not simply *think* or *believe* they are subject to the norm.

Similarly, it's not clear that these women undertake any active voluntaristic process such as *adopting* or *taking up* the norm. In fact, let us stipulate that neither woman has reflected on their relationship to femininity or gender norms at all; both are simply moving through the world as seems right to them, without any particular beliefs or intentions relative to gender. Nevertheless, both women experience an embodied awareness where "on some level" they understand that the norm of feminine appearance *applies* to them--that "people like [them] are not meant to look like that" (411), and having hairy legs means that each of them is doing something *wrong* according to the norm.

I am not interested in defining gender identity. Thus, Jenkins' further argument that norm responsiveness constitutes gender identity is irrelevant to my discussion here. The phenomenon of *norm responsiveness* as she identifies it is nevertheless important. If we want to understand how individuals come to calibrate their behavior relative to gendered standards of behavior, we must pay special attention to the ways in which that behavior is in fact shaped by those standards. Moreover, I take norm responsiveness in itself to be a bare psychological fact. I want to recuse myself from the discussion on how individuals come to *acquire* norm responsiveness. I am not in the business of explaining the psychological mechanisms by which individuals internalize the social norms that are presented to them. My claim is just this: individuals tend to calibrate their behavior relative to gender norms, and that this calibration often does not involve any particular cognitive state or voluntary adoption.

Why might this be the case? How do paradigmatically social norms, such as gender norms, come to guide our behavior without our participation or consent? One answer comes from Witt (2011, 2020). In brief, individuals become *habituated* to the norms that are assigned to them. As others expect certain behaviors from us and thereby train us to inhabit certain

normative roles, we come to perform those behaviors regularly and often. Over time, we learn to do them without thinking about it, or by *habit*. This elegant and simple account is grounded in the Aristotelean idea of *excellence,* and has much to offer in terms of understanding gender norm responsiveness. Moreover, it dovetails nicely with a central insight from the social sciences: the *internalization hypothesis,* or the view that individuals become motivated to follow social norms because, over time, they internalize the expectations of others. I will expand on *habituation* and *internalization* in Chapter 3.

One major advantage of the habituation account is that it explains how individuals who don't identify with or prefer their assigned gender categories can nevertheless feel responsive to them. For example, as Witt points out, someone assigned the gender category of "woman" who does not prefer or identify with the associated norms, such as someone who rejects gender's normative power on political grounds, might nevertheless feel responsive to norms of femininity against their will. A feminist account of gender norms needs an explanation for this phenomenon. Recall from Chapter 1: according to Witt, this is the primary reason that *voluntarism* does not succeed. Ascriptivism and the *habituation* account, however, capture this easily and by design.[1] Individuals are assigned a gender category and an associated set of norms, and over time acquire the disposition to habitually and unthinkingly exemplify the behaviors recommended by those norms.

However, Witt's focus on the phenomenon of *responsiveness to* a gender norm illuminates a problem for the category-based view. In some cases, individuals experience themselves as responsive to norms which do not match their gender category; that is, they

---

[1] A brief reminder: Voluntarism is the view that individuals are subject to a set of norms because they *adopt* or *take up* the norms. Ascriptivism is the view that individuals are subject to a set of norms because those norms are *assigned* to them by others.

experience what I call *cross-category norm responsiveness.* In what follows, I will explain cross-category norm responsiveness, distinguish it from other closely related normative phenomena, and argue that it presents a significant challenge against the category-based view.

Recall the central claim of the *category-based view:*

The category-based view: Gender norms apply to individuals only if and because those individuals belong to the associated gender categories.

The category-based view, together with *habituation*, suggests the following: Only those individuals who belong to a particular gender category will be trained to follow the relevant norms, and so only individuals in a gender category will develop *responsiveness* to those norms. But this isn't always how gender norm responsiveness works. Some people find themselves responsive to norms associated with a category which has *not* been assigned to them. For example, some people who have not been assigned the category "man" might nevertheless find themselves responsive to masculine norms--such as trans men and transmasculine people, non-binary people who aren't assumed to be men, and butch or gender-nonconforming women. The analogous is true for many people who haven't been assigned the category "woman" but nevertheless find themselves responsive to feminine norms. Since this norm responsiveness occurs "across" gender categories, I call this *cross-category norm responsiveness.*

Accounts of cross-category norm responsiveness are quite common in the self-reports of trans and GNC people.[2] For example, Julia Serano describes her childhood experiences of

---

[2] I purposefully understand "gender-nonconforming" or GNC to be a very broad category, covering the lived experience of anyone who, for whatever reason, is perceived as non-normative on some dominant gender standard. This is not meant to be a natural kind or to suggest that there are not many importantly distinct categories under the GNC umbrella. GNC terminology exists in relation to a deeply problematic system of gender classification—dare I say an *un*natural kind—and is only meant to play a role in arguments about that system. Notice also that on some understandings, racialization, disability, body size, etc. will entail that one is gender-nonconforming. For example, Bey (2017) holds that Blackness is fundamentally at odds with dominant gendering, and thus "Black is trans". In 2.4 I discuss the differential

feeling responsive to norms that did not match the gender category that she and everyone else believed she occupied:

> I had an unexplainable feeling that I was doing something *wrong* every time I walked into the boys' restroom at school; and whenever our class split into groups of boys and girls, I always had a sneaking suspicion that at any moment someone might tap me on the shoulder and say, "Hey, what are you doing here?" (2007, 78, emphasis added)

Serano's normative phrasing here is telling. She experienced herself as doing something *wrong* when she violated a feminine norm--even though she and everyone else believed that those norms did not apply to her. These experiences can't be explained in terms of category-based habituation. Serano clearly knew that the boys' bathroom was the one others expected her to occupy, just as she knew that others placed her on the "boy" side of the gender division. According to the category-based view, she should not have internalized any norms that would explain her felt responsiveness. But clearly she did. Moreover, the situations she gives in this example are paradigmatically, and perhaps uniquely, gendered. That is, the primary, perhaps the only, norms in play in determining whether one goes into the boys' or girls' bathroom, or joins the boys' or girls' side of the classroom, are *gender* norms.

For Serano, this *wrongness* for violating feminine norms was matched with a *rightness* when she began to follow them:

> It became obvious to me that I wanted to be a girl and that, on some level, it felt *right*... Saying that I "wished" or "wanted" to be a girl erases how much being female made sense to me, how it felt *right* on the deepest, most profound level of my being (80, emphasis added).

---

enforcement of gender norms on those who do not comfortably fit into dominant gender categories. For now, understand "gender-nonconforming" as broadly as possible.

Although Serano is talking in terms of *category* here, I suggest that we can understand her description of "being" here in an active voice. *Being* a girl felt right for her, just as doing masculine-coded gendered behaviors felt *wrong* for her. This is illuminated by the following recollection:

> I found myself inexplicably compelled to remove a set of white, lacy curtains from the window and wrap them around my body like a dress. I walked toward the mirror. Since I was a prepubescent boy with one of those longish boy haircuts that were popular in the late '70s, the curtains alone were sufficient to complete my transformation: I looked like a girl. I stared at my reflection for over an hour, stunned. It felt like an epiphany because, for some unexplainable reason, *seeing myself as a girl made absolutely perfect sense to me.* (2007, 79, emphasis added)

Experiences like Serano's are fairly common among trans and GNC people, as is the use of normative terminology to describe them. This sense of *rightness/wrongness* or *making sense* here is key. Cross-category norm responsiveness is a distinctive and theoretically interesting phenomenon because it manifests, not merely as a want or a wish to follow some norms or occupy some category, but as a felt sense that one *already is* responsive to some norms associated with that category. To feel responsive to some norm is to feel as if one should exemplify the properties and behaviors recommended by that norm, and that failing to exemplify those properties and behaviors is violating that norm. But, as noted by both Witt (2011) and Jenkins (2016), a person cannot violate a norm unless it applies to them. And, according to the category-based view, the norms of femininity did not apply to Serano as a child; she was not initially assigned the category of "girl" or "woman" by others, nor did she identify that way for many years thereafter.

In her memoir *Redefining Realness*, Janet Mock describes a similar pattern in her interactions with the norms of femininity:

Like most teen girls (whether they're trans or cis), I had a vision board of my ideal, pulled mostly from the pop-culture images that MTV fed me. I wanted Halle Berry's or Tyra Banks's breasts, Britney Spears's midsection, Beyonce's curvy silhouette and long hair, and I prayed that I wouldn't grow any taller so I didn't tower over the petite Asian girls who were the barometer of beauty in the [Hawaiian] islands. (122-123)

Mock fully understood that others were holding her to standards of masculinity. Earlier in the book, she recounts a childhood experience of being "certain I was a boy.... It was the first thing I'd learned about myself as I grew aware that I existed" (15). As a child in her kindergarten classroom, she understood, for example, that boys used blue cubbyholes and girls used red cubbyholes (15). Nevertheless, even as a small child, she experienced a "faint desire, whose origins I can't pinpoint to a pivotal *aha!* moment" to "step across the chasm that separated me from the girls--the one who put their sandals in the red cubbyholes" (16).

Mock's account sheds light on the distinction between two closely related phenomena: the *desire* to do something, and the sense that one *ought* to do it. She *wanted* to behave and look feminine, but she also experienced herself as *already* responsive to normative standards of femininity, as we can see from her interactions with the "vision board of her ideal." This distinction is drawn out better by Serano, who sometimes experienced herself as responsive to feminine norms even though she *didn't* particularly want to engage in the associated behaviors: for example, she writes "I never really *wanted* to take part in girlish activities, such as playing house" (2007, 79, emphasis added). Contrast this with her experiences of *rightness* and *wrongness* relative to certain gendered behaviors.

This discussion points to the need to distinguish between cross-category norm responsiveness and some closely related phenomena, each of which falls under the category of *desire*. The broad point here is that cross-category norm responsiveness is not simply a *desire* or

57

a *wish*. It is something more normatively substantive. Philosophers sometimes speak of the *normative pull* of certain norms, where a norm tells us what we *ought* to be doing--often understood in contrast to our desiderative states. One way to understand this is in terms of the kinds of *reasons* we have to act. I will discuss this further in Chapter 4. For now, I will make the preliminary claim that norm responsiveness is not a mere desire. To draw out this point, I will investigate the difference between cross-category norm responsiveness and three distinctive kinds of desires relative to gender norms: (1) the desire to occupy a category that one does not currently occupy; (2) the desire to engage in some behavior for the sake of that behavior in itself, rather than because of its relationship to some norm; and (3) the desire to violate or rebel against a norm to which one is subject. Cross-category norm responsiveness is none of these things, although it often coincides with all three.

Consider the following analogy. Suppose that I want to be a professional rock climber. I have a strong desire to perform that function, be subject to the associated norms, and engage in the behaviors characteristic of professional rock climbers. I want to be able to climb challenging rock faces like El Capitan in Yosemite National Park; to make a living solely on athletic sponsorships and prize money; to compete in world-class rock climbing competitions; and so on. There are norms for behavior on the part of those who occupy the category "professional rock climber". Professional rock climbers should train their bodies to peak rock climbing performance, prioritize rock climbing over other kinds of activity, consume certain kinds and quantities of food and drink and avoid others, and so on.

However, suppose that, although I want or wish that I were a professional rock climber, I don't have any important relationship to the norms that apply to professional rock climbers. It doesn't seem to me that I *ought* to follow these norms. When I take a few weeks off from

training for rock climbing to focus on writing my dissertation, I don't feel as if I've done something wrong. When I am not invited to compete in a rock climbing championship, I don't feel as if I've been passed over. Even though I wish that I occupied this category, the associated norms just don't apply to me, and not following them doesn't feel like a norm violation. In this case, I want to occupy a particular social role that I currently don't occupy, but I do not feel the normative pull of the associated norms. Cross-category norm responsiveness is the experience of feeling *already* responsive to a particular norm or set of norms that don't appear to apply to you. My desire to be a professional rock climber may involve a *desire to be* responsive to and evaluable under a particular set of norms, but it does not involve the felt sense that I already am. Cross-category norm responsiveness therefore comes apart from the desire to occupy a social category that one does not currently occupy.

Moreover, and similarly: suppose that I want to engage in the behaviors that are characteristic of professional rock-climbing for reasons that have nothing to do with the norms of being a professional rock-climber. Perhaps I want to climb El Capitan in Yosemite National Park because it is challenging and beautiful and I value the associated sense of accomplishment and bragging rights. Perhaps I want to have a professional athletic sponsorship because it is prestigious and the income is good, and perhaps I want to be invited to rock-climbing championships because I like to travel. However, suppose that I do not feel as if the norms of professional rock-climbing have anything to do with my desires. That is, I don't feel like I *should* act like this; I simply want to. In this case, I want to engage in many of the behaviors associated with professional rock climbing, even though I don't feel responsive to associated norms. Cross-category norm responsiveness therefore comes apart from the desire to engage in certain behaviors associated with a particular social category for their own sake.

Finally, suppose that I achieve my dream and I become a professional rock climber. However, once I belong to this category, I learn that professional rock climbers are expected to enter into partnerships with companies who are responsible for massive amounts of environmental damage, in order to improve the reputations of these companies. This norm now applies to me as a professional rock climber; suppose that I feel that I ought to comply. Nevertheless, I don't want to partner with these companies. I want to *violate* this norm. I understand that it applies to me; I am responsive to it; and I nevertheless prefer not to comply with it. I am responsive to a norm that *corresponds* to my social category, but I have other motivational states that pull against it, even in the face of my responsiveness to it. Call this the "phenomenology of rebellion." To have the phenomenology of rebellion, I must both e*xperience myself* as responsive to a norm, and *desire* not to comply. This is different from, but not incompatible with, experiencing cross-category norm responsiveness. If I am responsive to a norm which corresponds to a category to which I don't belong, I can have any number of other attitudes *about* that responsiveness. I might prefer to comply straightforwardly with the norm, to creatively reinterpret it, to rebel against it, and so forth. But norm responsiveness in itself is distinct from any of these preferences *about* one's norm responsiveness. Cross-category norm responsiveness therefore comes apart from the phenomenology of rebellion.

Let's apply these distinctions to the case of gender norms. First, individuals may experience a desire to occupy a gender category without feeling cross-category norm responsiveness. For example, a person who occupies the gender category of "woman" may wish or want to occupy the gender category of "man" without in fact being moved by its norms. In deeply misogynistic societies, for example, women may want or wish to be men because of the social privileges afforded to men. Consider pop songs like Beyoncé's "If I Were A Boy" or

Taylor Swift's "The Man" as examples. In both songs, the protagonist fantasizes about how much easier life would be for her if she were a boy or a man instead of a woman. In both songs, the protagonist seems *interested in* being subject to masculine norms rather than feminine ones; but in neither case are we led to believe that she *does* feel as if she's subject to them. These are not cases of cross-category norm responsiveness, but rather the desire or wish to occupy a different category than one currently occupies because of facts about that category. This need not involve any particular norm responsiveness. Moreover, desire to occupy a particular gender category can, and often does, happen in the case of those who experience *cross-category* norm responsiveness. For example, a transfeminine person who is responsive to norms of femininity might nevertheless desire very strongly to be a man, because life would be very much easier for her if she *were* to comfortably occupy the category of man.

Second, individuals may want to engage in gender-coded behaviors for reasons that have nothing to do with their felt responsiveness to gender norms. For example, in a context where skirts are coded feminine, a man may want to wear a skirt for any number of reasons, including pragmatic (because it is hot outside and skirts are less efficient at trapping body heat than pants), aesthetic (because a particular skirt looks good and accentuates his figure), political (because he wants to push back on gender norms), or hedonistic (because skirts are fun and twirly). He can easily feel this way and engage in this behavior without feeling like he *should* be wearing a skirt. There are many reasons to engage in behavior that happens to be recommended by some gender norm, and many or most of them are unrelated to one's felt responsiveness to that gender norm.

Engaging in behavior that specifically contradicts the norms assigned to you will typically involve a sense that one is *flouting* or *rebelling against* gender norms; that is, it will

involve what I have called the *phenomenology of rebellion*. Importantly, this can occur concurrently with the sense that one is obeying an incompatible gender norm. A person may feel responsive to both masculine and feminine norms at the same time. In wearing a skirt, for example, someone may feel that they are complying with a feminine norm to which they are responsive while also flouting a masculine norm to which they are (perhaps differently) responsive. Given the power of habituation and gendered training, many people who experience cross-category norm responsiveness will also experience an incompatible *in*-category norm responsiveness; however, they may experience one set of gender norms as more *authoritative* over their conduct. I expand on this in Chapters 3 and 4.

Third, and relatedly, individuals may want to rebel against the gender norms that others expect them to follow. Recall Jenkins' (2016, 411) example of a woman who feels responsive to the norm that women ought to shave their legs, but nevertheless does not shave her legs. This woman knows ("on some level") that she is subject to the feminine norm that recommends leg-shaving. However, she doesn't really want to shave her legs, and so she doesn't—even though she feels as if she ought. This woman experiences what I will call the "phenomenology of rebellion"; she feels responsive to a norm that she refuses to follow. Jenkins uses the example to explain how responsiveness to a norm comes apart from other kinds of desires and resultant behaviors. As Witt points out, "it is not possible to flout a norm that does not apply to oneself" (2011, 45).

These three desiderative phenomena about gender norms can, and often do, occur together with cross-category norm responsiveness. Many trans and gender-nonconforming people express or experience some combination of all of these phenomena when it comes to gender norms. Since I am not giving an account of what it is to be trans or gender-

nonconforming, I will not discuss whether any or all of these is necessary or sufficient for belonging to those categories. My aim in this chapter is to argue that the category-based view cannot explain cross-category norm responsiveness. The category-based view, however, *can* explain the three desiderative phenomena described above. One can want or wish to occupy a different gender category (and want or wish to be subject to different norms), want to engage in a behavior that is gender-coded irrespective of the categories with which it is associated, or want to reject or rebel against a norm which *is* associated with one's gender category, all while being all and only responsive to the norms which are associated with one's assigned gender category. As I have noted, each of these phenomena is *also* associated with cross-category norm responsiveness. The key point is that desiderative states and norm responsiveness come apart.

To recap: According to the category-based view, individuals are subject to gender norms because they occupy a particular gender category. On this view, then, a person's gender category membership determines which gender norms apply to them. The category-based view as defended by Witt (2011) can explain felt norm responsiveness in terms of *habituation*; since norms apply to an individual on the basis of their gender category, that person internalizes those norms and comes to follow the norms by habit. This view entails that individuals will be responsive only to those norms which are associated with their gender category, since only those norms can apply to them. Those who belong to the category *man* will be responsive to only masculine norms, those who belong to the category *woman* will be responsive to only feminine norms, and so forth.

Having identified the phenomenon of cross-category norm responsiveness, I will now show that it presents an objection to the category-based view.

There are, I have suggested, two broad desiderata for explaining gender's normative power over us. One is to explain the way agents are responsive to gender norms. The other is to explain the way that entities are evaluable under gender norms. In this section I will argue that the phenomenon of cross-category norm responsiveness reveals that the category-based view cannot meet the first desideratum.

Many dominant systems of oppression work in part by representing themselves as *natural* or *normal.* That is, it is a key feature of the perpetuation of harmful and hierarchical social practices such as gender and race that they do not appear to be *perpetuated* at all, but rather appear to arise naturally, or at the very least proceed unproblematically from natural features (see Collins 2000, Haslanger 2012, Ásta 2018). In the case of gender, this is achieved in part through the normative power of gender norms. The dominant gender categories of "man" and "woman" are taken to be natural kinds, and therefore the associated standards of femininity and masculinity are understood to be, at weakest, unproblematic trackers of essential features that naturally occur in sexually dimorphic human animals, and, at strongest, helpful normative guides for good behavior corresponding to the proper functioning of the two natural categories of human being. On either understanding, the pressure to be "normal" gender-wise affects individual agency and practical deliberations. Non-conformity is strictly punished, so, in general, individuals learn to calibrate their behavior relative to gender norms. As Witt argues, over time, individuals become *habituated* to this behavior, and do it without thinking.

This habituation, in turn, contributes to the appearance that gender is natural. A casual look at a world where many or most of the people visible in public life "do gender" habitually or without thinking will produce the observation that gender is intrinsic to many or most

individuals. Add to this the systematic destruction and often total erasure of gender-variant people and communities, and we get the following common pre-theoretical belief: binary, gender-conforming behavior is in some sense a natural or unavoidable element of human behavior. The normative pressure to conform to gendered standards creates gendered behavior and then makes itself invisible, such that the gendered behavior appears to arise without prompting. Pre-theoretical evaluations of gendered behavior tend to ignore this normative pressure, and instead see gendered behavior as *itself* innate or natural, obscuring the mechanisms which make it so.

As discussed in Chapter 1, there are two ways for socially grounded norms to apply to the individual: *thin ascriptivism* and *voluntarism.* For brevity, I will henceforth refer to *thin ascriptivism* as simply *ascriptivism.* Put briefly, ascriptivism is the view that norms apply to individuals because of the social practices of others, while voluntarism is the view that norms apply to individuals because of their own self-regarding attitudes or practices (e.g. choice, preference, self-identification, adoption, etc.). I will argue that, on either understanding, the category-based view cannot explain cross-category norm responsiveness.

First, an ascriptivist category-based view cannot explain cross-category norm responsiveness. On an ascriptivist category-based view, individuals will be only responsive to the norms corresponding to the category assigned to them by others. But this is obviously false. When a child whom everyone believes is a boy feels that they *ought not* be using the boys' room, or when they feel that they *should* look like the girls, they are demonstrating responsiveness to feminine norms which do not correlate with their assigned category. The category-based view cannot explain this. Cross-category norm responsiveness therefore represents a challenge to the category-based view's ability to explain *responsiveness.*

Importantly, however, a voluntarist category-based view does not fare better. Gender norm responsiveness does not consistently track self-identification, choice, or preference, but rather seems wholly distinct from it. Consider, for example, a case where voluntarism appears to have the advantage. A transgender man is classed as a woman in certain contexts, but nevertheless feels responsive to norms of masculinity. He experiences himself as responsive to norms that correlate with the gender category with which he has identified, rather than the category which others ascribe to him. However, as demonstrated by the experiences of Mock and Serano, many people experience a cross-category normative pull very early in life, long before they identify with or adopt any gender category other than the one they were assigned. A closer look at the case above, then, shows that does not so straightforwardly support voluntarism. For example, consider the following account by P. Carl:

> I had been scrutinizing masculinity my whole life, trying to perfectly replicate it in my gestures and clothes and physiques. I stayed very trim, wore only men's clothes, studied the latest short-hair styles, tried to keep the tenor of my voice low, and always played the roles that I thought men played. I earned. I mowed the lawn. I kept track of the finances. I filed the taxes. I shoveled the snow. I lugged the air conditioners from the basement to the bedroom windows every summer. I always drove. I was grossly deficient at housecleaning. I owned only one bathroom towel when my wife, Lynette, first moved into my bachelor pad. But I insisted that I was not a man. (Carl 2020b, 10)

Carl's responsiveness to norms of masculinity did not come as a result of his self-identification as a man. Rather the opposite: they were the impetus for it. After decades of experiencing norms of masculinity as guiding his behavior, Carl finally found himself unable to avoid the conclusion that he was a man. "And as much as I had done every single thing to look like a man and live like one," he writes, "I denied wanting to become one because I didn't want to become my father or lose my lesbian lover or be a failed feminist and intellectual" (2020a). For an experience like

Carl's, the voluntarist view gets the explanation the wrong way around. It is the norm-responsiveness that leads to the self-regarding gender practices, rather than the reverse.

Moreover, a voluntarist view fares worse with respect to explaining other kinds of gender norm responsiveness. As we saw in Chapter 1, a key advantage of ascriptivism is that it can explain the experience of feeling responsive to a norm with which one does not identify. It is demonstrable, for example, that people who reject gender categories and their associated norms on political grounds nevertheless sometimes feel as if they ought to follow certain associated norms. Voluntarism cannot explain this, as it grounds gender norm responsiveness in one's voluntary self-identification with or adoption of a gender category.

Furthermore, not everyone who experiences cross-category norm responsiveness identifies or comes to identify with the normative role associated with the norms to which they are responsive. Many gender-nonconforming people who intentionally and voluntarily take themselves to occupy a particular gender category may feel responsive to norms not associated with that category. Consider, for example, the rich history of butch women within queer cultures. Many butch women consider themselves to be women, and, moreover, are classed as women by the others in their social contexts. However, many butch women experience themselves as responsive to norms of masculinity, irrespective of their category membership or self-regarding practices.

In general, then, whether it is understood as voluntarist or ascriptivist, the category-based view about gender norms is ill-equipped to explain cross-category norm responsiveness. This is problematic for two reasons. First, trans and GNC people have a distinctive perspective on the domain of gender norms. This gives us an epistemic authority which is indispensable to constructing a theory of gender norms that meaningfully captures the way that they shape

agency. Many trans and GNC people have experiences of being subject to a variety of different normative gender standards at various times or in various contexts. Many of us also have the experience of failing to meet the minimum conditions for inclusion in any of the normative gender standards, by being excluded from gender intelligibility altogether. (I'll say more about this in 2.3.) Trans and GNC people as an epistemic community therefore experience gender norms "from all sides", as it were. If we are committed to understanding the distinctive ways in which gender norms guide behavior, as well as the distinctive ways in which they can oppress and harm, the perspectives of trans and GNC people must be centered.

Second, and relatedly: Attempts to understand gender phenomena within feminist philosophy are largely undertaken with explicitly liberatory aims in mind. That is, such projects are not merely descriptive; they are also *normative*, with the aim of identifying the distinctive harms caused by gender oppression in order to change them. (This project, at least, is understood thusly.) A feminist account of gender norms therefore must not exclude those who are most marginalized according to gender norms. The goal is to theorize gender oppression in a way that helps us to understand and undermine it. If our liberatory theories do not capture the experience of those who are most at risk, they are failing to meet this key desideratum.

In light of this consideration, I claim that the category-based view falls short of explaining gender's normative power over our behavior, because it cannot capture the phenomenon of cross-category norm responsiveness. Since this is a relatively common experience among trans and GNC people, the category-based view is insufficient to explain the way individuals actually respond to gender norms and incorporate them into their normative deliberations. I call this the *responsiveness objection* to the category-based view.

The second way in which the category-based view falls short is in its failure to capture the way that gender norms are enforced. In the next two sections, I will articulate this as wha I call the *evaluability objection.*

2.3. WHITE COLONIAL GENDER AND CATEGORY EXCLUSION

In the next two sections (2.3 and 2.4), I will give several reasons to think that the category-based view does not explain *evaluability* under gender norms. If we want to understand how entities in the social world come to be judged according to gender norms and punished for failing to meet normative gender standards, a focus on gender categories will miss the experiences of those who are often most targeted and harmed by these standards. In this section (2.3), I make a relatively broad argument that gender categories as a focus of analysis do not always help us explain gendered exercises of power, in part because this focus can easily ignore the experiences of those who are specifically *excluded from* inclusion in dominant gender categories. In the next section, I will draw on this broader analysis to specifically defend what I am calling the *evaluability objection* to a category-based view.

In her work on gender and coloniality, Lugones (2007, 2010) argues that gender categories are themselves a product of colonial thinking. The broadest version of the point is that we should understand "categorical, dichotomous, hierarchical logic as central to modern, colonial, capitalist thinking about race, gender, and sexuality" (2010, 742). That is, the claim is not just that gender categories are insufficient as explanatory tools for understanding the variety of gendered phenomena in the world, but rather that thinking of the social world primarily in terms of categories is a distinctive framework that proceeds from *coloniality,* a way of conceptualizing, structuring, and dividing the world that is fundamentally unequal, vicious, and

destructive. Specifically, coloniality posits a series of hierarchical, dichotomous categories: human/animal, white/non-white, man/woman, and so on. Categorization, according to Lugones, is part of the conceptual framework that works to enable colonialism and capitalism, by positing the world as something that is naturally measurable, quantifiable, and thus divisible (2007, 192). Gender categories are one features of this approach. Lugones therefore calls on a project of feminist resistance to specifically resist this Eurocentric epistemic framing, and instead to think what she calls "non-modern" terms. Rather than beginning from gender categories and hierarchies and attempting to theorize the different ways this has been done across history, she suggests that we begin by looking at the way non-colonial cultural contexts have understood the world, and their corresponding systems of social organization, most of which do not center or even involve individual gender categories. Only then can we begin to see that the colonial construction of gender is fundamentally a "disruption" of the self (2010, 749-50). This is central to Lugones' call for a "decolonial feminism" (2010).

