HOW TO BE SOCIAL

By

DANIELLE MARIE PAYTON

A dissertation submitted to the School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Philosophy

Written under the direction of

Theodore Sider

And approved by

_________________________________

_________________________________

________________________

________________________

_________________________________

New Brunswick, NJ

May 2021
ABSTRACT OF THE DISSERTATION

How To Be Social

By

DANIELLE MARIE PAYTON

Dissertation Director:  
Theodore Sider

Many people think there are socially constructed properties. For example, quite plausibly, gender properties like the property of being a woman are socially constructed. But what is it for a property like this to be socially constructed? What is it for a property to be a social property in the first place? And, where we are primarily concerned with investigating the nature of gender, why ask about properties at all? Why not, for example, simply ask about the meanings of our gender terms, like ‘woman’? How To Be Social offers answers to each of these questions.
DEDICATION

This dissertation is dedicated to the members my committee: Karen Bennett, Liz Camp, Kate Ritchie, Jonathan Schaffer, and especially Ted Sider, whose generosity, guidance, and support has been invaluable to me throughout graduate school.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER ONE: Real Definitions for Real Social Constructs</td>
<td>3</td>
</tr>
<tr>
<td>CHAPTER TWO: Constructing Properties with Words</td>
<td>26</td>
</tr>
<tr>
<td>CHAPTER THREE: Gender Metaphysics</td>
<td>42</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>66</td>
</tr>
</tbody>
</table>
INTRODUCTION

The claim that gender is socially constructed is plausibly the most central claim operative in feminist projects both within and outside of the academy. There are many ways to interpret this claim: some people think it means that gender is “all in our heads”, and because of this, gender is not “real” in a substantive sense. Call this position gender anti-realism. In contrast with gender anti-realism, many gender realists think that gender is real in a substantive sense. There are a number of ways to be a gender realist, but on many of these views gender is real at least in the sense that there are gender properties. Combine this with the claim that gender is socially constructed, and we get the view that there are real, socially constructed gender properties.

Which properties are the gender properties? What is it for a property to be socially constructed? What makes a property social in the first place? How To Be Social, addresses those questions.

In the first chapter, “Real Definitions for Real Social Constructs”, I develop a view of the social construction of properties. On my proposed view, a property is socially constructed just in case the real definition (“essence”) of that property includes social things. This raises a question: which things are the social things? In the second chapter, “Constructing Properties with Words”, I use the view developed in the first chapter to propose a sufficient condition on the property expressed by our predicate ‘social’. So, the first chapter offers an analysis of social construction and the second chapter offers some insight into sociality. The third chapter, “Gender Metaphysics”, draws on insights about the nature of socially constructed properties delivered by the previous two chapters, and applies them to discussions about gender. There I argue that questions like ‘what is it to be a woman?’ do not have to be treated as questions about the meaning of ‘woman’; they can also be treated as questions about the nature of the properties that predicates like ‘woman’ express.
While I am personally motivated by feminist concerns more broadly, *How To Be Social* is primarily a project in social metaphysics. In this dissertation I am largely concerned with developing and motivating metaphysical tools that can be put to work in service of feminist projects. To this end, the views I develop are assembled with the help of concepts routinely employed in “mainstream” metaphysics (e.g., essence, grounding, metaphysical explanation), and I apply these views to discussions in feminist philosophy regarding the social construction of properties and social kinds.
CHAPTER ONE

Real Definitions for Real Social Constructs

Abstract. Many people think there are socially constructed properties. For example, quite plausibly, the property of being a woman is a socially constructed property. In this paper I defend the view that a property is socially constructed just in case the real definition (or essence) of that property makes reference to social factors. I argue that this view has advantages over the view that socially constructed properties are constituted by social factors, as well as the view that socially constructed properties are grounded in social factors.

§1 Introduction

This is an essay about social construction. It is common to think that certain properties—like the property of being a woman, or the property of being Black—are socially constructed. In this paper I am not going to argue for the view that certain properties are socially constructed. Instead, I will develop and motivate one view of how some properties might be constructed, given that they are. That is, I’m interested here in developing a metaphysical tool, one which I will argue proves useful to social ontologists and others interested in the social construction of reality.

§2 Relations of Social Construction

Generally speaking, the literature on social construction respects a distinction between socially constructed representations (e.g., ideas, concepts, predicates), and socially constructed entities (e.g., individuals, categories, events, properties). In this paper, I’ll be concerned exclusively with the social construction of properties. Unsurprisingly,
philosophers disagree about how entities (like properties) are socially constructed. But there is some consensus, and I’ll highlight that here.

To start, it is common to claim that socially constructed entities exist only contingently, in the sense that the fact that they exist at all (at our world) is contingent upon the social practices and relations (perhaps among other things) in play here. Relatedly, the claim that \( x \) constructs \( y \) often conveys that a change in \( x \) will produce a change in what \( x \) constructs, with the consequence that \( y \) itself is slightly different, or else that \( y \) ceases to exist altogether, and something “new” exists in its stead (Hacking 1999). Furthermore, in many (but notably not all!) cases, the relation that holds between constructing entities and what they construct is taken to be asymmetric: so, \( x \) constructs \( y \), but not the other way around. This is in part because the relation that holds between \( x \) and \( y \) is also taken to be explanatory: the constructing entities play some role in explaining what they construct. Finally, this relation is also taken to be one of both generation and dependence: \( x \) is at least part of what makes it the case that \( y \), and for a similar reason, \( y \) at least partly depends upon \( x \) (Ásta 2015; Mallon 2007). This list can be distilled into four claims, each of which should follow from the claim that \( x \) socially constructs \( y \):

**CONTINGENCY:** \( y \) is contingent.

**EXPLANATION:** \( x \) at least partially explains \( y \).
GENERATION: where \( x \) and \( y \) are objects or properties, \( x \) is at least part of what makes \( y \) the entity that it is; or, where \( x \) and \( y \) are facts, \( x \) is at least part of what makes it the case that \( y \).

DEPENDENCE: where \( x \) and \( y \) are objects or properties, \( y \) at least partly depends upon \( x \), such that (necessarily) if \( x \) doesn’t exist then \( y \) doesn’t exist; or, where \( x \) and \( y \) are facts, \( y \) at least partly depends upon \( x \), such that (necessarily) if \( \neg y \) then \( \neg x \).

In addition to these commitments, it is also common to distinguish between entities which are causally socially constructed, and those which are constitutively socially constructed. Here, I’ll follow Haslanger (2012) in understanding this distinction as follows:

- **Causal Construction:** Something is causally constructed iff social factors play a causal role in bringing it into existence, or to some substantial extent, in its being the way it is.