We might be tempted to think, then, that a theoretical focus on gender categories aims to understand some of the social problems endemic to a world shaped by the legacy of colonialism. That is, taking seriously the claim that gender is a harmful colonial construction, we must therefore understand the significant impact that this construction has had on our world. As Lugones writes: "I am certainly not advocating not reading, or not ''seeing'' the imposition of the human/non-human, man/woman, or male/female dichotomies in the construction of everyday life, as if that were possible. To do so would be to hide the coloniality of gender" (2010, 750). Theorizing about gender categories therefore has its uses, insofar as we understand that those categories (and perhaps categorization itself) are not universal or trans-cultural, but rather are culturally and historically contingent, and fundamentally pernicious.

Indeed, most feminist metaphysicians have understood gender categories as both contingent and pernicious. In a (perhaps the) foundational work of analytic feminist metaphysics, Haslanger (2000) posits that gender is not just characterized, but rather *constituted,* by inequality. That is, the gender categories of *man* and *woman* are positions within a hierarchical class system, where those observed or imagined to have male bodies are classed as men and are thereby dominant or privileged, while those observed or imagined to have female bodies are classed as women and are thereby subordinated or oppressed. This approach is heavily influenced by the work of Mackinnon (1989), who argued that gender is constituted by sexual dominance-- specifically, the systematic sexual dominance of women by men. Thus, there is no such thing as "gender equality"; gender *is* an inequality. This approach has been taken up by many other prominent theorists in this tradition, such as Barnes (2019), Dembroff (2018b, 2020), and (with some caveats) Jenkins (2016, 2018). That is, much feminist metaphysics is committed to the basic view that dominant gender is unequal and harmful, and many such views hold that this inequality and harm is not just a contingent feature of gender but rather exists at its root.

But this is not sufficient to meet Lugones' challenge. By and large, feminist metaphysics has focused on the ways in which dominant gender is *binary, hierarchical,* and *patriarchal.* That is, it singles out and explores only one aspect of colonial gender: what Lugones calls the *man/woman* dichotomy. While this is certainly a fundamentally hierarchical and harmful system, it is not the only element of gender as a system of colonial power--nor is it the most basic. Lugones sketches the ways in which this gender binary in turn rests on a more fundamental division: the *human/non-human* dichotomy, where *humans* are white European colonizers, and everyone else is denied this status. Lugones calls this the "central dichotomy of colonial modernity" (2010, 743). As colonialism spread across the world, colonizers created a conceptual

division between themselves--white, European, heterosexual, Christian, "rational," subject/agent

males--and everyone else. To be sure, part of this story is the way in which white

European *women* are constructed as passive, weak, and fundamentally in service to white

European men. This is the core of the man/woman dichotomy, which Lugones acknowledges as

a real and harmful element of the gender system. However, and crucially, occupation of these

dichotomous categories is only available to those who are already sorted as *humans*--that is,

those who are white and European. Only colonizers count as *men* or *women.* Colonized people

are "bestial and thus non-gendered" (2010, 743); they are *males* or *females,* as in the

classification of non-human animals, but not *men* or *women*. She writes:

> Beginning with the colonization of the Americas and the Caribbean, a hierarchical, dichotomous distinction between human and non-human was imposed on the colonized in the service of Western man. It was accompanied by other dichotomous hierarchical distinctions, among them that between men and women. This distinction became a mark of the human and a mark of civilization. *Only the civilized are men or women.* Indigenous peoples of the Americas and enslaved Africans were classified as not human in species—as animals, uncontrollably sexual and wild. (2010, 743, emphasis added)

And, even more bluntly:

> The semantic consequence of the coloniality of gender is that 'colonized woman' is an empty category: no women are colonized; no colonized females are women. Thus, the colonial answer to Sojourner Truth is clearly, 'no.' (2010, 745)

According to Lugones, then, understanding gendered harms in terms of the harms specifically

done to those who occupy the dominant category *woman* is ignoring the majority, and the most

brutal, of the gendered harms that there are. These are visited upon those who are ejected from a

gender category altogether.

We can see a similar theme in the work of Spillers (1987), who articulates the ways in

which the Black enslaved body, particularly the female body, is *ungendered* through the legacy

of chattel slavery. Spillers understands gender here as a function of *domesticity;* the body

becomes gendered through domestic roles and relations within the family and household. Therefore, the disruption of familial bonds and the objectification and mutilation of the Black body under chattel slavery "ungenders" Black people, in part by dehumanizing them and reducing "body" as a coherent subjecthood to "flesh," a physical object that can be manipulated and abused (67-68). Thus, Spillers writes: "Under these conditions, we lose at least gender difference in the outcome, and the female body and the male body become a territory of cultural and political maneuver, not at all gender-related, gender-specific" (1987, 67). Here, again, we see the following pattern: *gender = humanity* and *humanity = whiteness*.

For Spillers, those who are ungendered are still targeted in normatively gendered ways. However, this plays out in ways that are not at all predicted by the gender categories we might want to (perhaps wrongly) impose. She argues that feminist theorizing of the female body has systematically forgotten that the "African female subject, under these historic conditions, is not only the target of rape--in one sense, an interiorized violation of body and mind--but also the topic of specifically externalized acts of torture and prostration that we imagine as the peculiar province of male brutality and torture inflicted by other males" (68). That is, the female Black enslaved "flesh" is subjected both to harms that the category system would code as feminine (e.g. rape, forced impregnation, domestic servitude), and harms that it would code as masculine (e.g. "externalized" physical violence, brutality, and torture).

Drawing from Spillers' notion of *ungendering*, Pritchett (2019) discusses the way that enslaved female bodies were ungendered through the *commodification* of their reproductive capacity, in much the same way that humans commodify the reproductive capacities of animals. At the same time, enslaved women were not permitted to marry--a distinctively human capacity. Like Spillers, Pritchett points out that Black women were prohibited from taking part in the the

domestic performativity of gender. "One of the fundamental markers of gender performance for women was marrying a man and taking care of a home, especially in the Antebellum South, and black women did not have the option to do so" (Pritchett 2019). The legacy of these practices is far-reaching and persistent. Tellingly, Pritchett calls attention to a common Jim Crow-era division of bathrooms into the following categories: "men," "women," and "colored." These examples starkly demonstrate the ways in which Black people are actively and materially through the enforcement of law *barred* from taking part in the social practices and performative acts that mark one as a member of a gender category.

The example of ungendered Jim Crow bathrooms is also noted by Gossett (2016). Gossett points out that recent anti-trans bathroom legislation in the United States is intimately connected with the history of bathroom policing against people of color. Various efforts to criminalize transness cannot be understood apart from "how sexual difference itself has been weaponized as an instrument of antiblack and colonial power and of white sovereign embodiment" and "how the Lacanian 'sexed body' is always already a racialized body and a colonized body, and how Black and/or indigenous peoples have always figured as sexual and gender outlaws to be disciplined and punished" (Gossett 2016). This discussion points to a phenomenon articulated at greater length by Bey (2017); transness and Blackness are conceptually intertwined, as both represent disruption, disturbance, and displacement of the "purity" of full humanity, as distinguished by white, binary, colonial gender categories.

In this context, Gossett's language of "outlaws" is telling. It mirrors Bornstein's (1994)'s *Gender Outlaw*, an anti-essentialist (white) trans manifesto about transitioning and living at the margins of gender categories. However, where Bornstein's use of the term is triumphant--reflecting an identity of active and creative gender resistance to the binary, as one who initially

occupied a position of significant privilege *within* it (that is, initially being read as a white man)--

Gossett's use describes a much more brutal status, one marked by discipline and

punishment while simultaneously being constituted by ejection from binary categorization. To be

a racialized "gender outlaw" is to be disciplined and punished specifically as an outsider to a

gender category, in order to demarcate the boundaries of the "normal"--i.e. the intelligibly

binarily-gendered and white. Thus, as Patricia Hill Collins writes,

> Heterosexuality itself is constructed via binary thinking that juxtaposes male and female
> sexuality, with male and female gender roles pivoting on perceptions of appropriate male
> and female sexual expression. Men are active, and women should be passive. In the
> context of U.S. society, these become racialized—White men are active, and White
> women should be passive. Black people and other racialized groups simultaneously stand
> outside these definitions of normality and mark their boundaries. (2000, 83)

The uses of "outlaw" here are potentially ambiguous. A more pointed term for the phenomenon

of trans and GNC exclusion from dominant gender categories can be found in Wilkins (1997).

Wilkins refers to themself, and other trans or GNC people who are not consistently and

comfortably categorized as men or women, as *gendertrash.* This terminology calls attention to

the ways in which many marginalized people are *rejected* or *thrown out* from the dominant

categories. According to Wilkins, *gendertrash* is consistently the target of gendered exercises of

power, by everyone from medical professionals to teenage boys on the street corner. Living in

the world as someone who is not intelligibly gendered as a member of a dominant, binary

category means being punishable under *all* of the gender norms at once; one is neither feminine

enough nor masculine enough, and thus one is failing on *both* standards. This leaves one open to

punishment from multiple fronts.

My point here is not to argue that there is no theoretical benefit to focusing on gender

categories. I recuse myself entirely from that question. Rather, I want to focus on the ways in

which membership in a particular gender category is not necessary for being the target of a gendered exercise of power. Broadly speaking, gender categorization itself reproduces a broader legacy of injustice--not just by dividing people into binary categories of *dominant* and *subordinate,* but by dividing people into *those who occupy dominant gender categories because they are treated as fully human,* and *those who don't because they are not.*

In what ways, then, does the focus on gender categories fail to explain the effects of gender norms beyond the boundaries of those categories? In the next section, I will show how the discussion here motivates what I am calling the *evaluability objection.* Entities are not held as evaluable according to gender norms on the basis of gender categories, as the category-based view entails. There are a variety of examples which can demonstrate this. I will focus on the three examples named above: *dehumanization, ungendering,* and *gendertrash.*

2.4 THE EVALUABILITY OBJECTION

A brief recap: I have articulated the category-based view as a way of understanding how the assignment and enforcement of gender norms translates into action. That is, it aims to explain how, as Witt puts it, we are "responsive to and evaluable under" particular gender norms. In 2.1 and 2.2, I motivated the *responsiveness objection.* I argued that a category-based view cannot explain how agents are *responsive to* particular gender norms, because it cannot explain *cross-category norm responsiveness*. Drawing on the ideas discussed in 2.3, I will now motivate the *evaluability objection.* I take this objection to show that a category-based view cannot explain how entities are *evaluable under* particular gender norms.

It is certainly acknowledged within analytic feminist philosophy that gender norms apply differently on the basis of other intersecting social factors, such as race, disability, and so forth.

One way of making sense of this from within the category-based frame has been to hold that binary gender categories are applied differently to different people on the basis of *other social categories*. For example, Dembroff writes: "The construction of binary gender kinds (men/women), the latitude individuals have within them, and how individuals are socialized into them, vary dramatically across intersections with other social identities, such as race, class, and disability" (2020, 17). In a footnote, they hint at a stronger version of this claim: "note that other, intersecting social identities–such as race, class, sexuality, and disability—often disqualify or prevent persons from attaining these so-called 'natural' features, rendering them 'unnatural' and devalued (2020, 16). While this seems in spirit compatible with much of what was discussed in 2.3, note that it still proceeds from the assumption that *gender* is the baseline category, and "other" social identities can intersect with gender and "disqualify" individuals from certain kinds of normatively bounded ("natural") gender presentations.

Similarly, Haslanger writes of the differential impact of gender norms based on the other categories to which one belongs:

> Imagine race, gender, and other social positions to be like gels on a stage light: the light shines blue and a red gel is added, and the light shines purple; if a yellow gel is added instead of the red, the light shines green. Similarly, gender is lived differently depending on the racial (and other) positions in which one is situated. Just as a light may appear different colors depending on which combination of gels it is filtered through, the gender norms for Black women, Latinas, and White women differ tremendously, and even among women of the same race, they differ depending on class, nationality, sexuality, religion, historical period, and so on. (2012, 9)

On this kind of analysis, one is a member of various categories--gender, race, class, disability, sexuality, and so on. As Haslanger's metaphor demonstrates, these categories are understood to *overlap* and affect how the others are experienced and interpreted. When we are trying to

understand how norms work, then, we must understand this overlap as part of the story.

Bernstein (2019) understands this as a metaphor specifically intended to communicate the *intersectional* nature of social categories, such that one cannot understand identities as *added together* but rather as *combined* or *mutually constitutive.* That is, intersectionality suggests "a complex, non-additive, intermingling metaphysical relationship between different aspects of social identity" (Bernstein 2019, 324). Both Haslanger and Dembroff explicitly point to this kind of understanding here.

First, notice that this way of thinking about things is grounded in the assumption that individual gender categorization is a basic part of evaluability under gender norms. Other social structures, such as race, are understood to interact with one's gender category, such that one's evaluability under gender norms is mediated by one's race (and perhaps vice versa--although things are rarely put in those terms within this literature). In terms of gender norms, we might infer the following: The gender norms that apply to someone, and their experience of being responsive to and evaluable under these norms, will also be mediated by other intersecting identities. A category-based view need not, of course, hold that the norms which apply cannot *vary;* it is only committed to the claim that, whichever norms apply, they do so *on the basis of* one's gender category. There is plenty of room within this view for an intersectional approach to the actual mechanics of gender norm enforcement, understood as a kind of "overlapping" of different categories. On this view, if the norms associated with *woman* if one is Black, middle-class, non-disabled, and living in Seattle, are wildly different than those associated with *woman* if one is white, poor, using a wheelchair, and living in Cairo, it is because the different categories *overlap* and affect the normative impact of one another.

In 2.3, I demonstrated that many marginalized people are *excluded* from membership in dominant gender categories. Based on these examples, I argue that the category-based view faces a real problem, even if it can accommodate intersectional "overlap." Recall that, for a view to explain *evaluability,* it must be able to capture the ways that agents are candidates for criticism under and enforcement of gender norms. The category-based view holds that evaluability under gender norms is grounded in the application of a gender category. But being excluded from membership in a dominant gender category does *not* entail being excluded from evaluability under the associated norms. Rather, as Lugones, Spillers, and Wilchins each point out, those excluded from dominant categories are the targets of harsher evaluation and more brutal enforcement according to those norms. I call this the *evaluability objection* to the category-based view.

Let me spell out the three versions of this objection in a little more detail. First, Lugones (2010) argues that the gender categories assigned to white colonizers are not assigned to the colonized; colonizers are *men* or *women,* while the colonized are not. There is some unclarity between Lugones' view on this in her (2007) and the later (2010). In the former, Lugones may be read as arguing that colonized people occupy subordinate racialized gender categories (e.g. colonized male, colonized female). In the latter, Lugones seems to be claiming that the colonized have *no* gender categories at all (analogous to Spillers' account of the *ungendering* of Black enslaved female "flesh"). In either case, however, the objection is the same: The norms under which everyone, colonizer and colonized, is evaluated, are centered on the normative categories assigned to the *colonizers.* That is, *both those included in and those excluded from the dominant categories assigned to white colonizers are judged by the standards associated with those categories.* Lugones writes:

I propose to interpret the colonized, non-human males from the civilizing perspective as judged from the normative understanding of ''man,'' the human being par excellence. Females were judged from the normative understanding of ''women,'' the human inversion of men. From this point of view, colonized people became males and females. Males became not-human-as-not-men, and colonized females became not-human-as-not-women. Consequently, colonized females were never understood as lacking because they were not men-like, and were turned into viragos. Colonized men were not understood to be lacking as not being women-like. What has been understood as the ''feminization'' of colonized ''men'' seems rather a gesture of humiliation, attributing to them sexual passivity under the threat of rape. (2010, 744)

Here again we see the relationship between gender and *full humanity*; only the fully human are afforded the gender categories of "man" and "woman," and only white Europeans are fully human. (This also mirrors Mills' (1997) argument that the European philosophical understanding of the "social contract" only includes those classified as full humans--i.e. white men.) Note also that the brief discussion of *colonized men* here further undermines the analysis of different social categories "combining" to make different effects. We can't understand the humiliation of colonized men by combining race with gender, or even gender with colonialism. Rather, colonized men are understood as *outside* the category "men" but nevertheless failing at standards of masculinity, in part by having characteristics attributed to them that are incompatible with the standards set out for dominant white men, such as *sexual passivity*. This is a far more nuanced phenomenon than can be explained by simply mixing together different social categories. For this reason, I argue that a category-based view simply does not have the resources to explain Lugones' account of how colonized people are evaluated under gender norms, because it does not explain how gender norms can apply to someone in virtue of their *failing* to occupy the associated category.

Spillers' (1987) account of the *ungendering* of the Black enslaved body raises a similar issue. Through the legacy of chattel slavery, enslaved people were violently removed from

classification as gendered. Moreover, despite the legacy of ungendering, we can clearly see the ways in which Black females were and are held to what Deliovsky (2008) has called "normative white femininity." As Collins (2000) has argued, Black women are historically judged as *unfeminine* in virtue of being too "aggressive" or "dominant"--something Spillers also notes (1987, 74). That is, Black women are constructed as failing to meet the norms of a feminine gender standard that treats White women as the "benchmark" for womanhood (Deliovsky 2008, 58). Relatedly, Black men are consistently constructed as either *hyperagressive* (and thus frightening) or *insufficiently masculine* (and thus feminized or passive); on either understanding they are perceived as *incorrectly* masculine relative to the White-centric norms which are understood to delineate masculinity (see Curry 2017). This is done in conjunction with what was, according to Spillers, a history of active *ejection from* membership in the associated dominant categories. The category-based view cannot explain how those who are *ungendered* are nevertheless treated as evaluable under gender norms.

Relatedly, through Wilchins' narratives of existing as "gendertrash," we can see how exclusion from dominant categories is related to harsh enforcement of associated gender norms. Although Wilchins and many other trans and GNC people are rendered unintelligible by local gender classifications, the resulting status of "gendertrash" is nevertheless sufficient for harsh evaluation on the standards of *either* masculinity or femininity. That is, they are both insufficiently masculine *and* insufficiently feminine--or perhaps *too* masculine and *too* feminine at once. A person ejected from dominant categories as *gendertrash* is evaluated, not just for failing on a particular normative gender standard, but for failing on *all* of the standards.

On each of these examples, the gender norms that apply to the people in question are not determined by the gender categories that they occupy, but rather by their *exclusion* or *ejection*

from a gender category. There are, I think, many other versions of this point. For example, Thorn (2021) points out that, under fatphobia, body size can disqualify one from being included in dominant gender categories. However, it's not as if fat people are no longer evaluable under the norms of the gender categories from which they are excluded or ejected. Rather, they are harshly evaluated according to those norms. In this section, I have focused on what I take to be particularly clear and stark cases: Lugones' account of *dehumanization,* Spillers' account of *ungendering,* and Wilchins' discussion of *gendertrash.* The first two are cases raised by feminists of color in reference to racialized gender under coloniality and chattel slavery. I raise these particular cases, in part because the deep historical imbrication of gender and race is under-theorized in analytic feminist metaphysics, and in part because the associated harms are particularly widespread, systematic, and brutal. Relatedly, I raise the experience of trans and GNC people because we are often used as exemplars of how gender norms can cause harm. On the whole, however, the point is overdetermined.

In this chapter, I have largely been focusing on the normativity of gender norms in terms of their effect on individual people. This is so in part because, if the aim is to understand how the landscape of practical reason and choice is shaped by gender norms, it is often more straightforward to talk in terms of *individual agents* and their normative deliberations. That is, when we are talking of being *responsive to and evaluable under* a norm or set of norms, it is sometimes easiest to see this at the level of the individual, whose responsiveness and evaluability can be more easily articulated.

However, according to Lugones (2007), this is not the only, or perhaps even the primary, level at which colonial gender norms are enforced on the colonized. In what follows, I will try to reconstruct Lugones' argument that the harmful normative practices of the colonial/modern

gender system are not merely about individual evaluability, but rather affect groups, social structures, and societies. Put differently, she argues that colonialist attitudes take "macro-level" social entities, such as groups and their collective attitudes and epistemic practices, to be *evaluable* under colonial gender norms.

Lugones articulates the way in which colonial gender norms were and are enforced at a structural level on non-colonial *egalitarianism;* that is, ways of understanding the world and organizing the social life of a group that do not incorporate a binary, hierarchical, category-based social metaphysics. She discusses both *nongendered egalitarianism,* such as societies like the Yoruba which, prior to colonialism, did not organize their societies around anything like gender roles; and *gynecratic egalitarianism,* wherein many Indigenous societies rested on spiritual beliefs that elevated what we might call the feminine, believed that the primary organizing force in the universe was female, and so forth. To control these egalitarian societies, colonialism did not merely create a new hierarchical gender system. It had to disrupt the entire social order and replace it with one conducive to the binary, category-based way of thinking that it imposed, in order to then mobilize the justification of a patriarchal, hierarchical system of social organization. These egalitarian societies are therefore subject to what Lugones calls "the gendered construction of knowledge in modernity"--a colonial attack on *entire epistemic frameworks* that are potentially contrary to the project of justifying colonialism. Thus, the "inferiorization of Indian females was thoroughly tied to the domination & transformation of tribal life"--a project which requires not just brutality, but the more subtle normative undertaking of "image and information control" (Lugones 2007, 499). For colonialism to justify itself, it was necessary both to disrupt the traditions which value the feminine, *and* to undermine the systems of social organization and ways of knowing that support this.

On this picture, it is not enough to say that the differential experiences Black and Brown colonized people and societies have of gender norms is due to the *overlap of categories* such as gender and race. The reality is broader and more complex than that. According to Lugones, colonial gender norms functioned to shape the very nature of the communities and epistemic frameworks that *constitute* non-colonial societies and individuals. Gender's normativity here is therefore far more complex than simply *those in the category (*or even *overlapping categories) are evaluable under the norm.* Rather, gender norms are woven into the colonial attitude, on which certain ways of existing in the world--for individuals as well as cultures and groups--are treated as normatively superior because they are friendly to a hierarchical, patriarchal framework, and others (i.e. egalitarian societies) are subordinated and/or brutalized because they are not. This is true not just because egalitarian or gynecratic practices promote the interests of individual females or women, but because a colonialist attitude and way of life is fundamentally incompatible with the prioritization and valuation of features that are, under that metric, coded feminine. The category-based view is simply too univocal to capture these nuances of gender normativity--in part, I suggest, because it proceeds from the experiences of those whose gender categorization proceeds from racial and colonial privilege.

In short, throughout the legacy of colonialism, gender norms are enforced not just on individual people, but on entire groups, cultures, and ways of life. The effects of this are visible primarily at the macro-level. The issue at hand is not just that race affects experiences of gender norms, like overlapping gels on a theater light. Rather, on Lugones' analysis, we should understand gender's normative power as a standard that guides how things ought to be in the "civilized" world--i.e. the white colonial world. Many non-colonial social organizations are incompatible with that standard, and as such are coded as normatively inferior. This is not

straightforwardly a matter of *man/woman,* but is gendered nevertheless, in a variety of complex

and intersecting ways--including, for example, ways where someone can be interpreted as

masculine and subject to masculine norms (as in the case of colonized men) and still brutally

subordinated.[3]

In short, I argue that we have reason to question a category-based view as a framework

for understanding gender norms (among other kinds of gender phenomena), because we have

reasons to think that such a view is primarily suited to explain the experiences of individual,

privileged people. This kind of framing cannot capture the way in which gender norms are

enforced on the dehumanized, the ungendered, or the "gendertrash;" nor can it begin to explain

the ways in which colonial gender as a normative framework shapes ways of living and knowing,

as in cases where it works to justify the violent restructuring and devÁstation of gynecratic or

egalitarian ways of life.

2.5 CONCLUSION

I have shown that the category-based view is insufficient to do the explanatory work

necessary to understand the normative power of gender norms. I have articulated two ways in

which it fails to do this. First, it cannot explain the way that agents are *responsive to* gender

---

[3] Curry (2017) defends an analogous point about the subordination faced by Black men. According to Curry, this subordination is explained, not by combining *Blackness* with the supposedly fundamentally powerful *maleness* as he argues that gender studies has historically done, but by understanding Black maleness as a distinctively subordinated and distinctively *gendered* or *sexed* subjectivity. I think again that a category-based view cannot capture this, at least not as it has been traditionally defended in gender metaphysics. This may be because that tradition owes much to Haslanger (2012), who famously argued that gender was defined by hierarchy, where one is a man iff one is systematically *privileged along some dimension* on the basis of one's observed or imagined sexed body (234). The broader point, however, is that if we want to understand how gender norms work to subordinate, gender categories seem insufficient as an analytical frame, because they are simply not flexible enough to capture the nuanced varieties of oppression gender norms can justify. More on this in Chapter 3.

norms. Experiences of *cross-category norm responsiveness,* where individuals experience themselves as responsive to gender norms which do not match their gender category, demonstrate this failure. I called this the *responsiveness objection* to a category-based view. Second, it cannot explain the way individuals, as well as groups, cultures, and other macro-level social phenomena, are *evaluable under* gender norms. I gave three examples where individual people are excluded or ejected from particular gender categories, but nevertheless evaluated according to the associated norms; Lugones' (2010) example of *dehumanization,* Spillers' (1987) example of *ungendering,* and Wilchins' (1997) example of *gendertrash*. I then briefly discussed the phenomenon of *macro-level social entities* such as groups and societies being evaluable under gender norms, and argued that a category-based view is incapable of explaining evaluability except in the case of individual people.

I have therefore given many compelling reasons why a category-based view specifically cannot explain how gender norms attach to agents and other entities in the social world. In the next chapter, I will lay out a *traits-based view,* which I will argue *can* do this.

3.0. INTRODUCTION

In analytic gender metaphysics, gender categories are often understood as a foundational element of the ontology of the social world. Gender categories offer us a clear story of how the social world is divided, in a way that helps us see how it is also *stratified.* As Frye wrote, "It is a fundamental claim of feminism that women are oppressed" (1983, 2). If we are to theorize outwards from this "fundamental claim," then perhaps we should proceed from the understanding that categories like "woman" and "man" are ontologically important, because they meaningfully divide the world into *those who are oppressed* and *those who are not*. The assumption, then, is that if we are going to understand the oppressive social power of gender norms, we should begin with the assumption that *gender categories* are central.

It is perhaps for this reason that feminist metaphysicians have by and large understood gender norms as applying in virtue of the gender categories that an individual occupies. For example, recall that Witt writes: "Women and men are responsive to their gendered social roles [sets of norms] and evaluable by others in relation to those roles *just by virtue of their social position occupancy*" (2011, 29, emphasis added). Haslanger writes: "The classification of

features as masculine or feminine...depends on prior social classifications" (2012, 46). And Dembroff writes that "norms...are imposed on people based on their gender classification" (2018, 3). I elaborated on each of these important views in Chapter 1, and argued that they share a commitment to what I called the *category-based view*. Recall:

> **The category-based view (CBV):** Gender norms apply to individuals only if and because those individuals belong to the associated gender categories.

The category-based view has several important theoretical advantages, most which I have already discussed. It gives us a clear and parsimonious account of the grounds on which gender norms are assigned, in a way that is consistent with several intuitively important features of gender norms: the fact that they are *socially constructed,* rather than natural, essential, or innate; the way in which they are often *nonconsensually assigned,* and exert force over us whether we want them to or not; the fact that they are *differentially applied* to different people, often on the basis of *sexed embodiment;* and the way in which, combining several of these features, they work to enable and justify oppression. However, in Chapter 2 I argued that a category-based view is too narrow, because it proceeds from the experiences of those who are already privileged within dominant gender categories. I have suggested that this is reason to reject the view.

However, rejecting a category-based view leaves us somewhat adrift. If we want to understand how gender norms exert normative force over us, we need an account of *how gender norms attach to us* that can capture the important features listed above, without centering the experiences of those who are privileged by a category-based view.

I will argue that we should understand we understand gender norms as socially assigned to individuals on the basis of perceived traits. Put very simply, traits are descriptive features of individuals (Mikkola 2016). Notice, however, that we sometimes also perceive larger-scale

social entities, such as *groups* or *societies,* to have traits (see Zheng 2016, Hofstede 2001). Call these *macro-level social entities.* On a traits-based view, particular traits are gender-coded-- usually as *masculine* or *feminine*. Since dominant systems of power often treat gender as binary, gender norms are accordingly expected to bifurcate into mutually exclusive, coherent clusters. That is, masculine traits are expected to occur together with all and only other masculine traits, and feminine all and only with feminine. The resulting clusters may take on further ontological significance, perhaps as *gender roles, gender identities,* or even *gender categories.* However, I will hold that the gender-coding of the trait is explanatorily primary to, and can occur without, this further significance. I want to remain neutral on the relationship between normative role and gender category--as well as, in general, on the nature of gender categories themselves. To reiterate, I am not giving an account of *what gender categories are*. Instead, I want to understand the way that particular gendered forces shape agency. Following Mikkola (2016), I take this project to both be possible without, and to have been inhibited in many ways by a theoretical preoccupation with, category-focused analysis.

I proceed as follows. In 3.1, I will articulate six key desiderata for the project at hand. The first two will call for the positive view to have the explanatory power to capture the experiences of marginalized people with respect to *responsiveness* and *evaluability* as I identified them in Chapter 2. The remaining four desiderata will consist of the "intuitively important features" that I identified above. In 3.2, I will lay the groundwork for the positive view by sketching some upshots of the desiderata. In particular, I will argue that a commitment to *social construction* and *nonconsensual assignment* suggests that we should adopt what I call *internalization/ habituation* about gender norms. In 3.3, I will outline the basics of my positive *traits-based view*. I will explain in detail what a "trait" is, and explain how constellations

norms *cluster* around particular socially relevant traits. In 3.4, I will explore the traits-based view further by showing, in turn, how it meets each of the six desiderata. This will also help me explore the view further. In 3.5, I consider some important objections to the view. In 3.6, I conclude.

3.1. THE DESIDERATA

3.1.1. In this section, I will identify six key desiderata for an account that aims to explain the way gender norms are assigned. The first two desiderata involve the view's ability to explain the experiences of marginalized people. That is, the view should be able to (1) explain the way that marginalized people are *responsive to* gender norms, and (2) explain the way that marginalized people are *evaluable under* gender norms. These correspond to the two major objections I raised against the category-based view in the previous chapter. The remaining desiderata are as follows: (3) *social construction* (the account must be consistent with the basic metaphysical position that gender is *socially constructed*); (4) *nonconsensual assignment* (the account must capture the way that gender norms exert power over us without our participation or consent); (5) *differential assignment and embodiment* (the account must capture the intuitively obvious fact that gender norms apply differently to different people, often on the basis of their observed or imagined sexed body); and (6) *oppression* (the account must enable us to explain why many domains of gender norms are either intrinsically oppressive or serve to justify and enable oppression). I will discuss and briefly motivate each of these desiderata in turn, and then elaborate on some of the commitments they generate for the positive view I will give.

3.1.2. Given my discussion in chapter 2, the first two desiderata need little further elaboration. Recall my formulation of Witt's question: What explains responsiveness to and evaluability under gender norms? In that chapter, I discussed the way in which established views have failed to answer this question with respect to the experiences of marginalized people. First, I discussed *cross-category norm responsiveness,* or the experience of being responsive to some gender norm which ostensibly hasn't been assigned to you. Thus, I argued that the category-based view does not explain *responsiveness*. Second, I discussed the impact of gender norms on Black and Brown people, societies, and ways of life, through the legacy of colonialism and chattel slavery. Thus, I argued that the category-based view cannot explain *evaluability*. Based on these examples, I argued that the category-based view fails.

Since I have argued that the view *fails* on these grounds, a successful view *should* be able to explain these phenomena. That is, it should be able to meet the following two desiderata:

> (1) *Responsiveness.* A view of gender norms must be able to explain how marginalized people are *responsive to* gender norms.
> (2) *Evaluability.* A view of gender norms must be able to explain how marginalized people are *evaluable under* gender norms.

This way of putting the desiderata is, admittedly, quite pointed. It might be objected that this is too narrow. The view should be able to explain why *everyone* is responsive to and evaluable under gender norms. The latter is, of course, correct. However, I still explicitly frame the issue in terms of marginalized people, for two reasons. First, I don't think it's necessary to specify that those who are privileged within dominant gender categories ought to be covered by the view. My discussion of desiderata 3-6 capture what I take to be the most intuitively important features of gender norms *for everyone.* I do not think there is risk of leaving out those who are typically centered in such accounts. Second, I *do* think it's necessary to specify that the experiences of

those who are *not* privileged within dominant gender categories be included. It is (clearly) too easy to ignore us otherwise. Doing so constitutes a serious theoretical oversight, because such experiences provide a unique perspective on dominant gender norms--one that can better help us understand how they work. Let me say a little more about why.

In general, those who are oppressed have a kind of epistemic privilege with respect to dominant social practices (see Janack 1997, Collins 2000). To navigate a world that is not built for your benefit--more to the point, one that is built specifically to constrain, control, and harm you--requires the ability to understand multiple intersecting and sometimes conflicting levels of social structure, as well as to inhabit the perspective of the privileged in order to navigate their exercises of power within that structure. The oppressed must therefore develop knowledge that is not available to those for whose benefit the world *is* built. To understand how to unpack those systems of oppression, then, we need this distinctive knowledge. We need to know what the harms are like and how they are visited on the disempowered in order to ameliorate them.

This is particularly apparent in the case of gender norms. Those who comfortably and consistently occupy dominant, hegemonic, white-centric gender categories can certainly experience some of the oppressions and harms they can perpetuate. However, this is only one dimension of those harms. Of course it is hard to be consistently gendered as a privileged woman in a privileged man's world, but at least it is a positionality for which there is a space, a name, and an intelligible set of gendered expectations (however self-contradictory or full of double-binds they may be). If we want to understand how gender norms shape behavior; if we want to know how they can weigh heavy against other considerations, even when we ignore, dislike, or actively disavow them; if we want to see the deepest and most profound harms that gender norm enforcement can cause; we need to listen to those whose experiences involving navigating

multiple, intersecting sets of norms, those for whom the dominant world has no space, no name, and no way to understand--and for whom *expectations nevertheless abound*. We need to center the experiences of those who have "reckoned with the boundaries and the dimensions of masculinity and femininity in ways [the privileged] have never had to" (Coates 2016). As Dembroff writes, "The best way to see gendered reality is to be a Tiresias, throbbing between overlapping but radically segregated worlds" (2018a).

I therefore take the first two desiderata to be especially important for two reasons. First, the experiences they reference are incredibly relevant and offer distinctive insight. Second, they are often overlooked within the tradition of feminist metaphysics. A theory of how gender norms attach to entities must be able to answer Witt's question specifically as it applies to the experiences of marginalized people.

3.1.3. The third desideratum for a theory of norms is *social construction:*

> (3) *Social construction:* the account must be consistent with the basic metaphysical position that gender is *socially constructed*.

Put differently: in explaining gender's normative power, the account must not appeal to intrinsic, essential, or pre-social features of individuals, groups, or humanity as a whole. In Chapter 1, I discussed at length the nature of "social construction" as it is being used in this project.[4] I positioned social constructionism in opposition to *biological essentialism* about gender, or the view that there is an "intrinsic essence" to individuals' genders, caused in some way by their

---

4 From Chapter 1: "social constructionism in feminist philosophy holds that gender is not a natural or fixed part of human life. Rather, gender arises out of contingent historical practices of assigning social significance to certain features. Those features themselves may be natural or fixed parts of human life, such as reproductive capacities (Witt 2011, 27-29; Ásta 2018, 3); however, the associated social significance is contingent, and varies across cultures and contexts."

biology (their chromosomes, hormones, reproductive roles, etc.). I did not give an extended

defense of social construction, nor do I intend to do so here. The importance of a social

constructionist approach has been thoroughly defended by others, and does not need more

attention here (see, among others, Beauvoir 1948, Frye 1983, Butler 1990, Collins 2000,

Lugones 2007, Haslanger 2012, Ásta 2018). It is a basic commitment of this project that a

feminist account of gender norms should maintain social constructionist commitments.

I will pause here to discuss the relationship between social constructionism and the other

desiderata I have identified. Specifically, I will discuss one objection to the possibility that they

can be addressed by the same theory. It is sometimes argued that social constructionism in

general is incompatible with trans and GNC experience. The thought is this: If social

constructionism is true, then individuals should be responsive to the norms that are socially

assigned to them. But trans and GNC experience seems to demonstrate that this is not the case.

Therefore, social constructionism is false.[5]

One prominent defender of this view is Serano (2007). In explaining her own experiences

of cross-category norm responsiveness, as well as other kinds of gendered feelings she

experienced from a young age, Serano defends what she calls the *intrinsic inclinations model*.

According to Serano, motivations to behave according to standards of masculinity and

femininity, which often appear at "a very early age" and continue into adulthood, constitute

---

[5] Another way of interpreting this conflict is to organize these premises differently, as follows: If cross-category norm responsiveness is a legitimate experience, then social constructionism is false. Social constructionism is true. Therefore, cross-category norm responsiveness is not a legitimate experience. Something like this seems to be going on in the most vitriolic versions of so-called "gender-critical feminism," or self-described feminist views (such as those defended by Raymond (1979), Bindel (2007), and more recently, Lawford-Smith (forthcoming)). I do not engage with this argument here, in part because I do not think it is worth attention; in part because it has already been thoroughly discussed and answered by others (see, for example, Stone 1987; Finlayson, Jenkins, and Worsdale 2018); and in perhaps largest part because to respect and take seriously trans and GNC experience is also a basic commitment of this project.

evidence that "certain expressions of masculinity and femininity represent deep, subconscious inclinations similar to those of sexual orientation" (98). Serano therefore argues that masculinity and femininity are deep, subconscious features of individuals (98-100). On Serano's view, cross-category inclinations are evidence that responsiveness to gender norms is not a response to social conditioning, but is already present prior to that conditioning. She writes:

> Many girls who are masculine and boys who are feminine show signs of such behavior at a very early age (often before such children have been fully socialized with regard to gender norms), and generally continue to express such behavior into adulthood (despite the extreme amount of societal pressure that we place on individuals to reproduce gender expression appropriate for their assigned sex). This strongly suggest that certain expressions of femininity and masculinity represent deep, subconscious inclinations similar to those of sexual orientation. (2007, 98)

Serano seeks to explain cross-category norm responsiveness by *jettisoning* social construction. She is quite explicit that this is her goal, writing that "a strict social constructionist model does not account for exceptional gender expression" (2007, 98)--by which she means gendered behavior that conflicts with one's apparent socialization and training.

An immediate problem with this model is that it assumes that the inclinations or behaviors in question and the gender norms with which they are associated are conceptually inseparable. Serano argues that children have deep masculine and feminine tendencies, and that these "gender inclinations are, to some extent, intrinsic to our persons" (99). "Intrinsic" here means something like *acquired without socialization;* she contends that motivations to behave in a masculine or feminine way "occur naturally (i.e. without social influence)" (99). However, Serano's argument only shows that certain *inclinations* occur naturally or without social influence. There is no reason to think that these inclinations are inherently gendered. On the contrary, feminist philosophers at least as far back as John Stuart Mill (1997 [1865]) have convincingly argued that masculinity and femininity are socially constructed normative

standards, rather than natural features of individuals. As Haslanger (2012) puts it, "Particular traits, norms, and identities, considered in abstraction from social context, have no claim to be classified as masculine or feminine" (46). This is evident in the fact that gender norms vary widely across cultures, and even across contexts within the same culture. Here is a plausible interpretation of the phenomena Serano points to, which is also consistent with social constructionism: children have certain inclinations, some of which are gender-coded *post hoc*.

A worry remains. Part of Serano's argument is this: if masculinity and femininity are socially constructed, how do they hold sway over young children before socialization can properly take hold? This is an empirical matter, and I can only gesture at the answer here. Constructed or not, gender runs deep. It is so thoroughly embedded in social experience that children are unlikely to acquire gender-coded concepts without also knowing that they are so coded. Moreover, gender training begins *early*. Children in the United States, for example, learn to identify the difference between masculine and feminine between 18 and 24 months, and they typically know their assigned gender by the age of 3 (Mayo Clinic Staff 2017). Gender is generally identified long before this; consider again the phenomenon of *gender reveal parties,* in which gendered features are identified before an infant is even born. I argue, then, that there is no intelligible time before gendered socialization takes hold. I conclude that we will not need to reject social constructionism to explain cross-category norm responsiveness.

3.1.4. Accepting *social construction* implies we should accept another desideratum. In Chapter 1, I argued that social constructionists should be (thin) *ascriptivists*[6] about gender norms. Recall:

---

[6] I have been calling thin ascriptivism simply *ascriptivism* for brevity.

Ascriptivism: Some gender norms apply to subject S because of the practices of others with respect to S, rather than because of S's voluntary self-identification, choice, or preference.

Ascriptivism is a way of satisfying the fourth desideratum for a view of gender norms:

(4) *Nonconsensual assignment:* the account must capture the way that gender norms can exert power over us without our participation or consent.

Nonconsensual assignment is not the same thing as ascriptivism. On ascriptivism, gender norms can in principle be assigned consensually. For example, ascriptivism can accommodate *self-ascription,* where the feature that others are trying to track when determining whether or not some norm applies to S is *whether S would prefer those norms be applied.* This is the case in some trans-inclusive contexts. More in this in 3.4.5.

As discussed in Chapter 1, there are excellent reasons to think that gender norms in dominant contexts are nonconsensually assigned. We generally do not have a say in which gender norms others apply to us, nor do we seem to have much say in the norms to which we feel responsive. We can see this because we have almost no control over such things. Simply refusing to acknowledge a gender norm's power over us does not mean that we are no longer responsive to its pull, or evaluable under it by others. If this were the case, then eliminating the normative power of gender norms would be as simple as refusing to acknowledge that power. To be sure, such a refusal can be a kind of resistance. For example, drawing from Spillers' (1987) discussion of violent racist *ungendering,* Bey (2022) argues in favor of an *active, subjective* project of ungendering oneself as a form of resistance to hegemonic gender (Bey 2022, 28-29). Similarly, Dembroff (2018, 2020) argues that adopting a *non-binary* or *genderqueer* gender identity can constitute resistance to dominant gender and its associated norms. However, this alone does not seem to remove us from the so-called "normative umbrella" of those norms (Witt 2011, 44). If it

did, then life for many trans and gender-nonconforming folks might be much easier. Including *nonconsensual assignment* in our theory of gender norms will therefore speak to desideratum (6): *oppression.* More on this in a moment.

3.1.5. Our fifth desideratum is also related to *oppression*. Gender norms are not assigned equally to all. One function of such norms is to *stratify* society, such that the normative permissions allotted to some are far greater and allow for far more latitude and control than for others. Often, this stratification seems to be connected to the kind of *sexed body* that one has. For example, consider again the phenomenon of a *gender reveal party.* During an ultrasound on a pregnant person, a fetus is identified as having either a penis or lacking one. Based on this identification, the parents will organize a social gathering centered around the "reveal"--a cutting of a cake, a release of balloons, an explosion, or some other equally dramatic and potentially dangerous event. What is revealed is traditionally one of two colors: blue for a fetus observed to have a penis, pink for a fetus not so observed. Such parties are usually deeply gendered in other ways. For example, cakes are often decorated with an oppositional gender-normative theme: "baseball or bows," "wheels or heels," "tractors or tiaras" (Incoherent Queer Screaming, 2020). Based solely on this relatively minor piece of information, a constellation of gender norms attaches to the infant before they are even born.

As discussed, feminist philosophers have often explained this attachment, in terms of gender category. However, this is not explanatorily sufficient. As discussed in Chapter 2, the gender norms which are assigned to those who are ejected from dominant categories, or "ungendered," may also be differentiated based on embodiment. For example, Spillers articulates the way that the Black enslaved female body ceases to become properly a *body* and instead

becomes *flesh*, which is then subjected to, among other things, norms of reproduction and sexual

engagement with white men (1987, 76). Similarly, Lugones describes gendered violence against

colonized people:

> The colonial ''civilizing mission'' was the euphemistic mask of brutal access to people's
> bodies through unimaginable exploitation, violent sexual violation, control of
> reproduction, and systematic terror (feeding people alive to dogs or making pouches and
> hats from the vaginas of brutally killed indigenous females, for example). The civilizing
> mission used the hierarchical gender dichotomy as a judgment, though the attainment of
> dichotomous gendering for the colonized was not the point of the normative judgment.
> (Lugones 2010, 744)

Again, Lugones is clearly arguing here that it is not the *gender category* of the colonized females

that determined the norms to which they were held. Nevertheless, *sexed embodiment* certainly

plays a role in how the norms attach and are enforced. This seems true in a wide variety of

incredibly disparate contexts and situations. I suggest, then, that a theory of how gender norms

attach to entities would be deeply unintuitive if it could not capture:

> (5) *Differential assignment and embodiment*: the account must capture the intuitively
> obvious fact that gender norms apply differently to different people, often, though not
> always, on the basis of their observed or imagined sexed body.

This description aims to be consistent with the position that biological sex is socially constructed

(a la Butler, 1993); that is, I don't want to claim that the "sexed body" is in any sense natural or

coherent as a category. Rather, I use this phrase to denote the fact that certain bodily features

have traditionally been *grouped together* under the umbrella of "biological sex," including but

not limited to: genital structure; secondary sex characteristics (presence or absence of breasts,

density and location of body hair); features presumed to be correlated with, but not directly

connected to secondary sex characteristics (height, weight, body shape); chromosomal makeup;

ability to seed or bear children. "Observed or imagined" is a phrase I borrow from Haslanger

(2012). It denotes that the important factor for gender norm assignment is not whether the

individual in question *really* has the feature(s), but whether they are *perceived* to have the features by others. Given the desideratum of *nonconsensual assignment*, this is an important caveat. (Consider, for example, a doctor who mistakenly identifies a fetus on a sonogram as having a penis, when in fact it does not. A gender reveal party in this case would attach masculine norms to that fetus. This may be rescinded when the mistake is corrected; however, the initial attachment holds until that point.)

It is clear that those who are observed or imagined to have certain sexed bodies are often thereby assigned certain gender norms. However, it is also clear that there are contexts in which sexed embodiment may be, as Ásta puts it, "highly irrelevant" (2018, 76). For example, consider contexts where gender norms are assigned to individuals based on their gender *presentation* rather than their sexed embodiment. In the context of a gay club, for example, a gay man, a straight man, and a drag queen may all be perceived to have the same sexed body. Moreover, each of these individuals may be perceived as, and perceived themselves as, *a man.* However, very different normative gender standards will apply to each, and each will likely be responsive to different norms. In these cases, there is likely some feature of the individual that fixes gender norm assignment; however, that feature does not appear to be either a gender category or a particular kind of sexed embodiment.[7]

---

[7] Note that, on one reading of Ásta's (2018) view, each of these individuals *does* occupy a different gender category. Ásta discusses the ways in which gender properties are conferred on individuals in a context based on the attempt to track some underlying property, such as sexed embodiment, self-identification, role in social organization, and so on. If we read *gender property* here as *gender category,* we might think that the context of the gay bar has (at least) the following three categories: *gay man, straight man,* and *drag queen.* However, I think this is not the right reading. Ásta's account of a *gender property* does not straightforwardly map onto what I am calling a *gender category.* I discussed in Chapter 1 the ways in which I think Ásta defends a category-based view. However, I do not think she is committed to the claim that there are a multiplicity of context-dependent *gender categories,* as I argued that, on her view, gender classifications are broader and more stable than *gender properties.* Moreover, whatever Ásta's commitments, I will argue that positing a multiplicity of context-dependent categories is not a good move for a category-based view. More on this in 3.5.1.

3.1.6. The final desideratum is politically motivated. This project, like most work in gender

metaphysics, is explicitly liberatory. The aim is not simply to describe or explain what exists, but

to target that description and explanation at the features of the world that are distinctively

harmful or oppressive, in order to identify the levers of change. To that end:

> (6) O*ppression:* the account must enable us to explain why many domains of gender
> norms, particularly the hegemonic norms that are operative in dominant contexts, are
> either intrinsically oppressive or serve to justify and enable oppression.

Importantly, this desideratum is framed to allow that *not all* domains of gender norms are

intrinsically oppressive or serve to justify and enable oppression. For example, some domains of

gender norms, such as those constructed in subaltern queer and trans communities (see Bailey

2011, Bettcher 2013), might arise out of a need to make intelligible the experiences of

marginalized people, and, concordantly, can specifically work in opposition or resistance to

dominant, hegemonic norms. This can serve as an important step to undermining the power of

dominant norms. Therefore, a view aimed at liberation ought not preclude the possibility of such

resistance. I discuss this further in 3.4.6.

Taken together, these six desiderata will help frame the positive view I am going to

construct. Notice here that the category-based view as I have framed it already meets several of

these conditions, in particular desiderata 3-6. In 3.2, I will draw on some of the resources those

view can provide, in order to lay the groundwork for the positive view I will give in 3.3

3.2. THE DESIDERATA AND THE CATEGORY-BASED VIEW

Let us consider the interplay between the factors listed in 3.1, with respect to a basic line

of questioning for this inquiry: What explains the normative force of gender norms? Witt's

earlier work (2011) has helped frame much of our investigation thus far. In chapter 1, I argued that the view defended in this work is not interested in what metaethicists call *robust* or *authoritative* normativity, but rather in a somewhat thinner social and psychological notion. However, in a later study of social normativity, Witt (2019) makes some headway on more substantive normative questions. She motivates this investigation with the following question: "How do we explain the idea that obligations or "ought tos" are somehow implanted in the social world, attached to its positions, practices, and structures?" (2019, 4) This framing focuses on *norm responsiveness,* as does much of Witt's view.[8] Explaining this is one of our desiderata. But what is the relationship between norm responsiveness and the other desiderata--in particular *social construction* and *nonconsensual assignment?* Put differently: How can a socially constructed phenomenon such as gender norms create reasons for agents to act, even when those norms are assigned without the agents' preference or consent?

Witt gives a simple and elegant answer. Her account of social normativity relies on Aristotelean notions of *function* and *excellence.* According to Aristotle, individuals have social functions which are inherently normative. For example, a *flute player* has the function of *playing the flute,* which "is associated with an excellence in relation to which the function is performed well (or badly)" (2019, 7). According to Witt, embedded in the Aristotelean concept of an *excellence* is the notion of *habituation.* To be properly excellent at a thing, one must habituate oneself to it. That is, one must do it regularly and often, such that one does it well without thinking about it. This understanding of Aristotle is perhaps best summed up by Will Durant, in a

---

[8] Recall from chapter 2 that one is *responsive to* a social norm when one experiences the norm as guiding one's behavior, which could mean that one experiences any of a "full range of possible reactions to a norm....from compliance to critique" (Witt 2011, 33). Again, this is compatible with, and indeed indicated by, what I have called the *phenomenology of rebellion*; the sense that one is *violating* a norm.

quote which captures the central idea so well that it is often misattributed to Aristotle: "We are what we repeatedly do. Excellence, then, is not an act, but a habit" (Durant 2006 (1926), 98).[9] Put differently, success at one's function is not merely something evaluated relative to one's performance, but rather an "acquired disposition" (Witt 2019).

For Witt, then, each social position has a function, and therefore an excellence. This is not to say that each social position is in itself excellent or normatively good (as is clearly false), but rather that each social position generates a role, or collection of norms, relative to its primary function. Each role can be performed well (or badly), and thus is normative for its constituents (2019, 10). To perform one's role well--to instantiate excellence relative to, say, the role of "teacher", "mother," or "woman"--one must acquire the relevant disposition; one must be *habituated* to follow the norms. The most excellent of mothers, for example, is fully habituated to this role and acts in accordance with the norms of motherhood, not because she thinks carefully about each norm and follows it out of conscious compliance, but because it is her habit; it is part of her.

Witt's Aristotelean account aims to explain how individuals come to be responsive to gender norms. If we combine this with her account in *Metaphysics of Gender,* we get the following picture: Someone is assigned to a gendered social position. Each gendered social position has a function. For Witt, since these functions originated with the human need to reproduce, the social positions are assigned on the basis of one's reproductive capacity (e.g. ability to bear, or seed, children). As she puts it, "Being a man and being a woman are social positions with bifurcated social norms that cluster around the *engendering function*" (2011, 40).

---

[9] This mistake is made even by Aristotelean scholars, to the point that some such scholars have recommended not correcting it. See Herron 2013 or Caelan 2017.

However, she acknowledges that the roles adapt to varying social contexts, and thus are both

conceptually distinct from and considerably more complex than the basic reproductive functions

which gave rise to them. That is, "the actual content of the social roles is variable" (2011, 40).

Call these two bifurcated norm clusters "masculinity" (the norms relevant to the role "man") and

"femininity" (the norms relevant to the role "woman"). According to Witt, individuals are

evaluable under these norms as soon as they are assigned the role (at birth, sometimes before).

Over time, they become habituated to their role, and come to act in accordance with its norms

without thinking. This involves being *responsive* to its norms (often also without thinking).

This understanding of gender norms closely echoes Haslanger's account of gender norms

as relating to the "proper function" for particular gender kinds (2012, 43). She writes:

> But if females are expected to perform the role of mothering and to perform it well, then
> rather than coerce them to fulfill this role, it is much better for females to be motivated to
> perform it. So the *norms must be internalized,* that is, they must be understood as part of
> one's identity and defining what would count as one's success as an individual. Ideally,
> one will develop unconscious patterns of behavior that reinforce the role in oneself and
> others and enable one to judge others by its associated norms. (Haslanger 2012, 10,
> emphasis added)

In addition to the significant similarity to Witt's view, notice also Haslanger's language of

internalization. This terminology points to a nice theoretical coherence across disciplines. The

notion of *habituation* fits well with what is known in the social sciences as the *internalization*

hypothesis. The internalization hypothesis holds that the motivational power of norms is

explained by the fact that individuals *internalize* the norms of their local social group (Scott

1971, Sripada & Stich 2007). According to the internalization hypothesis, individuals acquire

norm responsiveness from their surroundings through a process of *social reinforcement.* They

learn what a norm recommends by connecting it with some *sanctions*--i.e. rewards for

compliance and punishment for non-compliance. Through a process of social learning, sanctions

(both reward and punishment) encourage norm-compliant behavior while discouraging non-compliance. On this view, a norm counts as *internalized* to the extent that the sanctions which helped to install it no longer need to be regularly applied to ensure compliance--that is, when the norm is followed at a temporal distance from reinforcement (Scott 1971, 88). The internalization hypothesis explains why individuals treat social norms that are unique to their societies, contexts, and identities as authoritative over their conduct, and follow them even when there are no sanctions present. That is, it explains *norm responsiveness* in terms that are consistent with *differential assignment.* Moreover, according to some theorists, it explains the "intrinsic motivational effects" that socially constructed norms appear to have on the individual (Sripada & Stich, 2006).[10]

The internalization hypothesis is an empirically grounded analog of Witt's *habituation.* Early proponents of the internalization hypothesis even constructed it in direct opposition to *voluntarism* (Keeley 1973). Understood together, *internalization* and *habituation* nicely capture the desiderata listed above. Internalization is consistent with norms being *socially constructed* and *nonconsensually assigned,* while habituation explains how they are *relevant* to individuals. Internalization explicitly positions us to explain *differential assignment and embodiment;* if norm responsiveness is a function of social practices like learning and training, then gender norms may be assigned differently on the basis of different features, such as embodied features. Given this theoretical agreement across disciplines, and the agreement of both views with the desiderata

---

[10] Recall from the introduction that the "founder" of the internalization hypothesis, Scott (1971), specifically resists this interpretation, arguing that the internal states of norm followers are irrelevant to an empirical study and that we should instead speak only in terms of *behavior*. Behaviorism has since largely fallen out of favor in the social sciences, which might explain the difference in commitments between these accounts. In either case, however, *internalization* itself need not posit any internal states to be well-matched to *habituation*. It can merely function as the "external analog."

listed above, I suggest that we should understand gender norm responsiveness as a function of what I will call *internalization/habituation:*

> **Internalization/habituation:** In response to sanctions) individuals *internalize* gender norms (they respond to them even when they are not being enforced). Over time, they become *habituated* to following them (they do it by habit without thinking).