- **Constitutive Construction:** Something is constitutively constructed iff in defining it we must make reference to social factors. (Haslanger 2012, p. 87)

Causal social construction has fallen out of favor in recent years, in part because it does not satisfy all of the desiderata listed above.\(^1\) As such, this leaves all of the heavy lifting (at least with respect to the social construction of entities) to constitutive construction. And it really has been put to work: for example, it has been argued that gender and race properties are

---

\(^1\) Esa Díaz-León (2015) offers a full, paper-length critique of causal social construction (which doubles as an argument in favor of constitutive social construction).
constitutively constructed (Haslanger 2000, 2012). Disability might be constitutively constructed, as well as sex (Ásta 2012, 2015, 2018). But much more generally, it has been argued that all social facts, social properties and kinds, and a huge range of other social entities are constitutively constructed (Ásta 2018; Boghossian 2006; Epstein 2015; Haslanger 2012; Mallon 2008, 2015).

Given its central role in these discussions, it is important to have a clear articulation of constitutive construction on hand. It is common to understand Haslanger’s original definition in terms of constitution, an interpretation of constitutive construction which I’ll here call *constitutivism*.

*Constitutivism* Something is constitutively constructed just in case it is constituted by social factors.

For example, in a recent *Philosophy Compass* article, Ásta writes,

Intuitively, X is constitutively constructed just in case social phenomena make up or constitute X. The notion of constitution is the familiar one: just as a material object is constituted by the materials that make it up, arranged in a certain way, so a social phenomenon is constituted by the phenomena (physical and social) that make it up, again, arranged in a certain way. (Ásta 2015, p. 5)

I think there are a number of problems facing constitutivism, many of which stem from the fact that it is hard to see how things like socially constructed properties (e.g., the property of
being a woman) might be materially constituted in the classical statue-and-lump sense. I'll discuss this worry in the next section.

However, there is reason to think that Haslanger did not have constitutivism in mind when she first articulated these definitions. In a footnote immediately following the introduction of this definition and in more recent work, Haslanger clarifies that she intends constitutive construction to apply directly to the real definitions of entities, not just to those terms or concepts which refer to them:

The project is one of ‘real definition’ rather than ‘nominal definition’ [Rosen 2013] . . . My core-dependent accounts of gender and race are attempts at real definition not conceptual analysis in the traditional sense.

(Haslanger 2014, p. 29)

I maintain that my analysis of gender is descriptive, in fact, descriptive of the material reality of gender, even if it is not descriptive of anyone’s thinking about gender. In short, I am offering a theory of what gender is, not what some people think it is, or what they think it should be. This is the goal of real definition. (Haslanger 2014, p. 31)

These passages suggest that Haslanger’s original characterization of constitutive construction can be rephrased as follows:

Constitutive Construction (real definition reading) Something is constitutively constructed just in case its real definition makes reference to social factors.
This essay takes the real definition reading of constitutive construction seriously, and as such, endeavors to answer the following two questions: first, how does the metaphysics behind this real definition approach to constitutive construction work? Answering this question will involve saying a bit more about what it is for a real definition to “reference” social factors. And second, how does incorporating talk of real definitions into our theorizing about socially constructed properties help? Answering this question will involve saying a bit more about the goals of projects in social ontology.

In this paper I will do three things: first, I will argue that understanding constitutive social construction in terms of material constitution is problematic. Then I will develop an analysis of constitutive social construction in terms of real definitions and discuss how it succeeds where the constitution approach fails. Finally, I will consider one other popular way of thinking about social construction—social construction as grounding—and argue that the real definition approach to constitutive construction has advantages over that approach, as well.

§3 Three Quick Preliminaries

In this paper I am exclusively concerned with socially constructed properties—properties like the property of being a woman or the property of being Black. And I will for the most part speak roughly in terms of properties which are constructed by patterns of behavior (patterns which plausibly include the actions of individuals, their mental states, and the (social) environments, resources, and institutions with which they interact) (Haslanger 2007, 2012). This is of course a gross over-simplification, but I do this mainly so that I can phrase my discussion in more concrete terms. Additionally, throughout this essay, I will
often speak in terms of the property of being a woman in my examples. With this phrasing I do not mean to suggest that there is one single unique property of being a woman. This too, is a way of speaking that I’ve adopted in order to streamline discussion in this essay.

§4 Social Construction as Material Constitution

Constitutivism is the view that a property is socially constructed iff it is constituted by social factors. How should we understand the notion of constitution employed by constitutivism? The passage from Ásta (2015) quoted earlier suggests that we would do well to take this view rather literally, and understand the notion of constitution employed here to be that of material constitution.

Material constitution is commonly understood as a relation which holds between distinct objects that share all of their material parts, and so are located in the same place at the same time. More concisely then: material constitution is a relation which holds between spatially, materially, and temporally coincident things. Most proponents of material constitution hold that constitution is an asymmetric relation (one thing is constituted, the other does the constituting, and not vice versa) and it is not a relation of identity (the constitutor is not identical to the constituted). Additionally, a smaller subset holds that constitution is a dependence relation (the constituted entity depends in some sense upon its constitutor).² Material constitution has been championed as a solution to many tough puzzles in metaphysics, perhaps most famously, the puzzle of the statue and the clay—

Briefly, the puzzle is that the statue and the clay out of which it is made have different properties (e.g., the clay came into existence before the statue, the clay can survive changes in shape which the statue cannot, the statue and the clay seem to differ in their kind membership, . . .), and so they cannot be identical. And yet, the statue and the clay share all of their material parts and are perfectly co-located with one another. Typically, it is impossible for two material objects to be both spatially and materially coincident, and yet remain distinct. And so we can ask: do we have two objects here, or just one? Many proponents of material constitution take the first option: we have two objects, the statue and the clay, and the clay constitutes the statue. This view has been met with a number of difficult objections which have sparked several interesting and extended debates that I won’t be able to engage with here. Suffice it to say that, as with any substantive metaphysical theory, material constitution has its own battles and its own capable defenders.³

But on the assumption that material constitution is a viable view in its own right, the question becomes the following: can we understand the notion of constitution employed by constitutivism in terms of material constitution? I don’t think so, for two reasons: first, material constitution is typically taken to be a one-to-one relation, a relation which holds between two relata. However, the majority consensus is that socially constructed properties are constructed out of a great many things (e.g., behaviors, behavioral patterns, mental states, resources) (Ásta 2018, 2013; Haslanger 2016, 2018; Mallon 2003, 2016). That is, the relation of social construction is taken to be a many-one relation.

³ See references in fn. 2
Second, I think it’s doubtful that socially constructed properties and the behavioral patterns of individuals are spatially coincident. I take it that two things are spatially coincident at a time \( t \) just in case they occupy exactly the same region of space at \( t \). Behaviors are instantiated by individuals at a time and a specific place, and so for the sake of argument, let’s say that behaviors have locations. And, insofar as a pattern of behavior is made up of individual instances a given (sort of) behavior, it too might have a location (although the region at which it’s located is likely to be extremely disjointed). There are then two ways for socially constructed properties to fail to be spatially coincident with the patterns of behavior which construct them: (1) socially constructed properties might not have any spatial location at all, or (2) they might have some spatial location which is not shared by the patterns of behavior which constitute them.