Witt's version of *habituation* of course includes a commitment to her version of the *category-based view*. However, I argue that we can understand its appealing features independently of this commitment. The basic idea of *internalization/habituation* is just this: through regular performance, individuals come to respond to gender norms even when they are not being enforced, because they come to respond to them by habit, or without thinking. We might also think that, to do justice to the view's Aristotelean roots, we should be able to explain how this process relates to the notion of a *function*. But this does not need to involve a category. For example, a society which does not have a category for *servants* can nevertheless have someone who performs the function of service. Similarly, Scott (1971) seems committed to a category-based view when he writes that "as sex roles become differentiated, [a boy] will learn there are some acts for which his sister will be rewarded but for which he will be punished, and *vice versa*" (56). Again, despite the category-specific language, I argue that the intuitively appealing features of this view can be maintained without category-based commitments. Specifically, I will draw on Scott's account of *vicarious learning* to explain differentiated norm internalization without appealing to gender categories. I will say more about this in 3.3.

In preparing to construct the positive view, I will also call attention to the relationship between *nonconsensual assignment* and *differential assignment and embodiment.* That is, our positive view needs to explain how gender norms attach differently to different individuals, often on the basis of their perceived sexed bodies, without their consent.

Category-based views are incredibly well-positioned to handle some elements of this. If gender categories are nonconsensually assigned to individuals on the basis of their perceived sexed bodies, and gender norms come with gender category membership, then we have a clear story about how those norms attach to those perceived to have differently sexed bodies.

However, I think this story has an extra, unnecessary step. To discuss this, I draw attention to a point made by Dembroff in their discussion of what they call the "dominant Western gender system." According to Dembroff, one of the axes of this system is:

> The teleological axis: Someone's gender, by virtue of nature, determines a range of social, psychological, and physical features—e.g., sexual desire, occupation, family role, attire, comportment, personality features—that they either *must (are determined to) or ought to have*. Males naturally must or ought to have masculine features, females naturally must or ought to have feminine features. (2020, 15, emphasis added)

Dembroff is speaking here of the differential normative expectations we place on people relative to gender--that is, gender norms. They articulate this in a way that nicely captures both nonconsensual assignment and embodiment. As discussed, Dembroff elsewhere claims allegiance to a category-based view, writing that "norms...are imposed on people based on their gender classification" (2018, 3). In this passage, however, notice the language of *males* and *females*. Just previously in the same work, they suggest that, on the dominant ideology they are describing, "men have a male biological sex and women have a female biological sex" (2020, 15). That is, in describing the teleological axis relative to *males* and *females* (rather than *men* and *women*), Dembroff appears to speaking in terms of *biological sex* rather than gender category. They further note that biological sex is, *in most cases,* "understood as a feature determined by or reducible to external genitalia" (15, fn70). According to this view, then, the assignment of gender norms happens on the basis of biological sex, which is in turn understood as determined by or reducible to *external genitalia*--a paradigmatic feature of sexed embodiment.

This discussion gives us a clue as to the direction we should go. Elsewhere, I have read Dembroff as holding that a category-based view is necessary to explain gender norm assignment. However, in this passage, they do not refer straightforwardly to gender categories. Rather, they are pointing to a phenomenon wherein certain gender norms are assigned on the basis of a biological sex classification, which in turn reduces down to a single, embodied feature. But this raises a question: *What work is the classification doing in this process?* Why should we think that the gender (or sex) category is the key to assignment, *and not the feature itself?*

I suggest that we cut out the middleman. Rather than holding that gender norms attach to gender categories which are then assigned to individuals on the basis of their observed features, we can simply hold that gender norms are assigned to individuals *on the basis of those observed features.* That is, to explain how gender norms attach to entities, I propose that we shift focus from gender *categories* to gendered *traits.* The resulting *traits-based view* can pay heed to each of the six desiderata I have identified, as well as providing further attractive theoretical benefits. In the next section, I will elaborate.

3.3. THE TRAITS-BASED VIEW

To recap: I have identified some problems with the category-based view and articulated several desiderata for a view that will take its place. What we need is a view that can explain how gender norms attach to entities in the world. This view should capture how marginalized people are responsive to and evaluable under gender norms, and that they are socially constructed, assigned nonconsensually and differentially (often on the basis of perceived sexed embodiment), and, in many or most cases, oppressive.

I argue that we ought to adopt a *traits-based view* about gender norm attachment. What determines gender norms assignment is not the gender category that an individual occupies, but rather the *traits* that an entity is observed, imagined, or otherwise perceived or believed to display. It does not matter for my purposes whether the entity in fact *does* have the traits. What matters is how and whether those traits are perceived by the relevant others. In very broad terms: Traits are normatively gender-coded in various contexts, generally as *masculine* or *feminine.* As Mikkola puts it, "certain evaluations (like being judged to act, be or appear "feminine") can be seen to covary with particular descriptive traits" (2011, 69). If one is perceived to have or express a gender-coded trait, they are thereby subject to the associated norm. Where norms "cluster" around a particular social function, certain traits are expected to occur together with other traits; where this does not happen, punishment can follow.

3.3.1. Traits are non-relational, descriptive features of entities in the world. My understanding here is much indebted to Mari Mikkola's trait/norm covariance model (Mikkola 2011). Mikkola distinguishes between *descriptive traits* and *evaluative norms*, where traits "describe 'the way the world is'" and "include physical and anatomical traits (e.g., chromosomes, ovaries, testes, genitalia, body shape and size), one's appearance (e.g., one's clothing, make-up, haircut, amount of body hair), roles (e.g., whether one undertakes caretaking roles, engages in childrearing tasks) and self-conceptions (calling oneself a woman or a man)" (76-77). Mikkola is here describing the kinds of traits that *individuals* can have. However, I argue that entities other than individuals, such as groups, societies, cultures, and epistemic frameworks--what I call "macro-level social entities"--can also be interpreted as have traits. For example, these entities might be interpreted to have have physical traits (e.g. a nation-state has a trait of occupying a certain geographical

area); behavioral traits (e.g. we might understand U.S. foreign policy as having a trait of *aggression*); shared beliefs, values, or self-conceptions (e.g. a religious community has a trait of *believing* particular things); and so on.[11]

I understand the notion of a *trait* to be very thin. That is, traits are *bare descriptive facts* about some entity. For example, when I speak of a trait of *self-conception* like "calling oneself a woman or a man" (Mikkola 2011, 77) I certainly don't take this to be anything as substantive as a *gender identity* or *gender category*. Rather, I mean the bare descriptive fact that the individual, in fact, calls themself a woman or a man. Relatedly, notice that some traits, like *aggression,* are more like collections of traits. Call these *aggregate traits.* Properly speaking, aggregate traits are not "traits" in my sense, but rather are a handy way of describing a collection of traits that either commonly occur together or are perceived to do so. For example, the aggregate trait of *aggression* is really a collection of traits such as *being willing to attack other entities, exerting force in pursuit of one's interests,* and so on. These descriptive facts are perceived to occur together often enough that we have a name for the collection: *aggression.*

Many traits are gender-coded in various contexts. As Mikkola points out, some traits, like *being short-sighted*, are not gender-coded in any or almost any contexts (2011, 77). Other traits are widely gender-coded; *having ovaries*, for example, is coded feminine in almost all contexts.[12] Still other traits are only locally gender-coded. For example, in some contexts, *wearing jewelry* is coded as feminine; that is not the case elsewhere. However, *having ovaries* is coded as feminine near-globally. If an entity is interpreted as having a particular trait that is gender-coded

---

[11] This may be because the entity itself is perceived to have the trait, or because the individuals of which the entity is composed are perceived to largely share the trait. The truth of the matter doesn't depend on your view of collective agency, but rather on what is *perceived* to be true of these macro-level entities. It is therefore an empirical question which I don't address here.

[12] I discuss some exceptions to this, as pointed out by Bettcher (2013), in 3.4.6.

in their context, that entity is thereby evaluated under the associated gender norm. For example, a person who is interpreted to have the trait of *wearing jewelry* is thereby judged to be normatively feminine in many parts of the United States; a person who is perceived to have the trait of *having ovaries* is thereby judged to be normatively feminine in most contexts. In general, applications of gender norms follow from an initial perception of the entity as having particular, gender-coded traits.

It might be objected that it is not possible to draw a clean distinction between descriptive and evaluative here. Many gendered norms, particularly norms of appearance, are closely connected with descriptive features. Take, for example, "being skinny".[13] "Skinny" is ostensibly descriptive, but it also seems to carry some normative "baggage." That is, in some contexts, "skinny" is normatively feminine; however, it might be argued that this is not a matter of *gender-coding,* but rather is just part of the connotation of "skinny" in those contexts. This raises questions about the possibility of "pure" descriptive terms and concepts in the realm of gender. Examples like "skinny" appear to operate as *thick* terms or concepts. That is, they express both descriptive and evaluative content. Given how deeply gender is embedded in our thought and language, won't all or most descriptive terms and concepts which are gendered will have normative baggage? This is an interesting question that I aim to explore in further research. However, the question of whether there are "purely" descriptive gendered *terms or concepts* comes apart from the question of whether there are descriptive *features* which are distinguishable from their gender-normative upshot. The former is arguable, but I only need the latter. In fact, the normative connotation of terms or concepts like "skinny" might just be the normative coding I am talking about at work. Perhaps the traits have, in some contexts, become so closely

---

[13] Thanks to Charlotte Witt for this discussion and example.

connected with the norms that we cannot understand or speak about them separately. This is far from inconsistent with my view.

3.3.2. A key feature of this view is that gender norms tend to *cluster together*. That is, if a trait is gender-coded, it is often thereby socially expected to occur together with other, similarly-coded traits--masculine with masculine, and feminine with feminine. This social expectation may be enforced on those who display traits that do not cluster in this way, on a spectrum of severity ranging from mild surprise to brutal punishment. While many feminist philosophers have explained this clustering in terms of *gender categories,* as discussed, I argue that this is not necessary. Rather, we might understand gender norms as clustering around traits that have particular *social functions*, with the understanding that those social functions will vary widely and need not be unified by one's gender categorization.

Recall here Witt's discussion of norm habituation as related to Aristotelean excellence. She writes that "Being a man and being a woman are social positions with bifurcated social norms that cluster around the *engendering function*" (2011, 40). However, if we reject the category-based view, and remove the social positions of *man* and *woman* from this picture, we get the following: there are "bifurcated social norms ...that cluster around the engendering function." *Engendering* here is understood as the social process of "differentiating women from men according to their reproductive functions" in order to codify the social significance of those reproductive functions (32). Again, if we remove *women and men* from this picture, we get the notion of *differentiation according to reproductive functions*--that is, essentially *differential assignment and embodiment.*

Witt understands "reproductive functions" as involving a kind of *aggregate trait* including a collection of bare descriptive facts, such as producing certain kind of gametes (egg or sperm), possessing a gestational environment for offspring (having a uterus), and so on (36). According to Witt, these aggregate traits have a particular social function; they are necessary for a society to persist (100). Recall that, in later work, Witt also explains the normative force of social norms in terms of *functions*. That is, social positions have functions, and to be excellent at a particular function, you must be habituated to the norms that cluster around the function (Witt 2019). Again, we can understand this without involving the *social position.* That is, we can just hold that there are particular functions with associated excellences and thus clusters of norms. Putting this together, we have the following: Gender norms are social norms that cluster around the aggregate traits of *reproductive functions,* because those reproductive functions have a *social function:* they contribute to the persistence of a society.

I propose, then, that we adopt Witt's *habituation* view, but without including *social positions.* Instead, we should understand *functions* as involving collections of *traits* which are socially significant in the relevant context. Moreover, I suggest that we leave it open as to which traits can be socially significant for gender norm clusters. While Witt thinks that the *engendering function* is a necessary part of differential gender norm assignment, other philosophers disagree. For example, Ásta writes that, when assigning gendered constraints and enablements, "in some contexts, people are trying to track a sex assignment, in others a role in societal organization, a bodily presentation, a role in the preparation of food, a role in biological reproduction, a role as a sexual partner, and so on" (2018, 74). Notice that each of these examples, too, can be interpreted in terms of trait clusters with a particular social function. A role in the preparation of food, for example, includes a constellation of traits such as *planning meals, shopping for groceries,*

cutting vegetables, and so on. Collectively, these traits have a social function: those who have them meet the nutritional needs of the people in that society. As Witt points out, social functions are best performed by those who, habitually and without thinking, have all of the associated traits.

On a traits-based view, gender norms will tend to cluster around traits that are socially significant. These significant traits are what I call the *core traits* of a particular cluster. Generally speaking, core traits will have the most "weight" in determining whether an individual is expected to present all of the associated traits. For example, if some norms cluster around a core aggregate trait of *reproductive function*, then the traits which *constitute* reproductive function-- such as literally possessing the ability to reproduce, by bearing or seeding children--will outweigh other, more peripheral traits when determining how to evaluate someone.[14] The peripheral traits will be those that are perceived to *aid in* or *contribute to* the function; that is, those that are desirable, but not necessary. For example, if a core trait is *being able to bear children,* peripheral traits may include *being risk-averse* (to be more likely to survive pregnancy), *being inclined to nurture* (to better raise those children), as well as various traits that are considered desirable to *pass on* to those children (such as whatever traits are considered attractive or otherwise independently normatively valuable in the context).

---

[14] In some cases, such as this one, it is not always feasible to measure or directly observe the actual core traits. The contexts at hand may therefore assign gender clusters based on *markers* of these traits, with a built-in defeasibility condition. Those contextual markers are usually taken to reliably and non-accidentally predict the relevant characteristics, and will therefore only be investigated when suspicion is raised--if indeed they can be directly investigated at all. For example, norms that cluster around a core trait of the reproductive function *being able to bear children* may not rest on the actual perception that an individual can in fact bear children, as this can be remarkably difficult to verify. In lieu of this, the norms may temporarily attach to individuals according to other, more easily observable traits that are believed to "mark" that ability, such as *having breasts* and *being of childbearing age.* For more on the contextual nature of operative gender markers, see Ásta (2018).

3.3.3. Clustering is not a necessary feature of gender norms. Traits can be coded as masculine or feminine independently of how, or whether, they cluster around a particular social function. For example, gender-coding might denote a normative relationship between the trait and some overarching cultural ideals, such as the "eternal feminine." The trait could therefore be normatively feminine without being expected to occur together with other feminine traits. Moreover, clustering is not fundamentally harmful or oppressive. Consider, for example, Lugones' example of Indigenous *gynecratic societies,* in which we can understand gender norms as clustering around social functions that are treated as roughly *equitable,* and also relatively fluid: "most individuals fit into tribal gender roles on the basis of proclivity, inclination, and temperament" (2007, 199-200). Consider, also, the possibility of norm clusters in subaltern trans-inclusive communities, where the core trait of some clusters is *self-identification* (see Bettcher 2013). The harmfulness or oppressiveness of the cluster will depend on what the core traits are, and whether and how what I call *cluster coherence* is enforced.

I have argued that part of the function of most dominant gender norms is to *naturalize* certain behaviors--to make it seem as if gender is a unified natural kind that arises spontaneously from essential facts about the world, such as the so-called "biological features" of individuals (Lugones 2007, Haslanger 2012, Ásta 2018). In these contexts, then, there is a corollary expectation that the clusters are *mutually exclusive and coherent*. In order to successfully function to naturalize gender in this way, it must appear that individuals "naturally" express either all and only the traits associated with masculinity, or all and only those associated with femininity. That is, the expectation that masculine-coded traits occur with masculine-coded traits is backed up by a normative presumption of "natural"--meaning, of course, "normal" or "appropriate." Many dominant contexts work this way; non-coherence is often called "unnatural"

(abnormal or inappropriate) and severely punished. Since physical features such as sexed bodies are believed to be relatively stable (as traits go), those whose physical features are gender-coded in one way but are perceived to have other traits coded in a different way, such as trans and GNC people, are particularly vulnerable. Similarly, since a major relevant social function in *colonial* contexts is to *demarcate whiteness,* "being white" or "having white skin" will be core traits on the clusters in those contexts. Therefore, those who are perceived to have sexed bodies, but who are not interpreted as white, will be brutally punished for a fundamental kind of non-coherence. More in this in 3.4.

This view openly draws on elements of certain category-based views, most prominently Witt's (2011). This is so because, as argued in Chapter 1, these views have many important insights. The crucial difference between the category-based view and the traits-based view concerns the order of explanation. On a category-based view, the explanatory story starts with the existence of the categories; they are the basic unit of gendering, so to speak (hence the terminology "category-based"). Individuals are sorted into a category and then assigned norms and expected to display traits based on that assignment. On this view, then, the explanation begins by positing an entity called a gender category. Individuals either have the ontological status of belonging to that category, or they don't. Typically (though not always) the status is conferred by others on the basis of one's perceived sexed body. The category has associated norms. If one belongs to the category, one is assigned the norms.

On a traits-based view, however, no such ontological claims are made. The explanatory story starts with the normative gender-coding of traits, typically as *masculine* or *feminine*. Insofar as those traits are observed or imagined to occur, whether in individuals or in macro-level social entities, those individuals or entities are thereby evaluated as masculine or feminine.

Often, norms cluster around particular core traits that serve social functions. Sometimes, social practices will expect or mandate coherence and punish non-coherence to varying degrees of severity, in order to better promote fulfilment of the social function. A major social function in hegemonic colonial contexts is the stratification of society into categories that appear *natural* (e.g. into *human/nonhuman, man/woman*, etc.). This social function requires that the clusters of gender norms be *mutually exclusive* and *coherent*--that is, that masculine-coded traits occur together with *all and only* masculine traits, and feminine with feminine (while both clusters have *whiteness* at their core). In these contexts, non-coherence is a threat to the "naturalistic" normative justification for the entire social structure, and is brutally punished.

A traits-based view therefore makes no claims about the nature or ontological status of gender categories or their members. It does, however, leave open the possibility that a cluster might *gain* ontological significance. In individuals, for example, a cluster might become what Witt (2011) calls a *gender role* or a *gendered social position,* what Ásta (2018) calls a *gender property,* or what Haslanger (2012) and Dembroff (2018) refer to as *gender classifications*. However, a traits-based view need not be committed to any of these. It need not be committed even to the claim that the cluster has *any* particular ontological upshot, over and above being a contingent cluster of individual norms. At its root, a traits-based view is committed only to the claim that particular traits become gender-coded through social practices--and, of course, that normative gender phenomena, such as *responsiveness* and *evaluability,* proceed from this initial step in various ways. More on this in 3.4.

This discussion might seem to raise the following question: What makes these *gender* norms? That is, if the relevant gender-coding happens at the level of the trait, and the clustering is determined by a variety of disparate social functions, what makes it the case that we are

talking about *gender norms* and not some other kind of social norms, such as racial norms or norms of etiquette?

There are two ways of interpreting this question. We might interpret it as asking what the *ontological distinctions* are between gender norms and other kinds of norms or forces—that is, what *makes something* a gender norm? Alternatively, we might read it as asking *how we can meaningfully identify* gender norms as distinct from other kinds of norms, such that we can think or talk about them as gender norms and understand their effects. I want to sidestep the first version of the question altogether, as I don't think it's helpful. However, I don't think that stops us from answering the second version.

On the first reading of the question, it assumes that gender norms *are* ontologically separable from other kinds of social norms or forces. This approach gestures at the thought that gender metaphysics can be "joint-carving"--that the ontology of the social world is neatly separable, and particular social forces and entities are clearly distinguishable from other forces and entities. However, in the case of complex social phenomena, this approach is not always helpful. Many feminists have argued that the oppressive forces which shape the social world, such as gender, race, heteronormativity, capitalism, and so on, are both causally and constitutively interdependent (see, for example, hooks 1982, Spelman 1982, Haslanger 2020). That is, the causes of these forces are not cleanly distinguishable, nor are the features which make them what they are. Moreover, a long and storied tradition of *intersectionality* in feminism holds that trying to understand these forces separately often erases the experiences of those whose oppression is compounded by this interdependence (Spelman 1982, Crenshaw 1989). That is, even if such distinctions were possible, it is not always theoretically or politically helpful to make them. We cannot, for example, understand dominant hegemonic gender norms without also

understanding their racialization, their heteronormativity, their coloniality, and their imbrication

with, e.g., norms of "rationality" that enforce epistemic violence (Longino 1989, Lugones 2007).

So, if the question is asking how gender norms are ontologically distinguished from other kinds

of norms, one way of answering this is to say that they aren't. Many or most gender norms *are*

also racial norms, sexual norms, not to mention epistemic norms and norms of etiquette, etc.--

under different descriptions. Talking about them under *this* description is a useful point of entry

because of the way in which notions of masculinity and femininity have been *used* in the world.

But perhaps their ontological distinctiveness (not to mention their ontological *unity*) is merely

apparent.

I say "perhaps" here because I do not want to rule out the possibility that there *exists* an

explanation of the distinctiveness and unity of gender norms that is consistent with the

commitments I have articulated here. (For example, we might think that all of the various

socially significant traits around which other peripheral traits cluster are unified by a relationship

to the overarching system of gender in an important way--whatever that system may be.)

However, I don't think I need to give one here, because I don't think we need it to answer the

second formulation of the question: how can we meaningfully identify and theorize about gender

norms? I think we can answer this question even if we cannot answer--or we flat out reject--the

first. Failing or refusing to give a fleshed-out ontology of what makes a gender norm a *gender*

norm does not commit us to the claim that we cannot identify those normative forces which have

traveled under the banner of gender, and the way in which those forces and their classification as

gendered have perpetuated distinctive oppressions and harms. Following Mikkola (2016) and

Antony (2020), I suggest that we do not need to give a substantive account of what gender norms

are order to be able to meaningfully understand or talk about their effects in the world. Even

ordinary speakers generally do not have much disagreement about what they mean when they

say that something is a *gender norm,* or that it is related to a standard of *masculinity* or

femininity. In cases where there *is* disagreement, we might defer to those who have a distinctive

epistemic perspective on the matter. As Mikkola writes,

> ...if we wish to know about the ordinary usage of snow-related terminology, we should
> not look at ordinary language use of speakers who have no conception of snow. Rather, it
> would seem more appropriate to consult Finns, Greenlanders, Alaskans, and other
> northern peoples. This echoes Talia Mae Bettcher's (2013) view that in order to analyze
> gender concepts, we should not take the dominant cis conceptions as our starting point,
> but rather privilege resistant trans* conceptions (2016, 121).

To be clear, I am not trying to defend a kind of nominalism about gender norms on which gender

norms are distinguished and unified simply by the fact that they are all called "gender norms."

On a view like this, what *makes* something a gender norm is the fact that it is so named. But this

isn't my position, because I'd rather remain entirely neutral about what (if anything) makes

something a gender norm. Instead, I claim that we can direct our attention to the ways in which

we *use* gender norms. Throughout this project I have spoken of gender norms as *the evaluative*

standards associated with gender, and have talked in terms of the way entities are *responsive to*

and *evaluable under* those norms--often in ways that work to perpetuate or justify oppression and

harm. However, I have also resisted, and will continue to resist, addressing the question of *what*

gender is. The broadest version of the point, then, is that we do not need to say exactly what

gender norms *are* in order to theorize about what gender norms *do*. We can do the latter simply

by identifying those normative forces that have traveled under the banner of gender, and tracing

the accordant effects.[15]

---

[15] It may be important that individuals *understand* certain forces as being gendered; that might, for
example, explain how they become responsive to them. But, again, I think we can explain this in terms of
identification and use of gender norms--something we don't need a clear ontology to be able to do.

To summarize: I have sketched a *traits-based view,* on which gender norms apply on the basis of observed or imagined traits. Traits are descriptive features of individuals, groups, or other social entities. Gender norms tend to (but need not) "cluster" around particular social functions, like *reproduction* or *food preparation*. These social functions are themselves collections of traits that have some significance in the context; those traits constitute the "core" of the cluster. The clusters may (but need not) gain further ontological significance, as, for example, a gender role or a gender category. However, the connection between trait and norm is the "basic" level of gender norm assignment. This leaves it open that there is no ontological significance to the "gendering" of gender norms. Even if that's true, however, we can still meaningfully speak and theorize about them and their effects.

In the next section, I will discuss, in turn, how a traits-based view meets each of the desiderata I identified in 3.2. In doing this, I will further explore the view.

3.4. HOW THE VIEW MEETS THE DESIDERATA

In this section I will explain, in turn, how a traits-based view meets the desiderata outlined in 3.1. Recall the six desiderata: (1) *responsiveness,* (2) *evaluability,* (3) *social construction,* (4) *nonconsensual assignment,* (5) *differential assignment and embodiment,* and (6) *oppression.* Importantly, in meeting each of the desiderata, the view must also respect all of the the other desiderata. The view's account of *responsiveness,* for example, cannot conflict with *social construction* (as in Serano's *intrinsic inclinations* view). In discussing each desiderata and their interactions, I will further explore the commitments of a traits-based view.

3.4.1. *Responsiveness.* The account must be able to explain how marginalized people are *responsive to* gender norms.

The major challenge I raised for explaining *responsiveness* was the phenomenon of *cross-category norm responsiveness,* where many trans and GNC people experience norms that ostensibly have not been assigned to them as nevertheless exerting force over their behavior. A good explanation of responsiveness must also be consistent with other desiderata, in particular *social construction* and *nonconsensual assignment.* For example, a view like Serano's (2007) suggests that that cross-category norm responsiveness is fundamentally *incompatible* with these desiderata. Versions of this position have been suggested from a variety of perspectives, including other trans theorists such as Prosser (1998), but also including many so-called "gender-critical" feminists, who take social construction as paramount and conclude that cross-category norm responsiveness is either unimportant or spurious (see, for example, Bindel 2014, Lawford-Smith forthcoming). I suggest that it is a major advantage of a traits-based view that it can explain cross-category norm responsiveness in terms that are consistent with social construction and nonconsensual assignment, thus opening up space for inclusive theorizing.

Recall from 3.2 the discussion of *internalization/habituation.* Drawing on prominent perspectives in the metaphysics of gender (Witt 2011, Haslanger 2012) as well as insights from social science (Scott 1971; Sripada & Stich 2006), I argued that we should hold that agents become responsive to gender norms because they *internalize* those norms as a response to sanctions, and become *habituated* to them. That is, what explains social norm responsiveness of all kinds, including gender norm responsiveness, is the following pattern: Agents are trained that certain behaviors consistently lead to reward and others consistently lead to punishment. Over time, they *internalize* the pattern, and consistent reward and punishment are no longer required to

produce the behavior (Scott 1971). This involves *habituation,* or the psychological state where individuals calibrate their behavior according to the norms regularly and without thinking; they become *responsive to* those norms (Witt 2011). What do we get when we combine this with a traits-based view?

I argue that there are two ways in which agents can internalize gender norms and become responsive to them. First, they can be internalized because of *coherence pressure.* Second, they can be internalized through *vicarious learning.* I will discuss these in turn.

On a traits-based view, children will learn from their surroundings that certain *traits* are coded masculine or feminine. Put differently, sociocultural training instils a connection between *trait* and *norm,* and individuals with the relevant trait learn that they ought to follow the norm. Moreover, those raised in cultures where *cluster coherence* is mandatory will also learn that displaying traits which "mismatch" is punishable. However, it is abundantly clear that individuals are rarely disposed to express a perfectly coherent gender-coded set of traits. If a person expresses or is disposed to express a trait which is inconsistent with other traits that they are observed or imagined to have, and, importantly, if they have internalized a cultural presupposition that clusters of gender-coded traits are naturally coherent and mutually exclusive, they may feel that they *ought* to follow the norms associated with the traits they are disposed to express, in order to maintain coherence. For example, consider someone in dominant colonial culture who is perceived by others to have primarily masculine-coded physical traits, but is nevertheless disposed to display feminine-coded traits such as shaving their body hair, wearing a dress, or identifying themselves as a woman. For this person, the pressure to express a coherent set of traits may cause them to feel pressure to express other feminine-coded traits. Call this *internalization from coherence pressure.*

Consider an example. Recall Serano's report: "I had an unexplainable feeling that I was doing something wrong every time I walked into the boys' restroom at school" (Serano 2007, 78). Here is one traits-based explanation of this experience. Serano either had, or was disposed to have, some traits that are coded feminine; these might be behavioral traits, traits of appearance, physical traits, and so on. (Again, I am not trying to give an empirical explanation for why individuals have or are disposed to have the traits that they do. Serano, a biologist, gives an empirical explanation of this; see her 2007, 80-82. Let's assume that this or some other explanation is correct.) Since she had learned that gender-coded traits are only supposed to occur in coherent clusters, and, crucially, that *those who display non-coherence are punished,* she experienced *coherence pressure* to display other feminine-coded traits--such as using the girls' bathroom. That is, she learned that *those with feminine-coded traits go in this bathroom;* and insofar as she experienced herself as either having or being disposed to have some feminine-coded traits, she found herself responsive to that norm.

The second way in which agents can become responsive to gender norms is through *vicarious learning.* Drawing from *social learning theory,* Scott (1971) notes that norms can be internalized directly, as a response to sanctions applied to the individual themself; however, they can also be internalized *vicariously,* as a response to sanctions applied to someone else (55). That is, if an individual observes another person as being rewarded or punished for a behavior, "he learns that what happens to others can happen to him too" (55), and can internalize that norm himself. Importantly, for Scott, only observations of individuals who are relevantly similar to the individual will lead to vicarious learning. For example, "as sex roles become differentiated, [a boy] will learn there are some acts for which his sister will be rewarded but for which he will be punished, and *vice versa*" (56). If gender norms attach to individuals on the basis of their gender

124

categories, as Scott suggests, then vicarious learning cannot explain how individuals acquire cross-category norm responsiveness, for children who believe that they are boys will learn vicariously only from sanctions applied to other boys. However, if gender norms attach on the basis of *traits*, the relevant similarity will be trait, not category. On this picture, individuals identify others in their environment who exemplify traits that they either have or are disposed to have. They then learn vicariously from the rewards and punishments applied to those others, independently of their gender category.

> Consider an example. Recall Mock's report:

> I had a vision board of my ideal, pulled mostly from the pop-culture images that MTV fed me. I wanted Halle Berry's or Tyra Banks's breasts, Britney Spears's midsection, Beyoncé's curvy silhouette and long hair, and I prayed that I wouldn't grow any taller so I didn't tower over the petite Asian girls who were the barometer of beauty in the [Hawaiian] islands. (2014, 122-123)

I argued that we cannot understand this experience in category-based terms. As a trans child and teenager, Mock was *not* consistently rewarded for following feminine norms or punished for failing to follow them. Rather the opposite; throughout this memoir, she recounts the ways in which she was regularly punished for displaying feminine-coded traits, by her parents, siblings, and classmates. However, if we understand the norm-responsiveness she displays in the passage as based on *vicarious traits-based learning,* it becomes clear. Mock observed others in her environment as relevantly similar to her, and observed them being rewarded for displaying particular feminine-coded traits and punished for failing to display them. She then generalized to her own case; she learned, vicariously, to become responsive to those norms, because she too was disposed to express those traits.

At this point, it might be objected that it is unclear why an agent would become responsive to certain norms and not others. For example, someone like Serano or Mock may

have been disposed to express a variety of feminine-coded traits, but they would have been expressing a variety of *masculine*-coded traits as well--in particular the physical traits which led to their classification by others as boys. It might seem unclear, then, why they would learn vicariously but not directly, or internalize normative pressure to cohere with a feminine cluster rather than a masculine one. That is, why would a trans girl become responsive to feminine norms rather than masculine ones?

I first want to reject the thought that agents are responsive to some norms "instead of" others, as if they are *actually* mutually exclusive. Nothing about what I have said so far indicates that being responsive to different gender norms is impossible. My view suggests the opposite-- which is an attractive feature, because reality also suggests the opposite. It is fully consistent with the accounts listed here, and more broadly with the experiences of many trans and GNC people, that one can be responsive to a variety of apparently incompatible norms at once. For example, Mock describes her childhood self as being at times drawn towards masculine norms of dress and behavior, specifically because she did not want to disappoint her mother (2014, 98). Many trans and GNC people struggle to interpret our identities and figure out how we want to behave precisely because we are pulled by multiple, apparently incompatible, sets of norms at once. As children, we are trained by straightforward reward and punishment that we should follow some norms, even as we observe that those who are like us in the ways that seem most important follow other norms. We learn that non-coherence is unnatural or impossible, even as we experience unprompted non-coherence in ourselves. It is not surprising, then, that we will experience various norms as exerting force over us.

How, then, do we determine which norms will *really* guide our actions? If a trans or gender-nonconforming person is responsive to masculinity and femininity all at once, what's the

difference? Why do people often respond as if the cross-category norm responsiveness is more pressing? I will answer this question at length in Chapter 4. For now, let me just say that certain traits can feel closer to the core of *who one really is*; that is, expressing them can feel more *authentic* than expressing others. Moreover, given how strongly non-coherence is punished, perhaps many (apparently) cis or gender-conforming people experience a weaker form of cross-category norm responsiveness that does not end up guiding their actions, and is ignored in favor of the more socially acceptable responsiveness. That is, perhaps in a less strongly cis-normative society, displays of non-coherence would be much more common. This is, of course, an empirical question that I cannot address in full here. The basic point is this: I do not want to claim that being responsive to one set of norms is incompatible with being responsive to another set. I only want to say that being disposed to express traits which are gender-coded in a particular way can often lead to responsiveness to the associated norms, and that this can explain cross-category norm responsiveness.

3.4.2. *Evaluability.* The account must be able to explain how marginalized people are *evaluable under* gender norms.

In Chapter 2, I raised the *evaluability objection* against the category-based view, which is twofold. First, it cannot explain how those who are excluded from certain gender categories (such as colonized people who are not counted as "men" or "women" on dominant colonial frameworks, Black enslaved people who are forcibly "ungendered," or what Riki Wilchins called "gendertrash") are nevertheless subject to the associated gender norms, often in particularly brutal ways. Second, it cannot explain how gender norms are enforced at what I called the

"macro-level"--against groups, societies, and epistemic frameworks. A traits-based view, however, has no trouble explaining this. Here is how.

First, on a traits-based view, one need not occupy a particular gender category in order to be subject to the associated norms. It is sufficient for one to have, or rather be observed or imagined to have, the associated *traits.* In particular, being observed or imagined to have physical traits that are associated with sexed embodiment is often a key part of being responsive to and evaluable under particular sets of gender norms. (In 3.2 I argued that a view of gender norms which does *not* say this would be unintuitive.) Those who are excluded from dominant gender categories in various ways are often nevertheless still interpreted as having physical sex traits. It is no surprise, then, that they are held to the associated norms. Recall Lugones' discussion of colonized males and females who were historically constructed as "not-human-as-not-men" and "not-human-as-not-women" (2010, 744), but nevertheless subject to brutal gendered violence, in part for their *failure* to conform to the standards associated with the white-centric gender categories "man" and "woman."

Moreover, as discussed, dominant colonial gender categories notoriously center the traits associated with certain kinds of embodied privilege. This is likely because the *social functions* around which these norms cluster are intimately connected with the perpetuation of a falsely naturalized social order which privileges certain kinds of bodies, behaviors, and social organizations. The primary function of dominant colonial gender categories is to privilege white, non-disabled, cisgender, heterosexual maleness while also making that privilege, and these classifications themselves, appear natural or normal. The "core" of the associated clusters, then, will include traits of *whiteness* (as well as heterosexuality, non-disability, and so forth). Those who are interpreted as having some of the core traits of these clusters, such as sexed bodies, but

128

lacking other core traits, such as whiteness, are therefore exhibiting *non-coherence.* Moreover, since whiteness and sexed bodies are both understood to be relatively *stable,* this non-coherence is interpreted as permanent or fundamental. On a system which rests its legitimacy on the existence of naturalized, mutually exclusive, coherent gender clusters, permanent non-coherence is a grave normative error. However, since these norms exist precisely to encode the dominance of certain individuals and groups over others into naturalized law, we should not be surprised that these norms characterize subordinated individuals and groups as *permanently failing* according to this system, thus justifying systematic, brutal punishment.

Second, a traits-based view is well-positioned to be able to handle the evaluability of macro-level social entities. I have already discussed the ways in which groups and cultures can be interpreted as having traits. Importantly, recall here that it does not matter for the view whether these entities do in fact have these traits--or even *whether such entities can have traits at all.* What matters is whether the social entities are *interpreted by the relevant others* as having the traits at hand.

The phenomenon of macro-level social entities being evaluable under gender norms according to their perceived traits is well-documented. For example, Robin Zheng (2016) notes the ways in which entire races are gendered; for example, Asians as a racialized group are coded feminine "due to their purportedly shy, soft-spoken, submissive racial 'essence'," while Blacks as a racialized group are coded masculine "due to their purportedly aggressive 'essence'" (405-406). Similarly, social psychologist Geert Hofstede (2001) argued that different societies, ethnicities, and countries are gender-coded according to their perceived traits. According to Hofstede, societies which *prioritize achievement* and *organize structures around money and things* are "masculine," while societies which *prioritize nurture* and *organize structures around*

relationships are "feminine" (2001, 297-8). Again, it does not matter for my purposes whether these societies in fact have these traits--or, for that matter, whether Hofstede's taxonomy of gender-coding is accurate. What is important is that we see the following pattern correctly predicted by a traits-based view: Macro-level social entities, such as groups and societies, can be interpreted as having traits, which are then coded as masculine or feminine. The entities are thereby evaluable under the relevant gender norms.

3.4.3. *Social construction.* The account must be consistent with the basic metaphysical position that gender is *socially constructed*.

It should be obvious from the discussion so far that a traits-based view meets this desideratum. It does not suggest that gender is a part of any intrinsic, natural, or "pre-social" features of individuals, groups, or humanity as a whole. It does, of course, suggest that we often *perceive* these entities as having such features. For example, we may interpret entire races as having "essential" features that are coded masculine or feminine (see again Zheng 2016). However, such interpretations need not be posited as reliable trackers of real gendered features. Rather, we are conditioned by a deeply gender-coded world to see such features everywhere. It is beyond the scope of this view whether there is such a thing as an intrinsic, natural, or pre-social trait. Again, I am not in the business of explaining why entities have the traits that they do. I am arguing that certain traits--wherever they come from—are gender-coded by our social world.

In 3.4.1, I argued that it is a major advantage of a traits-based view that it can answer the argument that *cross-category norm responsiveness* and *social construction* are incompatible, because cross-category norm responsiveness suggests that certain responses to gender norms are not socially learned but are rather innate or pre-social (see Serano 2007). I pointed out that a traits-based view avoids this problem, by showing how those responses *can* be learned--via

coherence pressure and *vicarious learning.* A traits-based view is therefore not only compatible

with social construction, but uniquely well-positioned to balance it with other desiderata.

3.4.4. *Nonconsensual assignment:* The account must capture the way that gender norms can exert
power over us without our participation or consent.

A traits-based view is clearly compatible with nonconsensual assignment. Insofar as we

do not *choose* which traits we express or are disposed to express--and have even less control over

the traits that we are *perceived* to have--we do not choose which norms apply to us and exert

power over us. Moreover, the gender-coding of our particular traits is also not a matter of

individual preference and choice, but is determined by one's community and history.

Notice, however, that a traits-based view *can* allow that we have some participation in

this process, insofar as we do choose whether or how we express certain traits. For example,

certain traits of appearance, such as *wearing makeup* or having *a full beard,* are partially "up to

us;" insofar as we have access to the tools and capacities to express or not express these traits,

we can choose how to present ourselves for evaluation. Other traits are even more under our

control, such as the expression of our gendered self-identifications. In fact, there exist trans-

inclusive domains of gender norms where *self-identification* is the *core* trait for particular gender

clusters. (This is still consistent with *ascriptivism,* insofar as the norms still apply because of the

practices of others *with respect to* one's self-identification.) In these domains, the assignment of

gender norms is--or is at least working to be--*consensual.* The view therefore leaves open the

possibility that not all gender norm assignments are entirely nonconsensual, even if they are all

ascriptivist. Relatedly, subaltern communities may redefine gender-coding of particular traits, so

that, e.g. a full beard may no longer be coded masculine, but can be interpreted as neutral or even

feminine (Vaid-Menon 2020, 111). I will say more about subaltern communities and gender norms as resistance in 3.4.6.

3.4.5. *Differential assignment and embodiment*: the account must capture the intuitively obvious fact that gender norms apply differently to different people, often on the basis of their observed or imagined sexed body.

A traits-based view is fully consistent with this desideratum. Some of the most commonly gendered traits are *physical* or *embodied* traits, particularly those associated with the *sexed body*: genital structure, reproductive capacity, secondary sex characteristics, body shape and size, chromosomes, and so on. These traits are often--though not always--at the *core* of dominant clusters, sometimes operating under the aggregate label *biological sex* (as in the discussion of Dembroff in 3.2). On a traits-based view, then, gender norms will very often apply differently to different people on the basis of the sexed/embodied traits that they are observed or imagined to have. Moreover, these will often be the most oppressive and inflexible gender norms, as physical embodied traits are perceived to be *stable* or *immutable.*

A traits-based view has further advantages here. In particular, it leaves much room for flexibility and nuance in explaining differential assignment that does *not* involve embodiment. For one thing, it can explain the way that gender norms are assigned differently on the basis of other features that have nothing to do with sexed embodiment, such as when different people observed or imagined to have female-coded bodies are held to different standards on the basis of their race, class, sexuality, and so on. Similarly, it can explain the way that one's sexed embodiment can be interpreted differently in different contexts. This is so because traits, by their very nature, are often fluid, malleable, and interpreted differently in different contexts. I will say more about this in 3.5.1.

3.4.6. O*ppression:* the account must enable us to explain why many domains of gender norms, particularly the hegemonic norms that are operative in dominant contexts, serve to justify and enable oppression.

Throughout this project, I have discussed two prominent ways in which dominant gender norms do this. First, they work to *naturalize* gender. As agents become habituated to following gender norms, they come to do so automatically and without thinking. This makes the gender-normative behavior seem *natural* or *intrinsic,* and thus makes gender itself look natural or intrinsic. Thus, gendered behavior appears as a basic feature of human nature, one that we cannot, and perhaps should not, try to change. The behavior, and any oppressive effects it may have, will therefore be justifiable as "natural" and thus treated as immutable and even desirable.[16] Second, and ironically, the apparent "naturalness" of normative gender coding and norm clusters works to justify punishment of those who do not conform to them, on the grounds that non-conformity is *unnatural*. This is also linked with desiderata 4 and 5. Given that the physical traits of the body are believed to be fixed, pre-determined, and paradigmatically natural (as in "natural woman"), non-coherence is often most viciously punishable when the clusters in question are nonconsensually assigned on the basis of one's body. Moreover, I have argued that oppression on the basis of gender norms is visited against macro-level social entities, such as groups and societies.

A traits-based view is consistent with *habituation.* Children learn that certain traits are gender-coded, typically from a very early age. As they internalize the connection between trait and norm, they learn that those who exhibit the trait should follow the associated norm. Over

---

[16] Recall that, in Chapter 1, I explained that feminist social constructionism is largely motivated by the aim of "debunking" this justification, in order to resist the claim that the oppressive effects of gender are immutable or desirable.

time, they become habituated to navigating a world where those norms apply. Moreover, on many dominant gender norms, mutually exclusive and coherent gender clusters are enforced, such that agents are likely to internalize *coherence pressure* and become habituated to performing coherent clusters.

Moreover, a traits-based view is well suited to explain the way that macro-level social entities are oppressed according to their perceived gender-coded traits. The prioritization of white masculinity and all associated features (e.g. rationality, individualism, dominance) on hegemonic systems of norms suggests that the interpretation of any features determined to be non-white or non-masculine (inclusive) are appropriate for subordination. Thus, we can understand the justification of colonial dominance in terms of the traits that non-colonial groups and societies are interpreted to have. On a colonial gender system, there is a *right* and *wrong* way of organizing the social world, and this is justified in part by associations with the broad-ranging, traits-based *gender norms.*

One attractive further upshot of a traits-based view here is that it does not entail that *all* gender norms are oppressive. Clustering need not be naturalized or enforced as brutal. (In fact, gender norms need not cluster at all.) Therefore, it is in principle possible to have domains of gender norms that do not enable and justify enforcement of norms that are nonconsensually assigned on the basis of the sexed body. For example, the core traits of a cluster can allow for consensual assignment, by including traits of *self-identification* or *self-ascription.* Moreover, the clusters themselves can be modified to fit the needs of the community. For example, if some society needs to make trans experience intelligible*,* the gender norms in play can cluster according to the needs of those in that community.

Luckily for anyone interested in such a project, it is already underway. Bettcher (2013) points out that within queer and trans subcultures, gender and sex categories are fixed by self-identification, and their features are defined by those who thereby identify. For example, Bailey (2011) describes the normative gender practices at work in *ballroom culture,* queer and trans communities of color which exist in most major cities in the U.S. According to Bailey, identities and norms within these cultures are created out of twin needs: to make selves, bodies, and gendered experiences intelligible to one another, and to navigate the possibility of punishment and violence visited on the visibly queer in the broader world. The clusters that emerge are therefore, specifically and by design, tools of *resistance* against oppressive gender norms. (I say more about ballroom culture and resistant norms in Chapter 4.)

Some versions of the category-based view seem unfriendly to this possibility. For example, on Haslanger's (2012) view (understood very roughly), gender categories are classifications that create an oppressive hierarchy on the basis of observed or imagined sexed bodies. Gender norms exist exclusively to perpetuate and naturalize these categories. This view suggests that gender norms are irredeemable. If gender norms depend metaphysically not just on oppression, but specifically on nonconsensual assignment on the basis of the sexed body, the possibility of *resistant* uses of gender norms seems strange. In fact, Haslanger cautions against "theoretically appropriating" masculinity and femininity, on just these grounds (47). I worry, however, that a view which builds oppressiveness into the very nature of gender norms has problematic implications. On such a view, at best, all gender norms will force diverse individuals into unnatural boxes for no good reason; at worst, they are fundamentally oppressive and therefore cannot be used to "dismantle the master's house," as it were.

A traits-based view, however, does not have these implications. Although I have avoided making claims about the metaphysical nature or origin of gender norms (and will continue to do so), a traits-based view is clearly very conducive to the possibility of resistant *uses* of gender norms. That is, not only can it explain how many domains of gender norms are oppressive, it can also allow that some domains are constructed not to be oppressive, and indeed specifically work to resist oppression. I take this to be an attractive feature of the view.

I conclude that, not only can the traits-based view can meet all of the desiderata I have identified, it also has further benefits for theorizing resistance against dominant gender norms. In the next section, I will respond to objections against the view.

3.5. OBJECTIONS

In this section I consider three major objections against a traits-based view. Each objection will suggest that a category-based view has the advantage in explaining particular features of gender norms. I show how a traits-based view can answer each objection.

3.5.1. First, I consider the *objection from multiplied categories.* It might be argued that a category-based view can explain all of the phenomena I am describing, simply by positing a *wide multiplicity* of gender categories that are deeply context-dependent and vary according to other features, such as race, sexuality, disability, various elements of context, and so on. That is, rather than holding that the gender categories in play are just *man* and *woman,* perhaps they are more fine-grained. For example, in the case of dominant colonial gender norms, perhaps it's not that Black and Brown colonized people were excluded from gender categorization altogether; perhaps it's that they were simply assigned to specifically subordinated gender categories within

the system. That is, perhaps the categories at play were not just *man* and *woman,* but *man (white), woman (white), male (colonized, non-white), female (colonized, non-white),* and so on-- with increasing levels of granularity depending on the particular colonizing force and the existing systems in the non-colonial society. In fact, at times Lugones seems to be suggesting something like this in her earlier work (2007, 198-200). Similarly, perhaps Wilchins' *gendertrash* is its own gender category, and the norms under which Wilchins and others are evaluated are attached to this category.

Moreover, the objection might go, this account can explain certain features of gender norms *better* than a traits-based view can do. Ásta (2018) sometimes appears to defend a view which suggests this:

> Consider this scenario: you work as a coder in San Francisco. You go into your office where you are one of the guys. After work, you tag along with some friends at work to a bar. It is a very heteronormative space, and you are neither a guy nor a gal. You are an other. You walk up the street to another bar where you are a butch and expected to buy drinks for the femmes. Then you head home to your grandmother's eightieth birthday party, where you help out in the kitchen with the other women while the men smoke cigars. (2018, 73)

According to Ásta, this case demonstrates that the *gender properties* that are conferred on an individual by others--e.g. the property of *being a guy* or a *gal* or a *butch* or a *femme* or an *other*-- vary from context to context. Presumably, since this story is all happening in a single evening, our protagonist displays similar traits in each of the contexts described here. Nevertheless, the norms that apply to them seem to vary widely, even in different bars on the same street. Similarly, we can assume that they are perceived to have the same sexed embodiment as many of the others in this story. In the non-heteronormative bar, for example, the butches and femmes are likely read as having similar sexed features. Again, however, the norms that apply vary significantly. What could explain this? A defender of the category-based view might conclude

the following: Gender norms are assigned on the basis of fine-grained and context-dependent gender categories, such that one's gender category can change between different bars on the same street. At root, however, what explains the assignment of gender norms is the *category* into which one is sorted, rather than the traits that one is perceived to have.

In responding, I first will note that I do not claim that there are *not* a multiplicity of fine-grained, context-dependent categories. In general, I want to have as few commitments about the nature of gender categories as possible. I am only committed to the claim that gender categories do not explain gender norm assignment. So, it might be correct that one can change gender categories going from one bar to another, and that there are multiple gender categories in play relative to one's other intersecting features and identities. But I do not think these categories can explain the assignment of gender norms, for a few reasons.

First, I want to resist the specific claim that a category-based view can explain the enforcement of gender norms on people ejected from the relevant categories. For example, consider Lugones' argument about those who are not included in the colonial categories "man" and "woman". Even if we hold that colonized people belonged to racialized subordinated gender categories, Lugones is still very clear that the *norms* assigned to colonized people were *grounded in the standards assigned to dominant white colonizers*. She writes:

> Judging the colonized for their deficiencies from the point of view of the civilizing mission justified enormous cruelty. I propose to interpret the colonized, non-human males from the civilizing perspective as judged from the normative understanding of ''man,'' the human being par excellence. Females were judged from the normative understanding of ''women,'' the human inversion of men. (2010, 744)

According to Lugones, colonized people are judged according to the normative standards associated with the dominant categories *man* and *woman*. That is, even if there are different gender categories for everyone within a colonial gender system, the norms themselves are

applied from the "top" down. The problem remains: being ejected or barred from a dominant gender category doesn't entail not being evaluable under the associated norms. "Colonized females" are judged specifically for being *not women,* where "woman" is a status only available to white colonizers, and the purported features of women are the basis for the associated normative standard. The norm assignment occurs on the basis of the not-women's *lack* of gender categorization as women, rather than the presence of an alternative categorization.

Relatedly, the move to a multiplicity of categories cannot explain the enforcement of gender norms on those who are *fully* ejected from dominant categories--such as Wilchins' "gendertrash," or those who are, in Spillers' terms, "ungendered." Part of the specific harm that is being articulated in these cases is the harm that comes from being *removed* from gender classification. Spillers is very clear that the brutal violence visited on enslaved "flesh" is "not at all gender-specific" (1987, 67). Moreover, Wilchins' locution of "gendertrash" strongly suggests something similar; their social punishment is a result of being *thrown out* or *ejected* from gender altogether. Lopes (2019) argues that gender intelligibility is necessary for what she calls *full social standing*, or the status of being able to stand in certain kinds of social relationships--even the most basic ones--to other people. Being *ungendered* or *gendertrash* is particularly brutal because it leaves one open to evaluation and punishment for failing to do gender correctly, even as it makes that impossible by constructing them as outside of gendered intelligibility. If the objection claims that these systems *do* in fact have gender classifications for "gendertrash" or "the ungendered," it misses the point.

In short, the category-based view trades on the thought that what is harmful or oppressive about gender norms is the way in which they enforce gender norms on those who occupy a subordinate category. But cases of *non-inclusion* or *ungendering* are not explained by positing

that there are more subordinate categories than we originally thought. Rather, the thought is that

gender systems can oppress, not just by creating subordinate categories, but by *removing some*

people from categorization altogether, and then punishing them more harshly according to that

removal. Thus, it is partly the *exclusion* from categorization which enables gender norm

enforcement--and that cannot by explained by any number of gender categories.

Moreover, I think that the multiple-categories view is misguided, because it abandons the

key important attractive feature about a category-based view: its ontological simplicity. A

category-based view provides a parsimonious explanation for how gender norms are assigned to

individuals, one that easily explains certain commonalities across contexts (such as the

subordination of those perceived to have feminine-coded bodies). In order to accommodate the

objections I have raised, however, the multiplied-categories view will have to be maximally

flexible. It will have to allow that gender categories can vary widely in their natures, membership

conditions, and associated constellations of norms, based on a variety of intersecting features, as

well as context--that is, a huge constellation of disparate factors which share little or nothing in

common. But once a category-based view has conceded this, it loses that ontological

simplicity. It's not clear what advantage there is here to positing a *gender category.* On this

view, we are left with something like the following: gender norms apply to people on the basis of

their gender category, which is conferred on the basis of some features they are interpreted as

expressing in the context. But, again, this is just the traits-based view with an extra step. Again,

why not cut out the middleman? What is the gender category really explaining?

It might still be objected that a multiple-categories view has better resources to explain a

case like Ásta's, where the gender norms which apply to an individual change from context to

context. But I argue that a traits-based view has these resources as well. Traits, in general, are not

interpreted in isolation. Rather, they are interpreted both in the context of the other traits one is

understood to have, and in the specific context in which they are expressed. Lugones (1987)

points to this phenomenon in her discussion of *playfulness.* While Lugones' close friends insisted

that she was playful, her colleagues in academia maintained that she was not at all playful, and

was in fact very serious. She writes that "I was sure I had the attribute in question and, on the

other hand, I was sure that I did not have it. I remain convinced that I both have and do not have

this attribute" (9). On Lugones' account, this is not a contradiction, but rather a function of the

way in which she was differently interpreted in different social worlds. Her personality did not

change, but others' readings of her did—based not only on the context, but also on how her *other*

traits were perceived in that context. Lugones further writes of being understood as "intense" by

many white people, because they also interpret her as being Latin-American, and have

stereotypes about associations between being Latin-American and being intense (13). Certain

behaviors might not count as "intense" when expressed by a white person, but are counted that

way when expressed by Lugones, because of other traits that she has and the way the dominant

social world interprets them together. She therefore understood herself as both *playful* (in her

"home worlds") and *not playful* (in dominant worlds).

Something similar can be true of gendered traits. Just as the same behavior can be

counted as "playful" or "not playful" depending on the world one is in and the other traits one is

perceived to have, certain behaviors or features can be counted as differently-coded traits, partly

in virtue of the world one is in and the other traits that one is perceived to have. It is therefore not

inconsistent with a traits-based view that Ásta's protagonist would be subject to different norms

in different contexts. For example, at work they might be perceived to have a trait of *intelligence,*

while in their traditionally gendered family home, those same behaviors might be perceived very differently.