Possibly, properties have locations, so let’s grant that here (Armstrong 1978). But even given that properties have locations, it’s likely that they are not co-located with the patterns of behavior which construct them. Take the property of being a woman for example. If you think that property has a location, then you probably think it’s located where the women are, and many (if not all) women engage in the patterns of behavior which construct the property of being a woman. Accordingly, you might think that the property of being a woman and the patterns of behavior which construct it are spatially coincident at those regions.

But intuitively, the property of being a woman is not socially constructed by all and only the behaviors of individual women. It is constructed by the behaviors of people of all genders (as well as by institutions, norms, public policy . . . ) (Ásta 2018; Haslanger 2012). But if the property of being a woman is socially constructed by the behaviors of people of all genders,
then there are some patterns of behavior which (a) construct the property of being a woman, but (b) are not spatially coincident with that property. And if that’s right, then the property of being a woman is not materially constituted by patterns of behavior. Perhaps there are other ways that we could conceive of properties as having a location, but prima facie it seems that if the property of being a woman is located, it is not spatially coincident with the patterns of behavior which construct it.

The objections that I’ve considered here may not be knock-down arguments against constitutivism, but they are troubling. As such, I hope that discussion in this section has accomplished the following two things: first, that these paragraphs served to cast doubt on those interpretations which understand the notion of constitution employed by constitutivism in terms of material constitution. Second, and more importantly, I hope to have shown that the metaphysics behind the notion of constitution invoked by constitutivism is more slippery and elusive than it may have appeared on a first pass. And this, I think, is enough to motivate a search for alternatives.

§5 Social Construction and Real Definitions

According to the real definition reading of constitutive social construction, a property is constitutively socially constructed just in case its real definition makes reference to social factors. Does this way of thinking about the social construction of properties fare better than the constitution approach just considered? In order to answer that question, we’ll have to answer two further questions: first, how does the metaphysics behind constitutive construction work? And second, how does incorporating talk of real definitions into our
theorizing about socially constructed properties help? In this section I’ll address the first question and in the next section I’ll address the second question.

Saying more about the metaphysics behind constitutive construction will involve saying more about what it is for a real definition to “reference” social factors. I’ll start by assuming that the real definition of a property $F$ “makes reference” to a given social factor $S$ just in case $S$ is “in” the essence of $F$. In the literature on essence, this relationship is sometimes glossed in terms of “what it is to be” $F$. So in this case, we would say that part of what it is to be $F$ is to be $S$.

Real definitions are importantly different from so-called nominal definitions: for example, when we ask about the nominal definition for something like virtue, we will probably want to know what the term ‘virtue’ means, or how people use that term. In contrast, when we ask for the real definition of virtue, we want to know about the nature of the thing the term ‘virtue’ denotes— we want to know about the essence of virtue. Roughly put, nominal definitions are definitions of linguistic items (e.g., ‘woman’), and real definitions are statements of the essence of non-linguistic entities, like properties (e.g., the property of being a woman).

\footnote{For the rest of this paper I’m going to talk about social factors in terms of properties (e.g., the property of being a certain pattern of behavior is a social factor). I’m doing this in order to streamline the discussion to follow, particularly where it comes to evaluating the analysis of constitutive construction proposed here using the four principles listed earlier (§2). Strictly speaking, we don’t need to think about social factors in terms of properties however, and readers are welcome to bring their own account of social factors to this discussion and modify the conditions below accordingly.}

\footnote{Here I’m working with Kit Fine’s (1994) notion of constitutive essence.}
Bringing this back around to constitutive construction, we now have something like this:

\[
\begin{align*}
\textit{Constitutive Construction} & \quad \text{A property } F \text{ is constitutively constructed iff} \\
\textit{(real definition reading)} & \quad \text{there is at least one social factor in the essence of } F.
\end{align*}
\]

What does it take for a social factor to be in the essence of a constitutively constructed property? There are probably several good answers to this question. Here, I'll propose just one way of answering it, one which follows Rosen (2015). Start with the idea that for any given property, there are conditions which are both necessary and sufficient for possessing that property. For example, something is a triangle iff it is a closed, three-sided figure. And quite plausibly, this has to be true— for example, there are no possible worlds where it is true that something is a triangle iff it is a closed, four-sided figure. If that’s right then we might say that necessarily, something is a triangle iff it is a closed, three-sided figure.

Now, maybe you think it’s natural to say that what it is to be a triangle is to be a closed, three-sided figure. Rosen’s view works from that intuition: namely, that there is something about what it is to be a triangle which makes it the case that necessarily, \( x \) is a triangle iff \( x \) is a closed, three-sided figure. And further, there is something about the essence of trianglehood which makes it the case that if \( x \) is a triangle, then that is because \( x \) is a closed, three-sided figure. We might say that the fact that \( x \) is a triangle is grounded in the fact that \( x \) is a closed, three-sided figure and that grounding relationship is itself explained by the
essence of the property of being a triangle. Rosen calls this the *Grounding Definition Link* (Rosen 2015, p. 190).^6

In the example above, the property of being a closed, three-sided figure does a lot of work. So much work in fact, that it might be fair to say that the property of being a triangle is defined in terms of the property of being a closed, three-sided figure. In fact, on Rosen’s view, where a property \( F \) is partially defined in terms of another property \( G \) in this way, we can say that \( G \) is in the real definition (essence) of \( F \). Just so, we might think, if a constitutively constructed property \( F \) is similarly partially defined in terms of a social factor, then that social factor is in the real definition of \( F \). This would give us an answer to our earlier question regarding what it takes for a social factor to be in the real definition of a constitutively constructed property, and it would also give us some insight into the role that social factors play in constitutively constructing properties.

Combining Rosen’s view with the statement of constitutive construction immediately above yields the following:

---

^6 As I will use it here, grounding is a relation of metaphysical explanation which obtains between facts. When one fact, \( x \), grounds another fact, \( y \), the following things are true: necessarily, if \( x \) is true then \( y \) is true; \( x \) is more metaphysically fundamental that \( y \); \( y \) is true because \( x \) is true (but not vice versa); if \( y \) were not true than \( x \) would not be true. The distinction between full ground and partial ground also matters here. Roughly, \( x \) fully grounds \( y \) if \( x \) alone is sufficient explain \( y \), and \( x \) partially grounds \( y \) if \( x \) together with some other things is sufficient to explain \( y \).
Constitutive Construction

A property $F$ is constitutively constructed by a social factor $S$ iff for all $x$, if $x$ is either $F$ or $S$, then $Fx$ is grounded in $Sx$, and that grounding relationship is itself explained by the essence of $F$.

To see how this analysis might be applied to illuminate the metaphysics behind constitutive social construction, let’s work with another simplified example. Consider the property of being cool. Quite plausibly, this property is constitutively constructed (Haslanger 2012). For the sake of example, let’s say that necessarily, someone is cool just in case they are charismatic. Applying Rosen’s view here, we might say that what it is to be cool is to be charismatic, such that if someone (Suzy, say) is cool, then she is cool because she is charismatic. In other words, the fact that Suzy is cool is grounded in the fact that she is charismatic, and that grounding relationship is itself explained by the essence of the property of being cool.