Finally, notice that a multiplied-categories view still leaves out the evaluability of macro-level social entities. In order for a category-based view to explain this, it would have to say that macro-level social entities, such as groups and societies, can belong to gender categories. But that seems deeply implausible. To claim this would be to stretch the meaning (and certainly the usefulness) of a gender category beyond recognition. At its root, a category-based view is positioned to explain only how gender norms apply to individuals, because individuals are the kinds of objects that belong to gender categories. It might be objected, then, that a category-based view is *only* supposed to apply to individuals, and therefore it doesn't need to explain how gender norms are assigned to macro-level social entities. If that's right, then a traits-based view has a clear advantage over a category-based view here. It can do both.

I conclude that a multiplied-categories view fails to explain how gender norms work, and that a traits-based view still has the theoretical advantage.

3.5.2 Second, I will consider the *causal story* objection. According to this objection, a category-based view has the following explanatory advantage: It can give us a *causal story* about why particular features of the world are gender-coded and others aren't. This causal story generally begins with categories that are assigned on the basis of reproductive functionality. For example, Briggs & George (forthcoming) argue that gender phenomena are individuated "not by their contemporary connections to sex biology, but by their historical continuity with categories that were originally closely connected to sex biology" (2). That is, all genders and gender phenomena (including gender norms) are descended with the "right sort of causal continuity" from

"primordial" gender categories, which are constituted by, and assigned solely on the basis of, sex biology (20-21). This is similar to what Witt (2011) describes as the *engendering function* of gender categories. On a view like this, most societies have gender norms because they need to organize the social world around reproductive functions. Thus, they create *gendered social positions* (or gender categories) in order to delineate who plays what role. That is why those social positions, and the associated norms, usually are assigned on the basis of reproductive function. It also explains why gender norms typically tend to involve norms that are recognizable as related to reproduction, such as *chastity* for women. "The key point is that engendering is a social function with two primary social positions, and the associated social roles are specified in contrast to one another" (Witt 2011, 40). According to the *causal story* objection, it is intuitively central to the existence of gender norms that they have some important causal connection to reproductive roles. The category-based view can explain this and the traits-based view cannot. This is a theoretical advantage for the category-based view.

I do not think this is a real problem, for three reasons. First, as both Briggs & George and Witt acknowledge, even if there is a historical causal story related to sex biology about how gender norms came to be, the norms themselves take on a life of their own. That is, wherever the norms come from, I argue that they can become so separated from their origins that the causal story gives us little important insight. Moreover, if the sex-based biology view is right, then the origins of gender norms are so temporally removed from us that we likely have little epistemic access to what Briggs & George call "primordial gender categories." Even if these are the causal origins of gender norms, why should we think those origins tell us how they work now?

Second, there are very many domains of gender norms, and not all of them have the same causal origins. For example, if we take Lugones' arguments about the coloniality of gender

seriously, we get a causal story about many dominant domains of gender norms which merely *appears* to be about sex biology, but really is about racialization, Eurocentricity, and global control. In these cases, the "sex biology" explanation seems to be a smokescreen, meant to direct our attention away from the real underlying functions of those norms. At the very least, this shows us that gender norms may have many different causal origins. Why should we think that all gender norms have the same unbroken causal chain stretching back to primordial times? That is an empirical claim for which we do not have sufficient justification.

Third, even if it is conceded that gender norms do in fact have this causal chain and that a view of gender norm assignment should be able to explain it, I don't see why we need categories to do this. The views considered above assume that, since gender norms tend to occur together with gender categories, and gender categories seem importantly connected to reproductive function, it must be the case that the reproductive functions lead to the construction of the categories and the categories generate the norms. But I don't see why this follows. We might think, instead, that both gender norms and gender categories are *caused by some third factor.* That is, perhaps "gender" is a broad and multifaceted system of social organization that has various features, including both norms and categories. Suppose that this system of social organization does often have some important connection to reproductive biology; a "gender system" is however the society organizes itself and its values in connection with reproduction. Within this system, certain *traits,* such as the ability to seed or bear children, are very important. These traits therefore take on various kinds of normative significance within the society. They might, for example, become the core of normative clusters that are associated with various roles in child-rearing. Those clusters can take on the further the ontological significance of a gender category. Alternatively, these traits might also take on independent normative value, such as via

association with *divinity.* But none of the above suggests that we *need* to involve gender categories in the story.

Taken together, these three responses are sufficient to reject the *causal story* objection. I conclude that the objection fails.

3.5.3 The final point I want to consider is not an objection to my positive view so much as it is an objection to my wholesale rejection of the category-based view. It goes as follows: Even if we accept the claim that a category-based view cannot explain the experiences of marginalized people, we need not jettison it altogether. Perhaps a traits-based view is trying to do too much. Why should we not be be pluralists about gender norm assignment? That is, why not hold that gender norms are sometimes assigned on the basis of gender categories and sometimes on the basis of gender-coded traits? There are perhaps good reasons to think this. For example, gender categories sometimes seem critical in explaining how gender norms are enforced; people will say, for example, that *boys don't cry* or that *girls should be polite.* Why not think that a traits-based view explains some cases while a category-based view explains others, and adopt both? I will call this the *why not both* objection.[17]

In response, let me first say that, if this objection works, most of my view remains intact. At worst, I can bite this bullet. Instead of recommending *replacement* of the category-based view, I could simply adjust and say that the traits-based view *supplements* the category-based view. That is, even in this objection were to succeed, it would not be so bad for my position.

However, I don't think the objection does succeed. I concede that it sometimes appears that gender categories are important in explaining the assignment of gender norms. However,

[17] Thanks to Emily Tilton for this objection and name.

that is easily explained by a traits-based view. This sheds light on one further attractive feature of

a category-based view. My description of *traits* includes a wide variety of descriptive features,

including those which are often associated with a gender category--such as *thinking of oneself as*

a woman or *being called a boy by others.* As a result of enforced coherence, these traits will very

often occur together with other traits, such as physical features or gendered behaviors. When it is

said that *boys don't cry,* for example, we can understand this as claiming the following: those

who have traits like *being called a boy, having a masculine-coded body,* etc., should not also

have traits such as *crying.* This is non-coherence--perhaps of a particularly vicious sort, since

being called a boy and *having a penis* are at the core of many dominant clusters. Because these

count as *traits,* it will still come out as true that the things which are taken to ground gender

categories (such as sexed body, self-ascription, and so on) *do in fact* play a role in gender norm

assignment. That is, a traits-based view can not only describe many things that a category-based

view cannot, it can also describe everything that a category-based view *can*--but without the

theoretical baggage.

   If one is a nominalist about gender categories, perhaps this is enough for a version of the

pluralistic view to succeed. That is, if gender categories are what they are because they are all

picked out by the same term, then perhaps a traits-based view just encompasses a category-based

view here, because having a gender category *will* just be a trait (specifically, the trait of being

called a woman, a man, etc.). I have no issue with that outcome. What I want to resist, however,

is the claim that gender norms are assigned on the basis of an entity's ontological status as a

member of a substantive gender category, rather than on the basis of perceived traits--that

is, *bare descriptive facts* about that entity. This seems both unnecessary and, insofar as it bogs us

down in the weeds of category-based metaphysics, distracting. We can do all of the necessary explanatory work without making this move.

In short, I think that each of these possible rejoinders from the perspective of a category-based view fails. I conclude that a traits-based view is the best available explanation for the assignment of gender norms to entities in the social world.

3.6. CONCLUSION

In this chapter I have outlined my positive view about how gender norms come to exert force over entities in the social world; the *traits-based view*. I began by articulating six desiderata for such a view: it should explain (1) responsiveness to gender norms, (2) evaluability under gender norms, (3) social construction of gender norms, (4) nonconsensual assignment of gender norms, (5) differential assignment of gender norms often on the basis of sexed embodiment, and (6) oppressiveness of gender norms. I then described the *traits-based view* at length. On a traits-based view, entities (such as individuals, groups or societies) are responsive to and evaluable under gender norms because of the traits they are perceived to display. Traits are understood as bare descriptive facts, like *wearing a dress, having ovaries,* or *exerting force in pursuit of one's interests.* Particular traits are gender-coded as either *masculine* or *feminine;* if one has the trait, one is expected to follow the norm. What matters for gender norm assignment is not whether the entity in fact has the trait, but whether they are perceived to have the trait by others. Gender norms often cluster around particular traits which serve socially relevant functions, such as reproduction or food preparation. When this happens, *cluster coherence* is may be normatively enforced. That is, traits which "match" a particular cluster are expected to "cohere" or occur together, and non-coherence is punished to varying degrees.

I discussed how this view can meet the six desiderata I have identified. In doing so, I drew out further elements of the traits-based view, including the nuances of how it is consistent with *internalization/habituation,* and some possibilities it offers for theorizing resistance to dominant norms. Finally, I considered three objections to the traits-based view: the *objection from multiplied categories,* the *causal story objection,* and the *why not both? objection.* Each objection argued, from various perspectives, that the category-based view has an explanatory advantage over the traits-based view. In each case, I gave reasons to think that the category-based view does not, in fact, have any such advantage. I concluded that the traits-based view is a better available explanation for the phenomena at hand.

In the next chapter, I will explore what it means for a gender norm to be *authoritative* over one's conduct. Drawing from the existentialist tradition, I will articulate a notion of *authenticity* on which gender norms can be authentic while also being socially constructed. I argue that authenticity has normative authority over our conduct. Thus, gender norms can give us authoritative reasons to act.

4.0. INTRODUCTION

In his autobiography, *Becoming a Man,* P. Carl describes his experience navigating the

demands of masculinity as a transgender man. He writes,

> I am in one way "becoming" a man. But in another way, I have always been one,
> and I'm trying out all the ways to live as one, some good, some bad. One night I was in a
> Lyft talking to a guy who was a dental technician trying to join the Navy. He told me he
> was doing it "for his woman." "I think she's the one," he said tentatively. "They only
> want your money, and I've told her I haven't got any, but I'm making her sign a prenup
> anyway."
> I heard myself say "Yeah, man, I feel you—all that bullshit about women's
> rights." He laughed and said "Yeah, you know, my man, you know what I'm saying." I
> tipped him $10 and gave him five stars for letting me indulge my inner sexist jerk. (Carl
> 2020a)

This story will evoke a familiar, complicated feeling for many readers. Most of us find ourselves

subject to gender norms which demand that we behave in ways that are *bad,* for us or for other

people. Many gender norms are patriarchal, hierarchical, racist, sexist, and cisnormative; they

recommend behavior contrary to our moral sensibilities, our self-interest, our preferences or

needs. As discussed in Chapter 2, transgender (trans) and gender-nonconforming (GNC) people

often find ourselves responsive to gender norms which ostensibly have not been assigned to us.

Moreover, many such norms are deeply flawed in ways of which we are (painfully) aware.

Nevertheless, these norms can feel right; they can feel like *ours* in a way that the assigned norms

never did. Put differently: many trans and GNC people experience ourselves as caught between

normative standards, where morality, self-interest, or social expectations pull in one direction,

but a deep and powerful sense of *authenticity* pulls in another.

In Chapter 3, I argued that people are responsive to gender norms because of the traits

that they are disposed to express. I noted that, for many people, this may mean being responsive

to *multiple* sets of norms. For example, a transmasculine person might experience himself as responsive both to norms of masculinity and femininity, insofar as he has or is disposed to express traits that are masculine-coded and traits that are feminine-coded. However, this leaves behind a lingering worry in terms of explaining the experiences of trans and GNC people. For many of us, there is a distinctive normative difference between our experiences of responsiveness to different gender norms. Some varieties of norm responsiveness (often, to norms that match our assigned categories) feel perfunctory and instrumental. We follow them because we understand that we will face sanctions (disappointment, censure, violence, and so on) if we do not. Nevertheless, through long habit, we may still become *habituated* to following them and feel responsive to them. For example, Wilchins (1997) writes:

> To avoid displaying any of the "inappropriate" and prohibited signs about myself, I policed myself from feeling them, lest I give myself away with a gesture, a stance, or anything that would allow others to single me out and make me a target for social retribution…. At certain junctures I didn't need to police myself; there didn't seem to be any choice. I acted "masculine" in those circumstances not because I was forced to, but because that seemed, in some inexpressible way, to be what I was. (132)

However, other varieties of norm responsiveness (often, the norms that do *not* match our assigned categories) feel like they have *authority* over us; they "really" tell us what to do or give us reasons to act, in ways that guide us beyond, and sometimes even in conflict with, our fear of sanctions. This difference seems normatively significant. In this chapter, I argue that this sense of normative authority is explained by the *authenticity* of certain norms. Since authenticity has normative authority over our conduct, authentic norms can be reason-giving in a way that other norms may not be.

Most dominant gender norms are poor standards for conduct. Macroscopically, the norms encoded in dominant cultural practices encode a system of colonial gender hierarchy into

practical law, and thereby provide guidelines for the enforcement of that hierarchy. Microscopically, the demands of particular gender norms often conflict with what is morally good, or good for us as individuals. For example, some dominant norms of masculinity excuse violence and misogynistic behavior but discourage healthy emotional development, while corresponding norms of femininity promote passivity and self-abnegation. This creates a "boys will be boys" cultural context where masculinity is bound up with misogyny and poor mental health, while femininity encourages tolerance of bad behavior by privileged men at a cost to oneself. As Carl's account demonstrates, this context makes it difficult to be masculine and morally good—or to be feminine and maintain a healthy self-interest. For those of us with both feminist commitments and a persistent sense of the authenticity of certain gender norms, navigating this territory is a persistent theoretical, ethical, and practical challenge.

I hold that this felt sense of authenticity tracks something real. Gender norms, including harmful or oppressive gender norms, can be authentic for a person. This is potentially complicated by the fact that, as I have articulated in previous chapters, I am committed to the claim that gender norms are *socially constructed,* and, moreover, that many domains of gender norms are *oppressive.* If a gender norm is oppressive, then in many cases, following it will encourage one to do things that are morally or prudentially bad. Under these conditions, what does it mean for a gender norm to be experienced as *authentic*? What is the relationship between authenticity and other normative standards such that something morally or prudentially bad could be authentic? If some gender norms are morally or prudentially bad, can they nevertheless give us reasons to act?

These questions arise for anyone who experiences gender norms as authentic, regardless of whether those norms were assigned to them at birth. Trans and GNC people are not "problem

cases" here; gender norms affect nearly everyone. However, this inquiry will center trans and GNC people, for two reasons. First, the need to explain gender norm authenticity is more pertinent for those who do not fit comfortably into dominant gender categories and are brutally punished for it. Attending to our experiences of gendered authenticity can be the difference between finding the social and hermeneutical resources one needs to have a livable life, or not. Second, and pursuant to the first, many subaltern trans and GNC communities have engaged in important work towards solving the practical dilemmas created by gender norms. The discursive and normative practices present in many such communities therefore represent important sites of resistance to the harms of dominant norms.

In short, many (though by no means all) trans and GNC people experience some gender norms as authentic for us. However, the nature of this authenticity remains unclear. A better understanding is needed to make sense of these experiences and the prominent role they play in self-understanding and practical deliberation. Moreover, as I will show, certain interpretations of these experiences characterize them as anti-feminist. I will argue that this is a mistake which rests on confusions about the nature of authenticity as well as a mischaracterization of trans subjectivity.

The primary goal of this chapter is to get clear about what is being said when trans and GNC people claim that gender norms are authentic for us, in order both to explain how these norms can be reason-giving for us, and to demonstrate the *lack of tension* between these claims and central commitments in feminist philosophy. I defend the following main thesis:

> **Gender Norm Authenticity.** Gender norms can be authentic for a person, and thereby generate authoritative normative reasons for that person to act, *even though* some gender norms are morally or prudentially bad, and all are socially constructed.

I take Gender Norm Authenticity to be a desirable position for anyone interested in theorizing inclusively about gender liberation. To support this claim while maintaining key commitments from feminist philosophy, I defend the following two claims:

Social Authenticity. Authenticity is a project of constructing an intelligible self out of available materials from one's social context.

Authoritative Authenticity. Authenticity has *normative authority* over our conduct, but this does not mean the actions it recommends are (morally or prudentially) *good*.

On Social Authenticity, social phenomena can be authentic insofar as they are "owned" and incorporated into one's self-concept in the right way, an active and constructive process. I draw here from a family of views in existentialist philosophy, such as those defended by Heidegger (1962), Taylor (1991), and Guignon (2004). On Authoritative Authenticity, authenticity gives us powerful reasons to do what is authentic for us, even if there are other powerful reasons (such as moral or prudential reasons) not to do it. Together, these claims capture trans and GNC expressions of authenticity, while allowing that gender norms are socially constructed and that many domains of gender norms conflict with morality or prudence.

In 4.1, I articulate the apparent tension between commitments in feminist philosophy and Gender Norm Authenticity. I identify two important claims about gender norms in feminist metaphysics, and show how authenticity claims are sometimes taken to challenge both. In 4.2, I present Social Authenticity as a plausible positive view and show how it is already implicit in many trans and GNC narratives. In 4.3, I defend Authoritative Authenticity. I outline three ways of thinking about authenticity as a normative standard, and show how, on each view, authenticity need not be always morally good or good for us. In 4.4, I defend Gender Norm Authenticity. I synthesize preceding sections to outline the way gender norms can be authentic and thus authoritative over individuals' conduct, even when those norms are the appropriate target of

153

feminist critique. I articulate the practical dilemmas this can generate. I argue that these dilemmas are faced by all who function in a non-ideal world, and that trans and GNC people regularly do material, epistemic, and discursive work to solve them by creating spaces for agency and self-understanding within trans worlds of sense. In 4.5, I conclude with a brief discussion of authenticity's role in gender liberation.

4.1 FEMINIST PHILOSOPHY AND GENDER NORMS

In this section I articulate the apparent tension between key commitments in feminist philosophy and authenticity claims about gender norms, or "authenticity claims" for short. While authenticity claims are by no means universal in trans narratives, they are common—particularly when we talk about discovering, exploring, or constructing identities. People might talk of being *real* or *true to themselves*, or of reflecting *who they really are* in body, dress, presentation, comportment, name, pronouns, and so on. Gender norms are often divided into standards of *masculinity* and *femininity*. For example, imagine a black-tie event in New York City, where the feminine norm recommends wearing a gown and the masculine norm recommends wearing a tuxedo. An authenticity claim here might be the following: "I know that I am supposed to wear a gown to this event, but wearing a tuxedo feels more like *me*." Or: "I won't be wearing a dress tonight, because *that's not who I am*." Statements like this are often central to public-facing trans narratives. For example, the March 2021 issue of TIME Magazine features transmasculine actor Elliot Page on the cover, dressed in traditionally masculine clothes and sneakers, and sporting a flat chest and a short haircut, with a quotation which reads "I'm fully who I am" (Steinmetz 2021). Put simply: Authenticity claims hold that some gender norm or set of norms is or is not authentic for you, that the force of that norm is or is not "true" or "real" for you.

Broadly speaking, gender norms are *the evaluative standards associated with gender:* they are the expectations, constraints and enablements, and spoken and unspoken rules which apply to individuals in virtue of their position in gendered social space. Since gender practices vary widely across cultures and contexts, gender norms do as well. However, feminist analysis has paid special attention to what I call *dominant gender norms*. Dominant gender norms are the norms which operate in the mainstream cultural contexts that exert power over the vast majority of social and material resources. That is, they are the gender norms operative in what Lugones (2007) calls the "colonial/modern gender system"—a division of resources and roles according to hierarchical, constructed categories of sex and race. Due to a legacy of imperialism, psychological oppression, and epistemic violence, dominant gender norms infiltrate lives, minds, and practical decision procedures across contexts, even when we work to escape or resist them. As a result, many trans and GNC people are forced to grapple with dominant gender norms in order to make sense of ourselves and our experiences. We are raised in dominant contexts and shaped by them; we may have work, family, or other practical responsibilities there. In short, although dominant gender norms are not the only game in town, they are often the most powerful and pernicious.

As discussed, feminist philosophy has historically criticized dominant gender norms for contributing to the oppression of women. For example, Young (1980) argues that the enforcement of feminine norms in dominant post-industrial societies trains women to treat their body as a "fragile encumberance" (144), something to be "looked at and acted upon" rather than lived through (148); this creates a restricted and awkward mode of existence and restricts women's embodied agency. Similarly, Manne (2017) argues that the primary function of feminine norms is to circumscribe and enforce a subordinate social role for women and girls

across contexts. Moreover, Black feminists such as hooks (1982), Spillers (1987), and Collins (2000) have argued that gender norms have distinctively racialized harms. Black women have historically been subjected to strict codes of femininity, while at the same time being systematically represented in stereotypes which fail at those standards, so that it is "extremely difficult and oftentimes impossible for the black female to develop a positive self-concept" (hooks, 1982, 86).

More broadly, feminist philosophers have criticized dominant gender norms for being oppressive *at root*. As Lugones (2007, 2010) and others have argued, dominant gender practices are rooted in colonialism, and thus in the subordination of non-white colonized peoples. They also subordinate white women, although the experience of white women is integrated with racial privilege, and thus differs substantially both in kind and in level of brutality from the subordination of colonized peoples. Dominant gender norms maintain racist colonial structures of power by designating heteronormative whiteness as central to good gender performance, and thus to full humanity; non-white peoples are brutally punished for failing to conform to norms from which they are categorically excluded (Lugones, 2007, 205; see also Spillers, 1987; Espiritu, 1997). Within this framework, gender norms function to justify and reify interlocking systems of oppression across the colonized world.

We might conclude here, as many feminist philosophers have done, that we ought to eliminate, or at least change, these norms. But if it's true that we *ought* to change or eliminate them, it must also be true that we *can*. There is a widespread belief that gender-normative behavior is innate or essential to human existence, and thus cannot be changed. But this conflicts with *social constructionism,* which I have identified as a key feminist position, and thus a basic commitment of this project. As discussed in Chapter 1, feminists are therefore often engaged in

what Haslanger calls the *debunking project* (Haslanger 2012, Ásta 2018). To recap: the *debunking project* arises from a normative position about the badness of dominant gender norms, together with the metaphysical position that they are socially constructed. Debunking social constructionists "argue that the rituals or practices in question are unjust and should not be maintained in their current form and that the supposed metaphysical or natural justification for them is misguided" (Haslanger, 2012, 127). The aim is "to show that certain claims to objectivity are unfounded and that any social organization based on such claims is thus unjustified" (Ásta, 2018, 58). The debunking project therefore makes a normative claim that dominant gender norms *ought* to be changed or eliminated, because they are both bad in themselves and bad for people who follow them. Concordantly, it makes a descriptive claim that gender norms *can* be changed or eliminated, because they are socially constructed, rather than essential and inescapable facts about human nature.

Authenticity claims in dominant contexts appear to raise difficulties for a debunking project, for two reasons. Consider, first, the debunkers' descriptive claim that gender norms are socially constructed. The notion of authenticity is sometimes taken to appeal to one's "inner self," and to locate what is authentic for a person within that self. This inner self is understood as innate, immutable, and socially unmediated. If this is right, then authenticity claims about gender norms suggest that those norms are somehow innate, immutable, and socially unmediated. A debunking view, which places gendered phenomena in the external social world, precludes this understanding.

On the other hand, if gender norms are socially constructed, it is hard to make sense of authenticity claims. To see this more clearly, we can reverse the dialectic. A gender essentialist has an easy explanation for claims about gender norm authenticity. They can simply say that

people have essentially gendered features, and those features motivate people to respond to gender norms; one's "internal gender" matches up with some external norms. On this view, masculinity and femininity are essential features of the self, and the social aspects of gender, such as gender norms, arise out of a translation of those essential features into socially intelligible characteristics. This need not entail that gendered features are related to biological sex. For example, recall that Serano (2007) holds that there are essential masculine and feminine inclinations which occur irrespective of sex: "While variations in our sex characteristics and gender inclinations may occur naturally, the way we interpret those traits, and the identities and meanings we associate with them, can vary significantly from culture to culture" (101).

This would spell trouble for a debunking project. If individuals across contexts have innate, essential features that *just happen* to match up with socially constructed gender norms, that is either a stunning coincidence, or strong evidence that gender norms are importantly connected to innate, essential features of humans. But social constructionists deny just this. Put this way, it seems the essentialist has a better explanation of gender norm authenticity than the social constructionist does; gender norms feel authentic because they are the social expressions of deep, essential gendered features of individuals. This objection has been raised by trans theorists such as Serano (2007) and Prosser (1998). If authenticity is wedded to essentialism, and gender norms can be authentic, then something about gender must be essential. Conversely, if authenticity is wedded to essentialism, and gender is *not* essential, authenticity claims about gender norms either point to a truly remarkable pattern of coincidence, or they seem straightforwardly false. Either way, authenticity claims seem inimical to social constructionism.

The second apparent tension between the debunking project and authenticity claims is normative. Recall the debunker's normative claim; dominant gender norms ought to be changed

or eliminated. Authenticity is often understood as a moral notion. What is authentic is thereby taken to be *good*. To claim that any norm is authentic, then, seems to entail that it doesn't need to be changed or eliminated. Then, if the authenticity claims in question are about dominant gender norms, they are incompatible with the debunker's normative claim.

One immediate answer to this challenge is to point out that not all gender norms are created equal. While dominant gender norms are oppressive and harmful, other domains may escape, or even be constructed specifically to resist, these oppressions and harms. For example, drawing from Lugones' work on *worlds of sense* (1987), Bettcher (2014) argues that there are resistant trans contexts in which we negotiate "alternative gender practices" that center trans experience and subjectivity (389-90). In these worlds of sense, domains of gender norms—masculinities, femininities, or other standards altogether—may arise that are non-coercive, non-hierarchical, and non-oppressive. They may center rather than marginalizing trans experience; they may actively resist norms of whiteness and coloniality; they may prioritize self-understanding and authenticity rather than gendered and racialized hierarchy.[18] Many authenticity claims draw on the gender norms produced in these worlds. To ignore the rich histories of trans and GNC communities and interpretive practices is to actively misunderstand a great deal of trans and GNC subjectivity and self-understanding.

However, not all trans or GNC experience of gender norms is grounded in alternative worlds. Among the most difficult challenges many of us face is the fact that dominant, harmful gender norms have power over us and our choices. Consider, for example, P. Carl's account of

---

[18] For example, Marion (2011) describes alternative normative practices which arise from *ballroom communities*. According to Marion, gender practices in these communities actively subvert and resist dominant gender norms to create space for queer and trans subjectivity, as well as to prepare participants to move through dominant worlds. I say more about ballroom culture in successive sections.

responding to the demands of a toxic masculinity (Carl 2020a, 2020b). Carl describes himself, a

feminist with a PhD in gender studies, as nevertheless grappling with the persistent sense that

harmful, misogynistic, white American masculinity is a part of him—something he must reckon

with to understand and communicate himself as a man.

> Similarly, recall again Mock's account of her early experiences of femininity:
>
> Like most teen girls (whether they're trans or cis), I had a vision board of my ideal, pulled mostly from the pop-culture images that MTV had fed me. I wanted Halle Berry's or Tyra Banks's breasts, Britney Spears's midsection, Beyoncé's curvy silhouette and long hair, and I prayed that I wouldn't grow any taller so I didn't tower over the petite Asian girls who were the barometer of beauty in the [Hawaiian] islands. (Mock, 2014, 122-3)

The norms Mock is describing are not drawn from subaltern, trans-friendly standards, but

straightforwardly dominant standards guided by "pop-culture" and "MTV." This demonstrates

that, insofar as trans and GNC people find that certain gender norms resonate with us, this can

sometimes occur in dominant worlds. Most of us are not born and raised in resistant contexts.

The gender norms with which we have formative experiences are, very often, not expansive,

inclusive, or affirming, but restrictive, exclusionary, and coercive. As such, authentic self-

understanding can require reckoning with the power these experiences have over who we are.