In this case, the property of being charismatic is doing the work that the property of being a closed, three-sided figure was doing in our earlier example. But notice that the property of being charismatic is itself a social property. For example, it’s likely that charisma is constructed in different ways in different contexts, such that what it takes to be charismatic is itself dependent on social factors. This is an example in which a constitutively constructed property (the property of being cool) is partially defined in terms of a social factor (the property of being charismatic). And so, I suggest this is a case where a social factor is in the real definition of a constitutively constructed property.
§6 Evaluating the View

At this point, we have an analysis of constitutive construction on the table. Now it’s time to evaluate it. Is this a useful notion of social construction? Will it help us accomplish the goals of projects in social ontology? In order to answer these questions, we first have to have some sense for what makes for a useful notion of social construction in the first place. That is, what do we want from a relation of social construction?

In the beginning of this paper I went through a brief list of things that social ontologists typically want from a relation of social construction (§2). That list was distilled into four claims, each of which should follow from the claim that $x$ socially constructs $y$:

- **CONTINGENCY**: $y$ is contingent.
- **EXPLANATION**: $x$ at least partially explains $y$.
- **GENERATION**: where $x$ and $y$ are objects or properties, $x$ is at least part of what makes $y$ the entity that it is; or, where $x$ and $y$ are facts, $x$ is at least part of what makes it the case that $y$.
- **DEPENDENCE**: where $x$ and $y$ are objects or properties, $y$ at least partly depends upon $x$, such that (necessarily) if $y$ doesn’t exist then $x$ doesn’t exist; or, where $x$ and $y$ are facts, $y$ at least partly depends upon $x$, such that (necessarily) if not-$y$ then not-$x$. 
Now here is our question: Does the analysis of constitutive construction proposed in this paper deliver CONTINGENCY, EXPLANATION, GENERATION, and DEPENDENCE? If so, we are in good shape. For ease of reference, here again is the view on offer:

*Constitutive Construction* A property $F$ is constitutively constructed by a

*(real definition reading)* social factor $S$ iff for all $x$, if $x$ is either $F$ or $S$, then $Fx$ is grounded in $Sx$, and that grounding relationship is itself explained by the essence of $F$.

In the interest of keeping things concise, in this discussion I will use brackets like this [·] to indicate facts. So from here on, facts like the fact that Suzy is $F$ will look like this: [Suzy is $F$]. For the sake of this example, I’ll focus on two facts, [Suzy is $F$] and [society is $S$], where $S$ stands generally for the social factors which construct $F$. A bit more carefully then, here we’re considering the claim that $S$ constitutively constructs $F$, and we want to know whether it follows that:

**CONTINGENCY:** Facts like [Suzy is $F$] are contingent.

**EXPLANATION:** [society is $S$] partially explains facts like [Suzy is $F$].

**GENERATION:** [society is $S$] is at least part of what makes it the case that facts like [Suzy is $F$] are true.
DEPENDENCE: Facts like [Suzy is F] at least partly depend upon [society is S], such that (necessarily) if not-[Suzy is F] then not-[society is S].

Let’s start with CONTINGENCY. Does CONTINGENCY follow from the claim that S constitutively constructs F? Yes. Here, it matters that social factors like S are themselves contingent. If F is a constitutively constructed property, then social factors are in the real definition of F. And, given that social factors are contingent, if the social factors which are in the real definition of F fail to obtain, then nothing can have property F.

For example, let’s say that F is the property of being a woman, and for the sake of this example, let’s say that necessarily, someone is a woman iff they are taken to be feminine. If the property of being a woman is a constitutively constructed property, then part of what it is to be a woman (F) is to be taken to be feminine. Now of course, if no one takes anyone else to be feminine, then no one will be a woman, because what it is to be a woman is to be taken to be feminine. In this way, facts like [Suzy is F] are contingent, because the social factors which make these facts true only obtain contingently.

Note that EXPLANATION, GENERATION, and DEPENDENCE all follow from the claim that S constitutively constructs F for related reasons. On the proposed analysis of constitutive construction, where S constitutively constructs a property F, facts like [society is S].

---

7 Here, careful readers might notice a difference between the relata of the constructing relation (properties) and the four claims here (which are all about facts). This is the case because, given the analysis of constitutive construction on offer, one property constructs another property just in case a certain relation obtains between facts about the people/things with those properties. And this, I think, is a feature of social construction relations: that they obtain between many different sorts of relata.
S] ground facts like [Suzy is $F$]. And given the nature of the grounding relation, it follows that: [society is $S$] at least partly explains [Suzy is $F$], [society is $S$] is part of what makes it the case that [Suzy is $F$], and if not-[Suzy is $F$] then not-[society is $S$]. In sum, where the claim is that some social factors $S$ constitutively construct a property $F$, CONTINGENCY, EXPLANATION, GENERATION, and DEPENDENCE follow.

This tells us that the analysis of constitutive social construction just proposed promises to be useful to social ontologists, insofar as it satisfies the basic desiderata on a relation of social construction. As we’ve seen, this analysis of constitutive construction has advantages over constitution-based analyses of constitutive construction, at least insofar as it is not met with the same challenges. In the next section, I will argue that the analysis of constitutive construction proposed here also has advantages over a grounding-only approach to social construction.

§7 Social Construction: Real Definitions vs Grounding

At this point we have discussed three relations of social construction: the proposed analysis of constitutive social construction, constitutive construction as constitution (constitutivism), and causal social construction was mentioned briefly in the introduction. There are several other relations of social construction available in the literature, but one of them closely resembles the analysis proposed here. This is social construction as grounding, or the general idea that something is socially constructed just in case it is grounded in social factors (Epstein 2015; Griffith 2018; Mikkola 2015; Passinsky 2020; Schaffer 2017). For the purposes of our discussion here, call this view grounding construction.
As just one example of this view, consider Jonathan Schaffer’s (2017) characterization of grounding construction, according to which “to be socially constructed is to be grounded in social patterns” (Schaffer 2017, p. 2454). On Schaffer’s view, grounding is a relation of metaphysical explanation, generation, and dependence. As such, note that on Schaffer’s view, CONTINGENCY, EXPLICATION, Generation, and Dependence follow the claim that $x$ grounding constructs $y$ for the same reasons that they follow from the claim that $x$ constitutively constructs $y$. For example, where the claim is that some social factors $S$ grounding construct a property $F$, it is true that:

A. Facts like [Suzy is $F$] are contingent.

B. Facts like [society is $S$] explain why facts like [Suzy is $F$] are true.

C. Facts like [society is $S$] generate facts like [Suzy is $F$]

D. Facts like [Suzy is $F$] are dependent upon facts like [society is $S$], such that, necessarily, if not-[Suzy is $F$] then not-[society is $S$].

That is, where the claim is that some social factors $S$ grounding construct a property $F$, CONTINGENCY, EXPLANATION, GENERATION, and DEPENDENCE follow. So at this point, it doesn’t look like constitutive construction has a leg up on grounding construction. Furthermore, it looks like constitutive construction and grounding construction satisfy these desiderata for the same reasons, and so constitutive construction actually looks worse, because it includes some complicated metaphysical machinery (i.e., real definitions/essence) that it doesn’t use.
At this point in the discussion, you might wonder why the proposed analysis of constitutive construction appeals to real definitions (essence) in the first place. As I’ve just pointed out, all of the benefits of constitutive construction listed in the previous section were achieved by working with grounding, and all of them, it seems, could be had without appeal to essence. In response to this, I’ll mention two reasons for working with essences here, in addition to grounding.

First, consider the idea that some socially constructed things could be (and have been) realized by a number of distinct grounds. For example, Haslanger (2000) remarks that women have been systematically oppressed through the centuries, but different patterns of behavior have generated this oppression in each time period. Put in terms of grounding, the idea here is that the same systems of oppression have been grounded by different things at different times. Or in other words, these systems of oppression have been multiply realized by different grounds. Many grounding frameworks can accommodate the multiple realizability of ground, so that isn’t a problem. The difficulty here is one of accounting for which things are possible grounds for a social construct, and which aren’t. This matters because often in social ontology, we are interested in what it takes to eliminate bad or harmful social constructs. We might wonder, for example, whether certain oppressive social structures would still exist if we stopped engaging in certain patterns of behavior, or if that oppressive social structure would stick around, and just become grounded in something else. That is, we need it to be the case that a social construct can have multiple possible grounds, but it would also be helpful to have a principled way to say which grounds those are.
The claim that \( x \) grounds \( y \) is effectively a sufficient condition for \( y \): where \( x \) is a full ground for \( y \), if you've got \( x \), then necessarily, you get \( y \). This means that grounding alone won’t tell you which grounds (of all the possible grounds) are the ones that possibly ground \( y \), because that’s just not what grounding is designed to do. But real definitions can do this. The real definition of a social construct is a statement of its essence, and the essence of e.g., a socially constructed property (like the property of being a woman) will tell us all and only the possible grounds for being a woman. Why? Because quite plausibly, if the essence of the property of being a woman tells us anything, it tells us the full range of ways that possibly make someone a woman.

There is another reason to prefer constitutive construction here, too. When we are theorizing about some socially constructed entity \( y \), while we’re often interested to know why it is that \( y \) obtains, we’re also interested in the nature of \( y \). And what’s more, many times we’re interested in how what it is to be \( y \) might be connected to explanations as to why \( y \) is the case. Take gender for example: there are projects in feminist metaphysics which are concerned with giving an account of what sort of thing gender is, and many of those same projects also aim to explain how gender came about, as well. And, many times, accounts of what gender is are taken to do some work toward explaining how gender came about. For example, consider Haslanger’s (2012) account of womanhood: according to Haslanger, part of what it is to be a woman is to occupy a certain position within a hierarchy of power relations. Furthermore, the existence of such a hierarchy of power relations in which some individuals are oppressed and others are not (at least partly) explains the existence of women (Haslanger 2012, p. 234).
That said, grounding alone doesn’t open a route to understanding what it is for something to be \( y \), where \( y \) is a socially constructed entity. This is just because this isn’t something ground-theoretic frameworks are designed to do. Grounding tells us why \( x \)’s being the case gives us \( y \), but on its own, it doesn’t tell us about the nature of \( y \), and it also doesn’t tell us why it is that \( x \) grounds \( y \). This isn’t a problem for grounding, but insofar as many projects in social ontology are concerned with investigating the nature of socially constructed entities, a theory of social construction which only accommodates half of such projects falls short.

The analysis proposed here gives us a way to answer all of these questions however. Consider once more the property of being a woman, and say that it is constitutively constructed, such that what it is to be a woman is to be seen as feminine. So the fact [Suzy is a woman] is grounded in [Suzy is seen as feminine]. Why is Suzy a woman? Because she is seen as feminine. What is it to be a woman? To be a woman is to be seen as feminine. Why does the fact that Suzy is seen as feminine make her a woman? Because that’s just what it is to be a woman: to be seen as feminine. In other words, essences effectively give us a deeper story—they tell us that necessarily, Suzy is a woman because Suzy is seen as feminine (which is the grounding bit), and they also tell us that that fact (i.e., [Suzy is a woman because Suzy is feminine]) explained by the essence of womanhood.

§8 Conclusion

In this paper I hope to have accomplished three things. First, and most centrally, I developed an analysis of constitutive social construction in terms of real definitions:
Constitutive Construction  
A property $F$ is constitutively constructed by a social factor $S$ iff for all $x$, if $x$ is either $F$ or $S$, then $Fx$ is grounded in $Sx$, and that grounding relationship is itself explained by the essence of $F$.

I then considered reasons to favor this analysis, among which is the fact that CONTINGENCY, EXPLANATION, GENERATION and DEPENDENCE follow from the claim that $x$ constitutively constructs $y$. I also argued that this analysis has advantages over constitution-based analyses of constitutive construction, as well as grounding approaches to social construction. Of course, there is more to be said about how this analysis stacks up against other relations of social construction. But given the centrality of constitutive social construction to projects in social ontology, I hope the present paper has made some helpful headway in illuminating the metaphysics behind this relation.

Acknowledgements

Many thanks to Ásta, Matthew Andler, Robin Dembroff, Esa Díaz-León, Carolina Flores, Daniel Forman, Verónica Gómez, Sally Haslanger, Lauren Lyons, Rebecca Mason, Sean Nalty, Katherine Ritchie, Ezra Rubenstein, Jonathan Schaffer, Ted Sider, Neal Tognazzini, Danny Underwood, Ryan Wasserman, Isaac Wilhelm, and audiences at Rutgers University, Western Washington University, and the Bay Area Feminism and Philosophy Workshop (BayFAP) for helpful feedback on this essay.
CHAPTER TWO

Constructing Properties with Words

Abstract. It is common to think that predicates like ‘cool’ express socially constructed properties. In other words, the meaning of ‘cool’ is a socially constructed property. Consider two questions in this area: first, what makes it true that ‘cool’ expresses the property it does? And second, what makes it true that a given individual has the property ‘cool’ expresses? This paper begins with the observation that, in the social metaphysics literature, these two questions are often answered in the same way. I then suggest that this observation may lend us some insight into the nature of sociality more generally, by highlighting one special feature of social properties.

§1 Introduction

Consider the claim that Suzy is a woman. According to several prominent views of gender, when we say that Suzy is a woman, we’re saying that Suzy has a certain property (a gender property, if you will), and because she has that property, ‘woman’ applies to Suzy (Ásta 2018; Haslanger 2012; Jenkins 2016). A little bit more precisely: the predicate ‘woman’ expresses a certain property; Suzy has that property, and that’s what makes her a woman.