Put simply, while many authenticity claims are about gender norms that are not

the appropriate subjects of feminist critique, some *are*. As Carl and Mock both demonstrate, it's

not as if trans and GNC people don't know this. We are deeply, materially, brutally aware that

the dominant gender norms which sometimes shape our choices are harmful. Nevertheless, many

of us have had the unsettling experience of feeling as if some dominant gender norm, a piece of a

hegemonic and harmful system of which we are often the most prominent victims, is an authentic

part of us. What appears as a theoretical conflict for feminist philosophy is therefore matched by a genuine practical dilemma for many trans and GNC people.

In the next section, I reject popular "inner self" views of authenticity as implausible, and, drawing from existentialist views defended by Heidegger (1962), Taylor (1991), and Guignon (2004), articulate a view of authenticity as socially embedded. I show that what I call *social authenticity* is already implicit in many trans and GNC narratives. This will dissolve a tension I have identified between commitments in feminist philosophy and trans and GNC authenticity claims. Since authenticity is a socially embedded phenomenon, authenticity claims about gender norms do not entail the claim that gender norms are not socially constructed.

4.2. SOCIAL AUTHENTICITY

The intuitive idea of authenticity, in the words of Bernard Williams, is "the idea that some things are in some real sense really you, or express what you [are], and others aren't" (Jeffries, 2002). Colloquial definitions include "being true to oneself" or "being who one really is." Terms like "true" and "really" are sometimes taken to suggest the presence of something distinct from one's outward performances, which are false or façade. According to Guignon (2004), this "modern ideal of authenticity," as defended by cultural figures such as Oprah and Dr. Phil, relies on a presupposition that "lying within each individual, there is a deep, 'true self'—the 'Real Me'—in distinction from all that is not really me" (3). This "true self" consists of innate, socially unmediated traits which represent "who you really are". To live authentically, one must "find" or "get in touch with" oneself—indicating that there is something covered or hidden, with which one can be *out* of touch, and which exists apart from the influence of corrupting external factors such as relationships and social roles.

This picture is unsatisfying. Given humanity's deeply social nature, philosophers have questioned our epistemic access to this core self—if indeed there would be any recognizable "self" left after all of the "external" factors are stripped away (Adorno, 1973; Rorty, 1989). Others have argued that the "core self" is not merely a fiction, but a harmful one. Foucault (1983) argues that the myth of the hidden self encourages the individual to waste time trying to find it, rather than engaging in the crucial project of *creating* the self. Taylor (1991) argues that the notion of authenticity as *self-fulfillment*, understood as discovering and satisfying one's deep desires, is self-indulgent. Furthermore, Guignon (2004) notes themes in art, literature, and philosophy suggesting that humans are not fundamentally good or pure, but that "what lies within is characterized by aggression, cruelty and violence" (54). That is, even if there is a "pre-given self," perhaps it is something we don't want to be true to.

What I am calling the "inner-self view" is deeply implausible. As such, I argue that it is uncharitable at best to represent all authenticity claims as presupposing it. Many trans and GNC people understand our relationship to gender and gender norms as deeply embedded in and influenced by our upbringing and culture—and no less authentic or real for it. Consider, for example, Janet Mock's understanding of her own identity:

> I am aware that identifying with what people see versus what's authentic, meaning who I actually am, involves erasure of parts of myself, my history, my people, my experiences.... When I think of identity, I think of our bodies and souls and the influences of family, culture, and community—the ingredients that make us. (Mock, 2014, 249)

Mock does not understand her authenticity as stripping away her social context in order to uncover a gendered inner self. Rather, she describes her authentic self as in part *made of* social context; her history, people, experiences. This process is not seen as stable or complete, but as ongoing and constructive (230). At the same time, her authenticity is fundamentally in tension

with certain social expectations. Mock's identity as a Black trans woman is a way of representing herself as intelligible to others in dominant contexts; but this process requires obscuring elements of herself, and thus is never fully authentic. Mock's understanding of authenticity here is consistent with an intuitive understanding of authenticity as distinguishing what is *one's own* or what is *true to oneself* from what is not, while at the same time reflecting the socially embedded nature of the self.

To make sense of the richness and complexity of gendered experience, our philosophical view of authenticity ought to capture both of these features. I propose, then, that we adopt a view I call "social authenticity," on which the authentic self is constituted by a certain kind of relationship to one's social roles, relationships, and commitments. Views like this are defended by many in the existentialist tradition, particularly Heidegger (1962), Taylor (1991), and Guignon (2004). In what follows, I will articulate some common commitments among these views, and show how they satisfy these desiderata.

For Heidegger, humans are constitutively social beings. We do not exist prior to or apart from our situations. Rather, we "always-already" find ourselves embedded in a social context, with a past, a perspective, and roles, tendencies, and traits, all of which come loaded with social meanings. Indeed, the expression "find ourselves" is misleading; we do not *find* our "selves" as pre-given entities, but are constantly making ourselves through our decisions. We must then choose whether to take responsibility for these decisions, or not. As Carman (2003) puts it, "I am handed my existence, but then I have to face up to it or not: 'To be or not to be'" (289).

Heidegger's word for "authenticity", *Eigentlichkeit,* is perhaps best translated as "owning" or "being one's own" (Varga and Guignon, 2020). For Heidegger, authenticity is constructing oneself in accordance with one's "that-for-the-sake-of-which"—one's overarching

narrative, or life project. It is easy enough to get lost in the "average everydayness" and to just do "what one does;" this is *inauthentic,* and according to Heidegger, most of us are inauthentic most of the time. But authenticity does not require separating oneself from these concerns. Everyday tasks are done authentically insofar as they are "owned." I am authentic insofar as I recognize my activities, concerns, relationships, and roles as *mine,* done "for-the-sake-of" my life project, and insofar as I am willing to commit to and defend them as my own. Authenticity is a matter of "owning up to the concrete situation in which one finds oneself and understanding one's being explicitly *as* one's own" (Carman, 2003, 297).

In a similar vein, Taylor (1991) argues that human life is fundamentally shared:

> We become human agents, capable of understanding ourselves, and hence of defining an identity, through our acquisition of rich human languages of expression… The genesis of the human mind is in this sense not "monological," not something each accomplishes on his or her own, but dialogical. (33)

Taylor holds that the self is shaped by the background of values and meanings against which it is formed. The authentic self is therefore not an entity, but a process. In this process, one navigates the practical realities of existing as a somewhat disjointed socially embedded first-person perspective with an always-already existing plethora of everyday concerns, while concurrently trying to make sense of that existence. We aren't most like ourselves apart from our social context; rather, we are *made of* context, or, rather, constantly *making ourselves* from context. The "true self" of authenticity is not radically distinguished from the social world. It is a product and a part of the social world.

One might worry that this picture generates what Korsgaard (2009) calls the "paradox of self-constitution." How can we understand this project of making oneself, unless there is already a "self" there prior to the making to *do* the making? That is, how can one be "a craftsman who is,

mysteriously, his own product" (Korsgaard, 2009, 42)? In a post on social media coming out as trans, actor Elliot Page appears to evoke this concern when he writes "I can't begin to express how remarkable it feels to finally love who I am enough to pursue my authentic self" (Page, 2020). Here, Page is describing themself as both *already being* someone—that is, already having a self—and also beginning to *pursue* an authentic self. They do not, however, seem bothered by the paradox of self-constitution, and their phrasing can help us explain why. We can distinguish here between a broad class of facts about oneself, and the authentic subset of those facts. Authenticity is about owning up to and standing up for what may already be facts about oneself. When Page writes that he loves who he *is* (present), we can interpret him as saying that he loves his broader self; his traits, his values, and the other facts that constitute him. When they write that they mean to *pursue their authentic self* (future), we can interpret them as saying that they mean to own parts of themself which they have previously denied, a process which will help them make sense of themself.

It is important to recognize that social authenticity does not entail the view that the self is entirely freely chosen or "up to us." There is certainly an element of choice here—as Heidegger puts it, a "choosing to choose a kind of being-one's self" (1962, 314). Every choice I make is a choice about what kind of person I want to be. However, these choices are constrained by my *facticity*, or the set of social and psychological facts which constrain my options. I may be able to choose whether to marry Frida or Fred, but I cannot fully choose the implications of this choice, its effect on the world around me, or even what it means for my own emotions and dispositions. Moreover, one's choices have value only against background conditions of intelligibility—what Taylor (1991) calls "horizons of significance." For our choices to be meaningful, they must work with existing social meanings that others in our social milieu can understand and interpret. To

choose authentically, then, I must be aware: both of the significance of each choice, and that the choice itself is significant; it constitutes me. My choice is free, but not unmoored.

A key feature of the social view, then, is that authenticity requires a relatively clear-eyed understanding of one's own history, psychology, and relationship to others. One must know one's situation in order to choose well. Moreover, one be aware of one's own possibilities and limitations as one always-already finds them to be. Call this the *epistemic condition* on authenticity. You cannot own your facticity unless you know what it is you are owning.

However, an awareness of facticity does not commit us to merely accepting the world we are given. Ortega (2005) points out that our horizons of significance tend to be deeply flawed. We often "inherit certain possibilities that should not be repeated; we are members of communities with a past full of bloodshed, racism, and countless unmentionable acts" (28). Authenticity might therefore require *resisting* or *changing* certain shared social meanings or values. Taylor (1991) argues that the ideal of authenticity is admirable precisely because it enables the original contribution that each individual can, and should, make to the whole. In a democratic society, each of us is called on to contribute to the community of ideas from their own authentic commitments; this is essential to the public good. The ideal of authenticity therefore encourages each individual to take responsibility for their actions, and to critically evaluate, challenge, or change the norms and values which guide those actions.

This process therefore involves not just recognition of the self as a part of the world, but an active, creative process of engagement with one's world. Authenticity requires a productive originality that can only be realized by the person in question, in their own unique situation. Authentic selves create art, philosophy, scientific inquiry, and social change. The relationship between authentic self and context is a reciprocal one. When the self is fundamentally a part of

the world, to paint a picture or compose a symphony or propose an idea or start a movement is to create one's self *by* creating one's world.

This creative process can be crucial for building an intelligible life. The world does not give marginalized people the materials we need; so, we must make the world into what we need it to be. This, I suggest, is constantly ongoing in trans lives and worlds of sense, out of necessity. In *Gender Outlaw,* Kate Bornstein articulates this kind of construction of the self through art and activism that change her social context. Bornstein thoroughly rejects an understanding of her gendered self as "fixed" or "given." Instead, she describes herself and her gender as "patchwork"; she writes, "I learned to live my life like I'm making a collage" (148). The metaphor of "collage" here suggests a creative construction out of available elements. The created image is new, but it does not spring fully formed from the void. It comes together out of existing colors, shapes, and images, cut up and glued back together.

Queer and trans communities have long histories of engaging in this project. Consider the gender practices prevalent in *ballroom cultures*, queer and trans communities of color which have evolved in most major cities in the U.S. Ballroom cultures have their own gender classification system and accordant gender norms. In his study of Detroit's ballrooms, Bailey (2011) argues that these practices "are the result of a considerable amount of work, a form of discursive labor that often goes unnoticed and taken for granted by those outside of the community" (371-2). These identities and norms are created out of two needs: to make oneself intelligible to oneself and others both within the community and without, and to avoid the violence visited on the visibly queer in the broader world. Thus, constituents of ballroom culture subvert or creatively re-purpose dominant gender norms, at the same time as they prepare one another to function effectively within them.

The creative labor of building alternative gender practices, I argue, grows out of a need to own up to parts of oneself that are not intelligible to others in the dominant context. The process is beholden to shared intelligibility; it must trade on existing social meanings in order to make sense. For example, new pronouns are sensible only in an existing grammatical space, in contrast to existing pronouns. Just as one might compose a piece of music out of existing chords and tones, trans and GNC people historically create intelligible gender terms and identities for themselves, and spaces for cultural exploration, out of existing "gender materials"—terms, practices, categories, and norms.

In sum: On Social Authenticity, a person is given a situation, a set of personal, social, and psychological facts. That person becomes authentic by *owning up to* and *standing up for* themselves, an active, constructive, and often creative process (contrasted with the passive, reflective process of *finding* themself). This project is not undertaken in a vacuum of value, but rather against a background of shared intelligibility. Social authenticity doesn't posit a pre-given self, but neither does it mean that we have a radically free choice about what is authentic for us—insofar as we don't always choose what we believe or care about, or what our context is like. A social view captures the intuitive idea of authenticity without encountering the problems faced by an "inner-self" view. On social authenticity, what is authentic is yours not merely because it is given to you, but because it is *made* yours.

It should be clear from this discussion that, if we adopt Social Authenticity, feminist commitments about the social construction of gender norms are not in conflict with authenticity claims. Gender norms and our responsiveness to them are part of the "raw materials" of our facticity. We do not need to make any claims about their essential origins to acknowledge that they can be authentic for individuals.

In the next section, I will discuss the *authority* of authenticity as a normative standard. The fact that something is *authentic* for someone gives them a strong reason to do it. But what is the source of this reason? I will give several possible answers, and show how none of them entails that what is authentic must also be, on the whole, prudent or morally good. This will defuse the normative tension between feminist philosophy and trans authenticity claims.

4.3. AUTHENTICITY AND NORMATIVE AUTHORITY

Metaethicists sometimes distinguish between two kinds of normativity. All normative domains have the first kind: *formal normativity.* Something is formally normative just in case it is possible to succeed or fail according to its guidelines. Consider, for example, the normative domain "NCAA basketball rules." A group of people on a court with a basketball might follow the standard of NCAA basketball rules; they might follow some other standard, such as the norms of the game HORSE; they might make up their own rules; or they might follow no rules at all. The choices are entirely up to preferences and goals of the players in question.

However, certain normative standards, such as morality or prudence (self-interest), appear to have a special power over our choices; they "really" tell us what to do. That is, they give us reasons to act that weigh heavy against other reasons we may have. This is sometimes called "robust" or "authoritative" normativity.[19] As Paakkunainen (2018) puts it,

> Formal normativity is cheap: we can create new formally normative standards simply by inventing violable rules. Robust normativity is a seemingly more important phenomenon that many take to be associated with *normative reasons.* (403)

---

[19] These terms are roughly interchangeable in the literature. I use the terminology of "authoritative normativity" in this chapter to better parallel talk of reasons or standards having *normative authority.*

A standard might be authoritative in some cases, but not others. For example, college players during a championship game may have strong reasons to comply with NCAA rules, while children playing after school, or those same college players during the off season, may not. However, some normative standards seem to be authoritative *in themselves.* Morality or prudence appear to give us powerful reasons, full stop.

In this section, I hold that authenticity has normative authority, but distinguish this from the claim that this is *moral* authority. If authenticity has moral authority, then claims about authenticity are also claims about morality. This appears to put some claims about gender norm authenticity in tension with feminist commitments. I argue that this tension is illusory. I briefly explore three plausible views on which authenticity has normative authority. On each view, authentic norms can conflict with moral norms. Therefore, even if some gender norms are authentic and thereby provide authoritative reasons to act, this would not mean that they are morally good on the whole.

Authenticity is certainly formally normative. We talk of being more or less authentic, fully authentic, or altogether inauthentic. Moreover, authenticity appears to hold powerful sway over our conduct. The fact that something is authentic for us seems, by itself, to provide us with a powerful reason to pursue it. For example, suppose that I must choose between a career as an accountant and as a professional dancer. Suppose that I have a plethora of good prudential reasons to choose accountancy; it is a lucrative, respectable profession, at which I can anticipate a productive and stable life (all things that I want). Dancing does not have these attributes. However, suppose that being a dancer feels authentic for me, and accountancy doesn't. Intuitively, the choice is a difficult one precisely because authenticity generates strong reasons to act, reasons that weigh heavy against other strong reasons.

A full discussion of the nature of normative authority is beyond the scope of this project. However, it is worth noting that some prominent existing views would identify authenticity as authoritative. Varieties of *constructivism,* for example, hold that normative force derives from the agents who are bound by it. Roughly, the idea is this: since authoritative reasons are characterized by their power over our practical deliberations, normative authority for agents is importantly related to the practical interests of those agents. If you have an authoritative reason to φ, that is because of something about *you* and your relationship to φ. As Street (2010) puts it, "the bumper sticker slogan of constructivism is…'no normative truth independent of the practical point of view'" (367). On constructivist views, prudence and morality are authoritative because (or, depending on one's view, *insofar as*) we care about our own interests and those of other agents. Plausibly, on a constructivist view, authenticity is authoritative because (or insofar as) we care about *being* ourselves—being the kind of agent who has interests that are their own.

This is one reason authenticity claims have wielded such rhetorical power in discourse about queer and trans rights. If I say that I am coming out as queer because it is "who I really am" or that I am following some gender norms because they reflect my "true self," that provides an explanation of my reasons for doing it that calls for no further justification. An understanding of sexuality or gender identity as a part of a person's authentic self has lent legitimacy to a movement for social acceptance—precisely, I think, because we understand that authenticity gives us powerful normative reasons.

This is sometimes interpreted as equivalent to giving us *moral reasons*. But this needs further argument; not all powerful reasons are moral reasons. How, then, should we understand authenticity's normative force? Here are three plausible answers. 1) It is in one's interest to be authentic; authentic norms are prudential norms. 2) It is morally good to be authentic; authentic

norms are moral norms. 3) Authenticity's normative force is not reducible to morality or prudence.[20] In what follows, I will briefly discuss each possibility, and show that none of them entails that authenticity must always be in perfect lockstep with (all) moral norms.

First consider prudence. If the standard of authenticity is about respecting one's true self, perhaps authenticity generates reasons of *self-interest*, also known as *prudential reasons*. This is plausible enough. Even if this is right, however, authenticity and other reasons of self-interest can conflict. Abandoning a stable career as an accountant to pursue a risky career as a dancer may be deeply imprudent on the whole, *even if* dancing is genuinely authentic. This does not undermine the possibility that authenticity is *pro tanto* prudent. The authentic subset of one's prudential reasons may recommend a certain action, even if there are other prudential reasons against it. Sometimes different prudential norms simply recommend incompatible actions. Suppose I must choose between marrying Fred and marrying Frida (and I cannot marry both). Either spouse would be equally good for me, but for different reasons; with Frida I would live a life of excitement and adventure, with Fred I would live a life of comfort and stability. I have prudential reasons that pull both ways, but choosing either involves foregoing the other.

A second option is that authenticity is normative because it is *morally good*. This view is defended by Rousseau (1953). For Rousseau, the self is the source of goodness, and society is the source of evil; authenticity requires peeling away the distorting effect of society to uncover the pure moral power of the self. The self is imbued with a pure moral sensibility which is debased by society and its conflicting demands. To be authentic, to be good, one must turn inwards. Similarly, Taylor (1991) defends "the moral force of the ideal of authenticity" (17) on

---

[20] Note that (1) and (2) are not incompatible; authenticity may be authoritative because it represents some combination of morality and prudence (a moral duty to do right by oneself, for example).

the grounds that one has a responsibility to contribute to society from one's personal convictions. According to Taylor, there is something "noble, courageous, and hence significant to giving shape to my own life" (39). Authenticity is therefore fundamentally about moral value.

Notice, however, that authenticity can conflict with moral goodness even if authenticity is a moral ideal. The argument here runs parallel to the prudential argument above. Different moral norms may recommend incompatible actions; this is the source of moral dilemmas. For example, suppose that I have promised to pick my friend up on time from the airport. However, just as I am about to leave, I encounter a lost child who needs help finding his parents. I have, it seems, a *pro tanto* moral reason to keep my promise *and* a *pro tanto* moral reason to help the child. Assuming I cannot effectively do both, I have conflicting moral reasons; as a moral agent, I must weigh my reasons and choose. Similarly, if authenticity generates moral reasons, these reasons might conflict with other moral reasons. I might have an authentic reason to comply with a norm of masculinity that recommends misogynistic behavior, but a (different) moral reason not to do this. As a moral agent, I must weigh my reasons and choose.

A third option is that authenticity is irreducible to either morality or prudence. On this view, if something is authentic for us, that gives us reason to do it—not because authenticity is morally or prudentially good, but because authenticity alone is authoritative over us.[21] If this is right, then there is no theoretical issue if the demands of authenticity conflict with the demands of morality or prudence. On this view, then, what would be the source of authenticity's normative authority? If authenticity is moral or prudential, its authority is conferred by the authority of morality or prudence; but if it is neither of these things, then it must be authoritative

---

[21] Or, perhaps, because authenticity is a subset of some further authoritative normative domain. If this is the case, all of our theoretical problems are solved; authenticity need not be either morally or prudentially good, and it has normative authority.

for a different reason. However, this problem is not unique. Normative authority in general calls

out for explanation. There are intuitive reasons to think that morality and prudence are

authoritative; the fact that something is *morally good* or *good for us* seems, by itself, to give us a

powerful reason to do it. This alone does not explain the source of that authority; a major branch

of metaethics exists to tackle this very problem. I take it to be equally intuitive that authenticity

is authoritative, for similar reasons; if something is *authentic,* that seems, by itself, to give us a

powerful reason to do it.

In short, no matter how we understand the normative force of authenticity, we can

conclude the following: To say that a thing is *authentic* for someone, and that this authenticity

can generate authentic reasons for that person to act, is *not* to say that it is either prudentially or

morally good on the whole. A major element of the normative tension I have identified between

feminist philosophy and authenticity claims about gender norms is thus dissolved.

In the next section, I will combine the ideas of *social authenticity* and *authoritative*

authenticity to articulate influence of gender norm authenticity on our practical choices as moral

agents.

4.4. GENDER NORM AUTHENTICITY

One aim of this chapter is to show that the authenticity of particular dominant gender

norms can produce genuine practical dilemmas, and to show how trans and GNC communities

have done important labor to solve these dilemmas. I have argued that claims about gender norm

authenticity are not in tension with either normative or descriptive commitments of feminist

philosophy. However, this does not solve the following practical problem: What should we do

when gender norms which are authentic for us recommend morally or prudentially bad actions? How do we weigh those reasons?

To address this question, I will first say a bit more about the dilemma itself. While many feminist philosophers have highlighted the issue with *particular* gender norms, such as a masculine norm which recommends misogynistic behavior, I argue that is not the heart of the issue. Notice that it is just as easy to point to particular gender norms that are *not* bad in themselves—a masculine norm which recommends the genuine virtue of *courage*, for example. I contend that the more salient critique is about *entire domains* of gender norms. For example, Lugones' (2007) analysis focuses largely on the structural origins of dominant gender norms as a domain, specifically their construction as bulwarks of the racist, hierarchical "colonial/modern gender system" (187). Relatedly, Manne (2017) argues that dominant feminine norms work in concert to delimit a subordinate role for women across contexts. These are not claims about any particular norm in the set, but rather about what unifies or distinguishes that set.

A normative domain is a set of norms linked by some unifying activity or aim. For example, *formal etiquette* is unified by the aim of maintaining polite society. If the unifying aim of some domain D fails to meet the standard of some other domain E, then D is criticizable *as a domain* by the lights of E. For example, we might criticize the unifying aim of formal etiquette for being *classist*. "Polite society" is deeply racialized and carries undertones of wealth stratification and privilege; compliance with its norms signifies adherence to class hierarchy. Since classism is unjust, formal etiquette as a domain is unjust, and thus morally bad. This will be true even if particular formal etiquette norms on occasion recommend the same actions as particular moral norms, as they almost certainly will.

Many domains of gender norms are analogous to the domain of formal etiquette. For example, following Lugones, we might think that many domains of gender norms are unified by the aim of *maintaining the colonial social order.* Since this social order is unjust, these normative domains are morally bad. Similarly, since the colonial social order is harmful to marginalized people, many such domains will also be *prudentially bad* for those people; it is generally bad for their self-interest to comply. This is distinct from (although consistent with) the claim that individual norms within the domain are morally or prudentially bad.

We can see from this discussion that there are at two ways in which we can evaluate gender norms according to some other domain (such as morality): either the entire domain is evaluated, or individual norms within it are evaluated. Concordantly, there are two ways in which some gender norms can be authentic for someone, in virtue of being evaluated according to the domain of authenticity. First, some entire domain of gender norms can be authentic. For example, a particular domain of masculinity may be authentic for a person if they "own up to" that masculinity, by incorporating it as a standard which they take to be relevant to their behavior. This does not require following every norm in that domain wholesale. On the contrary, it may require that one interact critically with each norm, and balance it with other desires and obligations. Authenticity is difficult and may require negotiating between incommensurable demands. As Jenkins (2016) and Witt (2011) both point out, experiencing a norm as relevant to you is compatible with refusing to comply; in fact, one cannot rebel against a norm unless it applies to them.

Alternatively, individual norms within a domain may be authentic for someone, even if the entire domain is not. Consider, for example, someone who is genderqueer in the sense articulated by Dembroff (2020); they reject some or all domains of gender norms as binding over

their conduct. I suggest that rejecting the domain does not entail rejecting each of its norms. Such individuals may, and in my experience often do, still incorporate individual gender norms as a part of their authentic selves. For example, they may take themselves to have strong authentic reasons to *wear suits*. Insofar as the person understands that their local world of sense recommends *wearing suits* as a masculine norm, and they own up to this fact and understand themselves as masculine, they are incorporating this particular gender norm authentically into their life project.

This kind of authenticity is particularly apparent in the case of people who incorporate multiple gender norms that their world of sense takes to be incoherent together. As non-binary model and activist Alok Vaid-Menon writes: "Over time, I learned that where I was taught dissonance, I found harmony. This beard, this skirt, this love: There are no contradictions here, there is just someone trying to figure it out" (Vaid-Menon, 111). Elsewhere, Menon writes that their femininity is an authentic expression of themselves (26), even as they reject the power of dominant, hegemonic gender norms. There are "no contradictions" precisely because Menon rejects as illegitimate those domains which hold that skirts and beards are incompatible, even as specific norms of masculinity and femininity *within* those domains are understood as their own.

Perhaps most saliently, gender norms can also be *counterfactually authentic* for someone. If authenticity is authoritative over our conduct, we may experience normative pressure to *become* more authentic—that is, to own up to elements of our facticity which we currently do not acknowledge. We can imagine, for example, someone who is persistently responsive to a particular gender norm or set of norms as a matter of psychological fact, but has not yet owned up to this. As Jenkins (2016) points out, experiencing oneself as responsive to certain gender norms is often an impetus for understanding oneself as trans or GNC. This psychological fact

alone would not make a norm authentic in the sense I articulate. However, we might say that, *if* the person were to own up to this fact, it would be authentic for them. Moreover, one might find certain alternative norms strongly and persistently *inauthentic*—that is, impossible to make congruent with one's life project. That person would have authentic reason to *disown* certain gender norms. Often, communicating one's disownment of the norms assigned by others requires actively abiding by some alternative set of norms; for example, communicating one's disownment of femininity in dominant contexts may require performances of masculinity.

If dominant gender norms are morally bad, it is sometimes argued that we should disown them all. For example, Dembroff (2018) argues that we should reject all gender norms associated with a harmful binary. They acknowledge, however, that there may be practical reasons to acknowledge the impact of this binary. For some people, using gendered terms "is important for describing how they were socialised as children, how others interpret their bodies, or how they feel about their own bodies" (2018). Dembroff's account is therefore responsive to concerns of practical choice. They acknowledge that we do not always have the luxury of being practically guided by clean theoretical commitments.

This discussion reflects the realities of living authentically in imperfect worlds. Our normative reasons are not born in ideal settings. Dominant gender norms are ubiquitous, and their enforcement begins very early, as evidenced by the already-discussed case of the *gender reveal party*. Moreover, these norms are mandatory; failing to conform exposes one to censure and violence. Recall again the *epistemic condition:* authenticity requires understanding our surroundings and psychological state, including the norms which already move us. If one finds oneself responsive to a certain domain of masculine norms, for example, one might find that owning that domain, or certain norms within it, is the only way to make sense of one's facticity.

As Wilchins writes, "If the body is always a sign being read, then not communicating is impossible" (1997, 152). Put differently: gender norms in various forms are always-already there. One can own them, creatively re-imagine them, reject them, or be inauthentically swept along by them; but one cannot simply avoid them. Gender norms demand our attention, and the question of how to interact with them is not always easy to answer. One may find oneself always-already moved by them in ways that are impossible to ignore. Rejecting or disowning certain gender norms may be the inauthentic choice; and given high rates of trans suicides, inauthenticity may not be a livable option. If a prisoner faces a choice between complicity in an unjust regime, or a life of suffering escapable only by death, the fact that the regime is morally bad does not soften the difficulty of the choice.

Put simply, understanding dominant gender norms as *bad* does not tell us how to navigate worlds where they have power over us. What, then, should we do?

A major point to make here is that this problem is not unique to the experiences of trans and GNC people. As Watson (2016) and others have noted, the burden of pushing back is often unfairly laid on trans and GNC shoulders, when the vast majority of cis and gender-conforming people are complicit in oppressive norms. What is distinctive about trans and GNC people is not the gendered practical dilemma we face, but rather the work we do to navigate it. As Bailey (2011) points out, the creative labor of constructing and articulating alternative gender practices is undertaken by trans and GNC communities out of dual necessity. First, we need to understand ourselves when dominant standards actively erase us. Second, we need to pass through dominant worlds without being assaulted or murdered. The latter is one reason why many gender practices in these communities do not fully escape the influence of dominant norms; as Bailey notes, such communities may "end up re-inscribing and relying upon those same norms to view and judge

each other within the community" (382). However, Bailey explicitly resists the interpretation of gender norms in ball culture as grounded in the mere internalization of harmful dominant gender practices. Rather, the "gender and sexual performativity of ballroom culture emerges and functions at the interstices of hegemony and transformation to create new forms of self-representation and social relations" (384). Trans and gender-nonconforming people of color in ballroom culture are strategically appropriating, re-imagining, and deploying the norms which are weaponized against them in order to create possibilities for agency, self-expression and authenticity—within their own worlds, but also when they are forced to move through dominant worlds.

I want to close, then, by suggesting two things. First, trans and GNC people are actively and materially aware that dominant gender norms are harmful to us and to others, even as we often must engage with them in our project of authenticity. Trans worlds of sense therefore have a rich history of working towards practical solutions to this problem, through the creative construction of alternative gender standards which provide opportunities for authenticity that move away from dominant norms. Second, as Marion (2011) and Bettcher (2014) both suggest, these alternative practices and norms have great potential as sites of resistance, not only within trans worlds of sense, but through their influence on dominant worlds. Rather than charge trans and GNC people with harming ourselves and others through our authentic engagement with gender norms, theorists should look to those alternative practices as models for how to engage in a project of authenticity in a non-ideal social world.

4.5 CONCLUSION: AUTHENTICITY AND LIBERATION

I have argued that the following positions are compatible: 1. Gender norms can be authentic, and as such can give us powerful normative reasons to act. 2. Gender norms are socially constructed, and many or most of them are morally bad, prudentially bad, or both. Social authenticity tells us that the authentic self is constantly under construction, embedded in a context out of which it must pursue a life project that is uniquely its own. But most social contexts are imperfect. They are full of unjust power relations, material and social positionalities, and morally and prudentially bad norms. We would certainly be better off if we never had to grapple with this in our project of self-construction. But we do. It is not possible to live authentically without attending to this flawed context. In an ideal world, there might be no conflict between authenticity and morality or prudence. However, the world is non-ideal, and we must build our authentic selves with the materials available.

I have relied throughout, without argument, on the understanding that gender norm authenticity is something worth explaining; that it is both real, and important to inclusive theorizing. For myself, I take the point to be self-evident. I experience certain gender norms as authentic for and authoritative over me, despite belief in their social construction and a theoretical and personal distaste for them. From conversation with other people—trans, GNC, cis, or gender-conforming—I infer that I am not alone in this, and that the practical dilemmas this generates are persistent and challenging. I conclude that making sense of felt gender norm authenticity is crucial to finding the hermeneutical tools to understand our experiences.

Moreover, I believe that this approach can work to ease some tensions in liberatory discourse. Authenticity claims about gender norms are often treated as evidence of either gender essentialism, or of the moral goodness of gender norms. This leads some trans and GNC people to embrace gender essentialism and reject feminist critiques of gender norms, while at the same

time encouraging some feminists to reject authenticity claims. I think both moves are a mistake. I have tried to show that authenticity does not entail either moral goodness or essentialism, in the hope that this will help move us towards a more inclusive theoretical space with respect to gender and gender norms—a central aim of this project on the whole.

BIBLIOGRAPHY

Adorno, Theodor (1973). *The Jargon of Authenticity*, trans. Knut Tarnowski and Frederic Will. Evanston: Northwestern University Press.

Alexander, Michelle (2010). *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New York: The New Press.

Antony, Louise (2020). "Feminism Without Metaphysics or a Deflationary Account of Gender." *Erkenntnis* 85: 529-549.

Ásta (2018). *Categories We Live By: The Construction of Sex, Gender, Race, and Other Social Categories*. New York: Oxford University Press.

Bach, Theodore (2012). "Gender Is a Natural Kind with a Historical Essence." *Ethics 122*(2), 231–272.

Bailey, Marlon (2011). "Gender/Racial Realness: Theorizing the Gender System in Ballroom Culture." *Feminist Studies* 37 (2): 365–86.

Barnes, Elizabeth (2019). "Gender and Gender Terms." *Nous.*

Bartky, Sandra (1990). *Femininity and Domination.* New York: Routledge.

Beauvoir, Simone (1949). *The Second Sex.* Trans. Constance Borde and Sheila Malovany Chevallier, 2009. New York: Vintage Books.

Bernstein, Sara (2019). "The Metaphysics of Intersectionality." *Philosophical Studies* 177 (2): 321-335.

Bettcher, Talia Mae (2013). "Trans Women and the Meaning of 'Woman'." In A. Soble, N. Power & R. Halwani (eds.), *Philosophy of Sex: Contemporary Readings, Sixth Edition*. Rowan & Littlefield: 233-250.

Bey, Marquis (2017). "The Trans*-Ness of Blackness, the Blackness of Trans*-Ness." *Transgender Studies Quarterly* 4 (2): 275-295.

Bey, Marquis (2022). *Black Trans Feminism*. Durham: Duke University Press.

Bindel, Julie (2007). "My Trans Mission." *The Guardian,* 1 August.

Bornstein, Kate. *Gender Outlaw*. New York: Seal Press.

Briggs & George (manuscript). "Science Fiction Double Feature: Trans Liberation on Twin Earth." https://philarchive.org/rec/GEOSFD

Butler, Judith (1990). *Gender Trouble*. New York: Routledge.

Butler, Judith (1993). *Bodies That Matter: On the Discursive Limits of 'Sex'*. New York: Routledge.

Carl, P. (2020a). Becoming a man. *The New York Times,* January 21.

Carl, P. (2020b). *Becoming a Man: The Story of a Transition*. New York: Simon & Schuster.

Carman, Taylor (2003) *Heidegger's Analytic: Interpretation, Discourse, and Authenticity in Being and Time*. New York: Cambridge University Press.

Coates, Jennifer. "I Am A Transwoman. I Am In The Closet. I Am Not Coming Out." *Medium,* 11 March 2016. https://medium.com/@jencoates/i-am-a-transwoman-i-am-in-the-closet-i-am-not-coming-out-4c2dd1907e42

Collins, Patricia Hill (2000). *Black Feminist Thought*. New York: Routledge.

Conover, Adam, host. (22 March 2022). Alice Evans, guest. "Why Do Men Rule the World?" In *Factually!* Earwolf podcast. https://www.earwolf.com/episode/why-do-men-rule-the-world-with-dr-alice-evans/

Crenshaw, Kimberlé. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Policies." *University of Chicago Legal Forum* 1989 (1): 139-167.

Culture." *Feminist Studies* 37 (2): 365–86.

Deliovsky, Kathy (2008). "Normative White Femininity." Atlantis 33(1): 49-59.

Dembroff, Robin (2018a). "Why Be Non-Binary?" *Aeon,* October 30. https://aeon.co/essays/nonbinary-identity-is-a-radical-stance-against-gender-segregation

Dembroff, Robin (2018b). "Real Talk on the Metaphysics of Gender." *Philosophical Topics* 46 (2): 21-50.

Dembroff, Robin (2020). "Beyond Binary: Genderqueer as Critical Gender Kind." Philosophers' Imprint 2o(9): 1-23.

Dembroff, Robin (forthcoming). *Real Men on Top: How Patriarchy Weaponizes Gender.* Book in progress. Summary: https://www.robindembroff.com/real-men-on-top.html

Durant, Will. *The Story of Philosophy.* 2006 (first published 1926). New York: Pocket Books.

Ecklund, Matti (2017). *Choosing Normative Concepts.* Oxford: Oxford University Press.

Espiritu, Yen Le (1997). "Race, Class, and Gender in Asian America." In *Making more waves,* ed. Elaine H. Kim, Lilia V. Villanueva, and Asian Women United of California. Boston: Beacon Press.

Evans, Alice (forthcoming). *The Great Gender Divergence.* Princeton: Princeton University Press.

Finlayson, Lorna; Katharine Jenkins; and Rosie Worsdale (2018). "I'm Not Transphobic, But…": A Feminist Case Against the Feminist Case Against Trans Inclusivity." Verso: 17 October.

Foucault, Michel (1982). "The Subject and Power." *Critical Inquiry* 8 (4): 777-795.

Frye, Marilyn (1983). *The Politics of Reality.* Berkeley: The Crossing Press (a division of Ten Speed Press).

Gossett, Che (2016). "Žižek's Trans/Gender Trouble." *LA Review of Books,* September 13. https://lareviewofbooks.org/article/zizeks-transgender-trouble/

Guignon, Charles (2004). *On Being Authentic.* London: Routledge.

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

Haslanger, Sally (2020). "Why I Don't Believe in Patriarchy: Comments on Kate Manne's *Down Girl*." *Philosophy and Phenomenological Research* 101 (1):220-229.

Haslanger, Sally, and Ásta (2018). "Feminist Metaphysics." *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.). https://plato.stanford.edu/archives/fall2018/entries/feminism-metaphysics/

Heidegger, Martin (1962). *Being and Time.* New York: Harper & Row.

Herron, Frank (2013). "It's a MUCH More Effective Quotation to Attribute It to Aristotle, Rather than to Will Durant." *The Art Of "Quotesmanship" and "Misquotesmanship":* UMass Amherst. https://blogs.umb.edu/quoteunquote/2012/05/08/its-a-much-more-effective-quotation-to-attribute-it-to-aristotle-rather-than-to-will-durant/

Hofstede, Geert (2001). *Culture's Consequences*. New York: Sage University Press.

hooks, bell (1982). *Ain't I a Woman: Black Women and Feminism.* Boston: South End Press.

hooks, bell (2000). *Feminism is for Everybody.* Cambridge: South End Press.

hooks, bell; Marci Blackman; Shola Lynch; and Janet Mock, panelists (2014). *Panel discussion.* "Are You Still a Slave? Liberating the Black Female Body." 6 May, Alvin Johnson/J.M. Kaplan Hall, The New School, 66 West 12th Street, New York, NY. Retrieved from https://www.youtube.com/watch?v=rJk0hNROvzs

Huntress, Caelan (2017). "My Favorite Quote Of All Time Is A Misattribution." *Medium,* August 24. https://medium.com/the-mission/my-favourite-quote-of-all-time-is-a-misattribution-66356f22843d

Incoherent Queer Screaming (2020). "Gender according to the Cis, based on their cakes." Youtube, 14 June. https://www.youtube.com/watch?v=sP9Z8Xq4E24

Janack, Marianne. "Standpoint Epistemology without the 'Standpoint'?: An Examination of Epistemic Privilege and Epistemic Authority." *Hypatia* 12, no. 2 (1997): 125–39. http://www.jstor.org/stable/3810473.

Jeffries, Stuart (2002). "The Quest for Truth." *The Guardian*, online edition. November 30.

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

Jenkins, Katharine (2018). "Towards an Account of Gender Identity." *Ergo* 5.

Keeley, Benjamin (1973). "Review of *Internalization of Norms: A Sociological Theory of Moral Commitment*, by John Finley Scott." *Sociological Analysis* 34(3): 236–38.

Korsgaard, Christine (1996). *The Sources of Normativity.* Cambridge, Cambridge University Press.

Korsgaard, Christine (2009). *Self-Constitution: Agency, Identity, and Integrity.* Oxford: Oxford University Press.

Lawford-Smith, Holly (forthcoming). *Gender Critical Feminism.* Oxford: Oxford University Press.

Longino, Helen (1989). "Feminist critiques of rationality: Critiques of science or philosophy of science?" *Women's Studies International Forum* 12(3): 261-269.

Lopes, Filipa Melo (2019). *Recognizing Social Subjects: Gender, Disability and Social Standing.* Dissertation. https://deepblue.lib.umich.edu/handle/2027.42/151639

Lugones, María (1987). "Playfulness, "World"-Travelling, and Loving Perception." *Hypatia* 2 (2): 3-19.

Lugones, María (2007). "Heterosexualism and the Colonial/Modern Gender System." *Hypatia* 22 (9): 186-209.

Lugones, María (2010). "Towards a Decolonial Feminism." *Hypatia* 25 (4): 742-759.

MacKinnon, Catherine (1989). *Towards a Feminist Theory of the State.* Cambridge: Harvard University Press.

Manne, Kate (2017). *Down Girl.* New York: Oxford University Press.

Mayo Clinic Staff (2017). "Children and gender identity: Supporting your child." September 1. Retrieved from https://www.mayoclinic.org/healthy-lifestyle/childrens-health/in-depth/children-and-gender-identity/art-20266811

Mikkola, Mari (2011). "Ontological Commitments, Sex and Gender." In *Feminist Metaphysics: Explorations in the Ontology of Sex, Gender, and the Self.* Charlotte Witt, ed. Netherlands: Springer.

Mikkola, Mari (2016). *The Wrong of Injustice.* New York: Oxford University Press.

Mill, John Stuart (1997 [1865]). *On Liberty and the Subjection of Women.* Hertfordshire: Wordsworth Editions Limited.

Mills, Charles (1997). *The Racial Contract.* Ithaca: Cornell University Press.

Mock, Janet (2014). *Redefining Realness: My Path to Womanhood, Identity, Love & So Much More.* New York: Atria (Simon & Schuster).

Ortega, Mariana (2005). When Conscience Calls, Will Dasein Answer? Heideggerian Authenticity and the Possibility of Ethical Life." *International Journal of Philosophical Studies* 13 (1): 15 – 34.

Paakkunainen, Hille (2018). "Normativity and Agency." In *The Routledge Handbook of Metaethics,* eds. Tristram McPherson and David Plunkett. New York: Routledge.

Page, Elliot (2020). Twitter post, December 1, 12:10 pm. https://twitter.com/TheElliotPage/.

Pritchett, Clara (2019). "The Ungendering of Black Women." AFRI 0090 S01: An Introduction to Africana Studies. Published November 5. https://blogs.brown.edu/afri-0090-s01-2019-fall/2019/11/05/the-ungendering-of-black-women/

Prosser, Jay (1998). *Second Skins: The Body Narratives of Transsexuality.* New York: Columbia University Press.

Raymond, Janice (1979). *The Transsexual Empire: The Making of the She-Male.* New York: Beacon Press.

Rorty, Richard (1989). *Contingency, Irony and Solidarity*. Cambridge: Cambridge University Press.

Rousseau, Jean-Jacques, & Cohen, J. M. (1953 [1782]). *The Confessions of Jean-Jacques Rousseau.* London: Penguin Books.

Salamon, Gayle (2010). *Assuming a Body.* New York: Columbia University Press.

Scott, John Finley (1971). *Internalization of Norms: A Sociological Theory of Moral Commitment.* Englewood Cliffs: Prentice-Hall, Inc.

Serano, Julia (2007). *Whipping Girl.* Berkeley, CA: Seal Press.

Spillers, Hortense (1987). "Mama's Baby, Papa's Maybe: An American Grammar Book." *Diacritics* 17 (2): 65–81.

Sripada, Chandra Sekhar, & Steven Stich (2007). "A Framework for the Psychology of Norms." In *The Innate Mind: Volume 2: Culture and Cognition.* Peter Carruthers, Stephen Laurence, and Stephen Stich, eds. New York: Oxford University Press. 280-301.

Steinmetz, Kathy (2021). "Elliot Page is ready for this moment." *TIME*, March 16.

Stone, Sandy (1987). "The *Empire* Strikes Back: A Posttranssexual Manifesto." https://sandystone.com/empire-strikes-back.pdf

Street, Sharon (2010). What Is Constructivism in Ethics and Metaethics?" *Philosophy Compass* 5: 363-384.

Taylor, Charles (1991). *The Ethics of Authenticity.* Cambridge: Harvard University Press.

Thorn, Abigail (2021). "Food, Beauty, Mind." *Philosophy Tube* on YouTube. Uploaded August 28. https://www.youtube.com/watch?v=9MGUWtTakQc

Vaid-Menon, Alok (2020). *Beyond the Gender Binary.* New York: Penguin.

Varga, Somogy and Charles Guignon (2020). "Authenticity." *The Stanford Encyclopedia of Philosophy (Spring 2020 Edition)*, ed. Edward N. Zalta.

Watson, Lori (2016). "The Woman Question." *Transgender Studies Quarterly* 3 (1–2): 246–53.

Wilchins, Riki Anne (1997). *Read My Lips: Sexual Subversion and the End of Gender.* Ithaca: Firebrand Books.

Witt, Charlotte (2011). *The Metaphysics of Gender.* Oxford: Oxford University Press.

Witt, Charlotte (2020). "What Explains Social Role Normativity?" In *Social Functions in Philosophy: Metaphysical, Normative, and Methodological Perspectives.* Rebekka Hufendiek, Daniel James and Raphael van Riel, eds. New York: Routledge.

Wodak, Daniel (2019). "Mere Formalities: Fictional Normativity and Normative Authority. *Canadian Journal of Philosophy, 49*(6), 828-850.

Young, Iris Marion (1980). "Throwing like a Girl: A Phenomenology of Feminine Body Comportment Motility and Spatiality." *Human Studies* 3 (2): 137-56.

Young, Iris Marion (1990). "Five Faces of Oppression." In *Justice and the Politics of Difference* (Iris Marion Young, author). Princeton: Princeton University Press. 39-65.

Zazanis, Noah (2019). "On Hating Men (And Becoming One Anyway): Transmasculinity, Feminism, and the Politics of Online." *The New Inquiry,* December 24. https://thenewinquiry.com/on-hating-men-and-becoming-one-anyway/

Zheng, Robin (2016). "Why Yellow Fever Isn't Flattering: A Case Against Racial Fetishes." *Journal of the American Philosophical Association* 2 (3): 400-419.

ROWAN BELL

CURRICULUM VITAE

DEPARTMENT OF PHILOSOPHY
SYRACUSE UNIVERSITY
541 HALL OF LANGUAGES, SYRACUSE, NY, 13244

(423) 302-8467
HTTPS://ROWANBELL.ME
THE.REAL.ROWAN.BELL@GMAIL.COM

ACADEMIC APPOINTMENTS

2022-2023   **Postdoctoral Fellow**, Preparing Future Faculty: Faculty Diversity Program & The Department of Philosophy, University of Missouri (Columbia, Missouri)

2023 -   **Assistant Professor,** Philosophy & Sexualities, Genders, and Social Change, University of Guelph (Guelph, Ontario)

AREAS OF EXPERTISE

Areas of Specialization: feminist philosophy, trans philosophy, metaethics

Areas of Competence: social epistemology, social ontology, applied ethics, philosophy of race

EDUCATION

2022   **Ph.D.,** Philosophy (expected)
Department of Philosophy, Syracuse University
Dissertation Title: "Gender Norms and Gendered Traits"

2014   **B.S.,** Psychology and Philosophy: Magna Cum Laude
East Tennessee State University (ETSU)

RESEARCH

Forthcoming   "Being Your Best Self: Authenticity, Morality, and Gender Norms." *Hypatia.*

Under review   "'It's Just Science': Thick Concepts and Hermeneutical Injustice."

Draft   "Distorted Identities: Hermeneutical Injustice and Normative Social Roles."

Draft   "The Role of a Lifetime: Trans Experience and Gender Norms"

Published   "(Un)Radical Feminism: Gender and the Limits of Imagination." In *To Boldly Stay: Essays on the Series that "Went" Where No Trek Had Gone Before*. Sherry Ginn and Michael Cornelius, eds. McFarland Books.

SELECTED AWARDS AND FELLOWSHIPS

FINANCIAL AWARDS IN PARENTHESES

Humanities Center Dissertation Fellowship (funding for one academic year)
    The Center for the Humanities, Syracuse University: 2022-23. *Offered and declined.*

Summer Dissertation Fellowship ($4000)
    The Graduate School, Syracuse University: Summer 2021

Graduate Dean's Award for Excellence in Research and Creative Work ($500)
    The Graduate School, Syracuse University: Spring 2021

SWIP-Analytic Annual Graduate Student Essay Prize ($300)
    Society for Women in Philosophy (SWIP) and the CUNY Graduate Center: Spring 2020

Outstanding Teaching Assistant Award
    The Graduate School, Syracuse University: Spring 2020 (standing title)

Faculty Choice: Outstanding Student Award
    Department of Philosophy & Humanities, ETSU, Graduating Class 2014

SELECTED RESEARCH PRESENTATIONS

* DENOTES COMPETITIVE SELECTION; † DENOTES INVITED TALK

"The Role of a Lifetime: Trans Experience and Gender Norms." Accepted for presentation at:

- * Social Ontology and Collective Intentionality 2022, Vienna, Austria, August 2022.

- * North American Society for Social Philosophy International Conference: Polarization, Reconciliation, and Community, Aston, PA, July 2022.

- * APA Pacific Meeting, Vancouver, BC, April 2022.

* "The Trans Authenticity Paradox." Presented at the LGBTQ+ Advocacy Committee Session, Society for Phenomenology and Existential Philosophy (SPEP) Virtual Meeting, September 2021

"It's Just Science: Thick Concepts and Hermeneutical Injustice." Accepted for presentation at:

- * APA Central Meeting, February 2021. (Held virtually)

- * The Association for Feminist Epistemologies, Methodologies, Metaphysics, and Science Studies 8. (Postponed indefinitely due to COVID-19)

* "Distorted Identities: Hermeneutical Injustice and Normative Social Roles." Presented at the APA Eastern Meeting, January 2021. (Held virtually)

* "Authenticity and Gender Norms." Presented at MAP Flash Talks, May 2020. (Virtual event)

† "Transgressing Normativity: The Metaphysics of Gender-Nonconformity." Presented at Trans Philosophy Spring Workshop Series, Queens University, Kingston, ON, February 2020.

"Inside and Out: Ascriptivism about Gendered Traits." Presented at:

- † EQS: Early Queer Scholarship Series, Syracuse University, Syracuse, NY, October 2019.

- * MAPWorks (Minorities And Philosophy Spring Workshop Series), Columbia University, New York, NY, April 2019.
- * Liberating the Future: Concordia GSPA Spring 2019 Conference, Concordia University, Montreal, QC, March 2019.
- * SWIPShop (Society for Women in Philosophy Workshop), Baruch College, New York, NY, February 2019.

* "Gender and Traits." Presented at Social Ontology, Tufts University, Boston, MA, August 2018.

Bell, R.A., Taylor, D.A., White, C.D., Shi, D., Martin, B.A., & Dula, C.S. "Relationship between Grades, Academic Self-Efficacy, Family Conflict and Church Attendance." Presented at Appalachian Student Research Forum, Johnson City, TN, April 2010. **First Place Winner**

Taylor, D.A., Bell, R.A., Martin, B.A., & Dula, C.S. "We Wouldn't Call It 'Fighting': The differences in how parents and students view conflict and the effects on grades." Presented at Tennessee Psychological Association, Nashville, TN, November 2009. **Third Place Winner**

Taylor, D.A., Martin, B.A., Gibson, B.W., Bell, R.A., & Dula, C.S. "Effects of Home Environment Type on Family Conflict and Extracurricular Activities." Presented at Mid-South Psychology Conference, Memphis, TN, February 2010. **First Place Winner**

TEACHING EXPERIENCE

Teaching Mentor, University-Wide TA Orientation, Syracuse University Graduate School, 2019-present. Competitive position.

Primary Instructor, Department of Philosophy, Syracuse University

Spring	2022	Social & Political Philosophy
Fall	2021	Philosophy of Feminism
Spring	2021	Logic
Fall	2020	Philosophy of Feminism (hybrid)
Summer	2020	Philosophy of Feminism (online asynchronous)
Spring	2020	Human Nature
Fall	2019	Philosophy of Feminism
Summer	2019	Philosophy of Feminism (online asynchronous)

Teaching Assistant, Department of Philosophy, Syracuse University

Spring	2019	Ethics and the Media Professions	Paul Prescott
Fall	2018	Introduction to Logic	Mark Heller
Spring	2018	Introduction to Logic	Michael Rieppel
Fall	2017	Philosophy of Feminism	Kara Richardson
Spring	2017	Human Nature	Neelam Sethi
Fall	2016	Theories of Knowledge and Reality	Robert van Gulick

GRANTS AWARDED

$2500. Engaged Humanities Grant, Syracuse University. For Syracuse University's *Philosophy Lab,* a branch of the [Philosophy Learning and Teaching Organization (PLATO)](#).

COMMENTARIES

Comments on "Is Gender Fictionalism Defensible and if not, How Should We Think About Gender?" by Lucia Dikaczova. Syracuse University Philosophy Graduate Conference, Syracuse, NY, March 2022.

Comments on "The Role of Survival in Constructivism," by Stacy Kohls. Syracuse University ABD Workshop Series, Syracuse, NY, March 2022.

Comments on "No Desires are Relevant to Well-Being," by Nikki Fortier. Syracuse University ABD Workshop Series, Syracuse, NY, November 2021.

Comments on "Two Conceptions of Essences," by Laura Nicoera. Syracuse University Philosophy Graduate Conference, Syracuse, NY, April 2021. (Conference held virtually.)

Comments on "A Sensibility of Humour," by Zoe Walker. Syracuse University Philosophy Graduate Conference, Syracuse, NY, August 2020. (Conference held virtually.)

Comments on "The Wrongful Inclusion Problem and Jenkins's Analysis of Gender Concepts," by Evan Woods. Syracuse University Philosophy Graduate Conference, Syracuse, NY, Mar 2017.

PROFESSIONAL SERVICE AND MEMBERSHIPS
* DENOTES PAID POSITION

Invited Panelist, MAP PhD Panel, Mudd Undergraduate Ethics Conference, March 2022

*President (elected), SU Philosophy Graduate Student Organization, 2020-2021

*Founder and Organizer, MAP Anti-Racist Reading Group, Summer-Fall 2020

Co-Organizer, MANCEPT Workshops, "What Is Gender?" series, Fall 2021 and 2022

Founder and Co-Organizer, Syracuse-Queens Social Metaphysics Group, 2020-2021

Peer Evaluator, Teaching Mentor Selection Committee, Spring 2020 & 2021

*Organizer, SU Philosophy Women & Gender Minorities Group, 2018-2020

*Co-Organizer, SU Philosophy Graduate Conference, 2017-2018

Activities Coordinator, SU Philosophy Graduate Student Organization, 2017-2020

Reviewer, *Feminist Philosophy Quarterly*

Reviewer, *Hypatia*

Reviewer, *Ergo*

Member, American Philosophical Association

Member, Society for Phenomenology and Existential Philosophy

Member, International Social Ontology Society

REFERENCES

Dr. Hille Paakkunainen
Associate Professor, Department of Philosophy, Syracuse University
hpaakkun@syr.edu

Dr. Katharine Jenkins
Lecturer, Department of Philosophy, University of Glasgow
Katharine.Jenkins@glasgow.ac.uk

Dr. Kara Richardson
Associate Professor and Department Chair, Department of Philosophy, Syracuse University
kricha03@syr.edu

Dr. Verena Erlenbusch-Anderson
Associate Professor, Department of Philosophy, Syracuse University
verlenbu@syr.edu

Dr. Luvell Anderson
Associate Professor, Department of Philosophy, Syracuse University
lander04@syr.edu