Of course, it is natural to say that the property ‘woman’ expresses is just the property of being a woman. I am happy with that view in general, but sometimes using ‘woman’ to talk about both a predicate and a property can cause confusion. So for the purposes of our discussion here, let’s call the property ‘woman’ expresses, ‘\(W\)’. Now we have two facts on the table:

A. Suzy is \(W\)

B. ‘woman’ expresses \(W\)
Many people think properties like $W$ are socially constructed. There are lots of views about how social construction works, and although they each say something different, there is rough consensus in the idea that people have socially constructed properties because of things that we do. For example, the fact that Suzy is $W$ is true because of things we do—the ways in which we think, talk, communicate, and coordinate with each other give rise to the social world, and with it, socially constructed properties.

In this paper, I’ll use brackets like $[·]$ to indicate facts. So from here on, facts like the fact that Suzy is $W$ will look like this: $[\text{Suzy is } W]$. Now here’s an important question: how does human social activity make a fact like $[\text{Suzy is } W]$ true? In reply, social constructionists will point to a number of things, but crucially among them are patterns of language use. That is, social constructionists will say that facts like $[\text{Suzy is } W]$ are true in part because of how we communicate with each other using natural language. In other words, the language we use to describe the social world plays a role in constructing the very world we are trying to describe (Hacking 1999; Mallon 2007; Haslanger 2012; Epstein 2015; Ásta 2015). On this view, $[\text{Suzy is } W]$ is made true by things that we do. And crucially, patterns of language use are among the things that make $[\text{Suzy is } W]$ true.

What about our second fact, $[\text{‘woman’ expresses } W]$? What makes that fact true? Many people working in the philosophy of gender and the literature on social construction are semantic externalists, in the sense that they think, roughly, that the fact $[\text{‘woman’ expresses } W]$ is explained by other facts about how we are related to the external world (Haslanger 2000, 2012; Mason 2020; Bettcher 2013; Dembroff 2018; Jenkins 2016). Here, semantic externalism is treated as a metasemantic view about how words (like ‘woman’) come to mean things (like $W$). What exactly about the world makes the fact $[\text{‘woman’ expresses } W]$ true? Here, people working in the literatures just mentioned will point to a number of things, but again, crucially among them are patterns of language use. That is, many semantic externalists in these literatures will say that facts like $[\text{‘woman’ expresses } W]$ are true.
in part because of how we communicate with each other using natural language (Kripke 1980; Haslanger 2012; Mallon 2017; Spencer 2019).

Interestingly, there is some overlap in the explanations for the two facts we have been considering. That is, the fact [Suzy is $W$] and the fact ['woman' expresses $W$] are explained by some of the same things. To see this, notice that we have just painted a very general picture according to which the following two things are true:

C. Suzy is $W$ because of patterns of language use.

D. ‘woman’ expresses $W$ because of patterns of language use.

This fact about a property [Suzy is $W$] and this fact about language ['woman’ expresses $W$] are made true by the same things.

Is this relationship, between [Suzy is $W$] and ['woman’ expresses $W$], somehow unique? In this paper I will explore one particular way of answering “yes” to this question. On the view proposed, this relationship between ‘woman’ and properties like $W$ is unique insofar as it is one that never obtains between putatively non-social properties, and the words that we use to talk about those properties (e.g., words like ‘electron’ and ‘proton’). And if that’s right, it suggests that the relationship highlighted here (between words like ‘woman’ and properties like $W$) might be a mark of sociality. In other words, perhaps this relationship obtains between [Suzy is $W$] and ['woman’ expresses $W$] but not between e.g., [electron, is $E$] and ['electron’ expresses $E$] because $W$ is a social property (and $E$ is not).

Here is what the rest of this paper will look like. In the next section (§2) I will work through a more specific example than the one given here, in order to work out the relationship sketched above in
more detail. I will then suggest that the presence of this relationship is a point of difference between putatively social properties and putatively non-social properties (§3). Then I will consider one benefit of this proposal (§4) and two objections to it (§5) before concluding (§6).

§2 A Cool Case Study

In this section I will work through an example involving the predicate ‘cool’. In §2.1 I will discuss various views according to which ‘cool’ comes to express a property (a property which we’ll call ‘C’). In §2.2 I will discuss one way of characterizing what it is to be C. Then in §2.3 I will bring these discussions together to highlight an interesting relationship that holds between facts about ‘cool’ and C.

2.1 Predicates & Properties

Consider the predicate ‘cool’ and suppose that it expresses a property, C. I’ve decided to work with ‘cool’ for this example because, in the literatures with which this paper is concerned, it’s relatively uncontroversial that ‘cool’ expresses a socially constructed property (C). And it is for at least this reason that C has become a stock example in social metaphysics. For instance, Sally Haslanger (2012) argues that C is a constitutively constructed property: people and things are C because we see them in a certain way, and so treat them in a certain way (and not because they have some special intrinsic quality that we are somehow latching onto with our usage of ‘cool’) (Haslanger, 2012, p. 89).

What makes it true that ‘cool’ expresses C? I said previously that many people think that facts like [‘cool’ expresses C] are true at least partly because of how we use natural language. For reasons that will become apparent later, I’m going to put this claim in terms of grounding.¹ Here the relevant

¹ As I will use it here, grounding is a relation of metaphysical explanation which obtains between facts (Rosen 2010). When one fact, x, grounds another fact, y, the following things
claim is this: the fact ['cool' expresses $C$] is partly grounded in at least one fact about how we use natural language. Call that fact about how we use natural language, ‘$F$’. When I say that ['cool' expresses $C$] is partly grounded in $F$, I mean that $F$ is just one element of a full ground for ['cool' expresses $C$]; things besides $F$ ground ['cool' expresses $C$]. For now, consider a set of facts which jointly fully ground ['cool' expresses $C$] and which also includes $F$, and call that set $\mathcal{G}$.

$\mathcal{G}$ and $F$ will be central to the project of highlighting the connection between $C$ and ‘cool’ in coming sections. But, in order to see this, we’ll first need to say a bit more about what $C$ is like.

2.2 What is it to be $C$?

$C$ is a property expressed by our predicate ‘cool’, and $C$ is also a constitutively constructed property. But, what does that latter claim really mean? Sally Haslanger (2012) first characterized constitutive social construction as follows:

\[ \text{Constitutive Construction} \quad \text{Something is constitutively constructed just in case in defining it we must make reference to social factors.} \]

(Haslanger 2012, p. 87)

There is some disagreement over how Haslanger’s original definition is best interpreted, but I favor a reading according to which a property is constitutively socially constructed just in case its real definition makes reference to social factors.\(^3\) That is, social factors are “in” the real definition of that property.

\[^2\] There may be more than one set which both consists of facts that jointly fully ground ['cool’ expresses $C$] and which also includes $F$.

\[^3\] Haslanger is more explicit about this position in her more recent work. See Haslanger (2014). See Ásta (2015, 2018) for alternative readings of constitutive social construction.
Roughly, the real definition of a property is just a statement of the essence of that property.\textsuperscript{4} Furthermore, on my preferred view, the essence of a property “contains” conditions under which that property is instantiated: every property has essential instantiation conditions which are “in” its real definition.\textsuperscript{5} For example, say that it is essential to the property $C$ that someone is $C$ just in case they have stylish hair. So the fact [$S$ is $C$ iff $S$ has stylish hair] is “in” the essence (real definition) of $C$. Note that if “having stylish hair” is a social factor, then in this example $C$ has at least one social factor in its real definition. On this view, social factors are essential to constitutively constructed properties, in the sense that the conditions under which these properties are instantiated include social factors. \textit{What it is} to be that property is to be instantiated under those conditions, conditions which incorporate social factors.

Back to the example at hand: which social factors are in the real definition of $C$? The same things mentioned earlier when it was said that what it is to be $C$ is to be treated in certain ways (Haslanger 2012). Put somewhat differently, those patterns of behavior (those “certain ways” in which $C$-individuals are treated) are the social factors “in” the real definition of $C$. Whether or not something is cool (i.e., has property $C$) is determined by those patterns of behavior.

2.3 Summing Up

Recall that $\mathcal{G}$ is a set that includes the things which (metaphysically) explain why ‘cool’ expresses $C$. And furthermore, $\mathcal{G}$ includes a fact about how people interact with each other, a fact we’ve called $F$. What I want to bring out here is that $F$ is likely in the real definition of $C$. In other

\textsuperscript{4} Real definitions are different from nominal definitions. In a nutshell, nominal definitions are generally taken to specify meanings of terms while real definitions are generally taken to specify the essences of the things to which those terms refer. (Rosen 2015)

\textsuperscript{5} Here I’m working with Fine’s notion of constitutive essence (1994a, p. 57). See also Fine (1994b).
words, facts about how we interact with one another (i.e., facts like \( F \)) contribute to making it true that a given individual has property \( C \). This is because \( C \) is constitutively socially constructed, and here, we’ve taken this to mean that social factors (like \( F \)) are in the real definition of \( C \). And given that facts like \( F \) are in the real definition of \( C \), they are among the grounds for facts like [Suzy is \( C \]).

So the same things that grounds facts like ['cool' expresses \( C \)] also ground facts like [Suzy is \( C \)].

\[ \text{§3 A Mark of the Social} \]

In the last section I used an extended example involving the predicate ‘cool’ and a property it expresses (\( C \)), to illustrate in further detail the observation sketched in the introduction. A new question guides the present section, where we are no longer concerned with what this relationship is like, but why it holds in the first place. In short, my proposal here is that this relationship holds in part because properties like \( C \) are social properties. That is, the relationship just explored between predicates like ‘cool’ and the properties they express may give us a bit of insight into the nature of sociality more generally.

Why think that? Well intuitively, the relationship just highlighted between ‘cool’ and \( C \) does not hold between putatively non-social properties, and the predicates we use to talk about them. For example, consider our predicate ‘electron’ and suppose it expresses a property, \( E \). In addition, let’s consider a particular electron which we will hereby name ‘electron\(_1\)’. In order for the relationship outlined in the previous section (§2) to hold here, the following two things have to be true:

E. \( [\text{electron}\(_1\) is \( E \)] \) is partially grounded in patterns of language use.

F. \( ['\text{electron}' expresses \( E \)] \) is partially grounded in patterns of language use.

\[ ^6 \text{Note that the claim here is not that the elements of } \mathcal{G} \text{ are the only things in the real definition of } C, \text{ just that they are among the things in the real definition of } C. \]
Plausibly, (F) is true and (E) is false. This is because, while the ways in which we communicate with each other plausibly play some role in making facts like ['electron' expresses $E$] true, facts like [electron$_1$ is $E$] have nothing to do with how we use natural language. That is, whether a given entity has property $E$ is wholly independent of how human beings use natural language. And in general, I think we would be hard-pressed to find examples of putatively non-social properties for which it is the case that patterns of language use partially determine which things have those properties. This is primarily because the real definitions of non-social properties (like the property of being an electron) don’t contain information about how we represent them: they are, by nature, independent of us.

If there is indeed an asymmetry between the case of ‘cool’ and $C$, and the case of ‘electron’ and $E$, then that is worth exploring. Why might this asymmetry hold? Here I suggest that it is because the relationship just explored between ‘cool’ and $C$ is a mark of the social. In what follows, I try to make this idea a bit more precise. To begin, consider the following proposal:

**SOCIAL**

For any property $P$, representation $R$, and fact about language use $F$, if

(i) $F$ is a partial ground for [$R$ represents $P$], and

(ii) $F$ is in the real definition of $P$

then $P$ is a social property.

**SOCIAL** is a sufficient condition on sociality for properties. What makes a property a social property? **SOCIAL** offers an answer. I’m using ‘representation’ here in a very broad sense: beliefs and concepts are representations in this sense, as well as lower-level mental states such as emotions and desires. Linguistic items are also representations in this sense, and this includes predicates like ‘woman’ and ‘cool’, predicates commonly discussed in the literature on social construction.
I have been working with examples of social construction that highlight this relationship as it obtains between predicates and properties, primarily because the literature on social construction has devoted most of its attention to these sorts of cases. The example with ‘woman’ sketched in the introduction was one such case. There, it turned out that the questions “Why does Suzy have property $W$?” and “Why does ‘woman’ express $W$?” had similar answers. And the same thing happened in the next section, when it turned out that the questions “Why does Suzy have property $C$?” and “Why does ‘cool’ express $C$?” had similar answers. That overlap, I suggest, is a mark of the social. In the next section I explore one fruitful application of \textit{SOCIAL}, before considering some objections (§5).

\section{Applications: A Question for Debunkers}

In this section I suggest that \textit{SOCIAL} can help us answer a particularly difficult question for proponents of debunking arguments in the literature on social construction. First, I'll explain the basic strategy behind debunking arguments; then, I will raise an objection for these arguments and point out that existing literature does not have the resources to respond to this worry without circularity; then, I will show how \textit{SOCIAL} can provide a non-circular response to this objection.

For a very long time, people thought sex and gender were the same thing. Indeed, some people still think this. And for many feminists, this is a problem. Conflating sex with gender leaves us with limited resources with which to characterize the manner in which certain people are oppressed, and to explain why this is the case. It also leaves us without the means to account for how people understand themselves and their relationships to their bodies in societies which place a heavy emphasis on sex and sexuality. And not only does the view that $\text{sex} = \text{gender}$ deprive us of these resources, it also makes it seem like we don’t need those resources in the first place, and in so doing obscures crucial elements of reality.
For at least these reasons, many feminist scholars think the view that \( \text{sex} = \text{gender} \) should be debunked as false: gender is not the same thing as sex. Arguments like this are sometimes called *debunking arguments*. A debunking argument aims to show that certain phenomena that we thought fell on the non-social side of this divide actually fall on the other (Haslanger 2003). Typically, this is done by arguing that the item in question is socially constructed. Simone de Beauvoir’s now-famous claim that “one is not born but becomes a woman” is one to this effect: it is social factors (and not biological ones) which are responsible for the fact that some people are women (de Beauvoir, 1949).

Note however that debunking arguments turn on a deeper distinction between the social and the non-social: that is, claims about the social construction of e.g., gender are effectively claims about how gender depends on human *social activity*. But then, what makes the things that construct gender social? Note that it won’t help to say that they are also socially constructed; that just pushes our question back a level. The issue here is this: we want to know what makes properties like the property of being a woman *social*, rather than not. Debunking arguments answer this question with the claim that the property of being a woman is socially constructed, and *that* is what makes it a social property. But note that with this response, the debunker invokes sociality, which is the very thing they purport to explain. If the idea is to answer the question of sociality by appeal to social construction, where social construction itself is analyzed in terms of sociality, that just won’t work, because it is circular. Therefore, it would seem that we need to appeal to something else to draw the line between the social and the non-social.

Despite its centrality to debunking arguments, this distinction between the social and the non-social is nearly always assumed and never argued for. There are a number of candidate reasons for this: for one thing, in many cases the distinction between what’s social and what’s not is intuitive enough for most of us that it needn’t be elaborated upon. For another thing, pausing to argue for this distinction
would take many projects too far afield from the task at hand, with marginal returns. These are pragmatic concerns, but there are theoretical worries, too.

In particular, social ontologists have previously expressed serious doubts about the prospects of drawing the social/non-social distinction in a principled way: for example, Brian Epstein states that he is “not confident” that distinguishing clearly between social and non-social facts can be done (2015, p. 102). Katherine Ritchie writes that “giving a noncircular definition of what it takes for something to be social is difficult, if not impossible” (2020, p. 404). And Sally Haslanger concurs, insofar as she thinks “it is unlikely that there is a non-circular definition [of social]” (2016, fn.8).

The question of sociality is an important question for debunking arguments, and for the social construction literature more broadly. I agree that this distinction between the social and the non-social is a slippery thing to pin down. My own view is just that, despite this, given the importance of this distinction, the more we can say about it, the better.

_SOCIAL_ offers one way to mark this social/non-social divide. The idea here is that properties are social if they satisfy the conditions laid out in _SOCIAL_. Note then, that this way of approaching the question of sociality differs from the strategy employed in most debunking arguments. Traditional debunking arguments aim to show that some property (like the property of being a woman) is social rather than not by showing that the property in question is _socially constructed_. But _SOCIAL_ offers a different strategy. Where the question is: Why should we think that some property \( P \) is social rather than not? The answer may be: because it satisfies the condition presented in _SOCIAL_. 
SOCIAL. For any property $\mathcal{P}$, representation $\mathcal{R}$, and fact about language use $\mathcal{F}$, if

(i) $\mathcal{F}$ is a partial ground for $[\mathcal{R} \text{ represents } \mathcal{P}]$, and

(ii) $\mathcal{F}$ is in the real definition of $\mathcal{P}$

then $\mathcal{P}$ is a social property.

What makes the property of being a woman a social property? SOCIAL offers a non-circular way of answering that question. That is, in saying that a property is social because it satisfies the condition in SOCIAL, we are not thereby invoking the very thing (i.e., sociality) that we’re trying to explain.

Crucially, SOCIAL is not a full analysis of sociality, and so it may be that some things are social which SOCIAL does not count as such. But just as importantly, SOCIAL gives us something to use to begin to answer this question. Namely, if a property meets the condition stated in SOCIAL, that means it is a property like the one expressed by predicates like ‘cool’ and ‘woman’. And if a property does not meet the condition stated in SOCIAL, that means it is not a property like the one expressed by predicates like ‘cool’ and ‘woman’. That is helpful information. Before, when asked “what makes the property of being a woman a social property?”, we had nothing non-circular to say in response. Now, with SOCIAL, we have something.

§5 Objections

In this final section, I’ll consider two objections to SOCIAL.

5.1 Some socially constructed properties don’t have representations.

For any given human language, there are potentially many social properties that go unrepresented by that language. These are social properties which exist, but we just don’t have words for them in our present lexicons. For example, the English term ‘sexual harassment’ was only introduced in the mid-1970s. This term refers to actions that have a certain property, but surely acts
had that property prior to the 1970s. And, surely, given that we consider sexual harassment to be a social phenomenon now, it was a social phenomenon prior to 1970—that is, the property didn’t suddenly become social with the introduction of the term ‘sexual harassment’.7 To suggest otherwise is to suggest something false.

This objection targets the fact that the discussion so far has focused on a relationship between predicates and the properties they express, and points out that not all social properties are expressed by a predicate. These properties look like they fall outside the scope of SOCIAL, and if that’s right, that is a problem. However, even though these properties are not represented by lexical items, they are still represented, and crucially, our representations of them play a role in determining the conditions under which they are instantiated. For example, an instance of what we would now call sexual harassment that happened in the 60’s would not have been represented by the harassee with the predicate ‘sexual harassment’ (or with its associated concept(s)). But, it would still be represented by the harassee as e.g., harmful, demeaning, and as a personal violation, and in virtue of being represented in that way, it would have a certain property (the property which is now expressed by ‘sexual harassment’).

I have focused on social properties that we do have words for (i.e., ‘woman’, ‘cool’) because these properties take up the most space in the literature on social construction. However, I think it’s likely that there are a number of social properties that we don’t (yet) have words for, and it is a feature of the more general analysis presented here that it can account for why some of those properties are social, too.

Finally, here it will be good to remember that SOCIAL is not a full analysis of sociality, but merely a sufficient condition on sociality for properties. There may be social properties which SOCIAL does

7 This example is adapted from Fricker (2007).
not count as such. And there may be other social entities which fall completely outside the scope of SOCIAL. But these are not counterexamples to the view. A counterexample would take the form of a putatively non-social property which satisfies both (i) and (ii) in SOCIAL. But as I pointed out earlier, there probably aren’t any properties like that (§3).

5.2 Socially constructed properties could have been represented differently.

There is another worry in this vicinity, one which turns on the idea that we could have used different predicates to express the properties that we do talk about. Here is how this objection goes:

I’ve said that a partial ground for the fact ['cool' expresses C] is in the real definition of C. Presumably, any full ground for this fact about ‘cool’ will include information about the word ‘cool’, and so, information about the word ‘cool’ is in the real definition of C. But intuitively, it seems that we could have used a word other than ‘cool’ to express C, in which case information about that other word would be in the real definition of C. But then, to change the real definition of C is to change what it is to be C, and so strictly speaking, we wouldn’t be using that other word to express C, but rather to express a different property which is similar to C. And that is counterintuitive. It seems like we could have used different words to talk about the same properties, not slightly-different-but-still-very-similar properties.

This is a tough objection to answer, but I think there is a bit of room to maneuver here. For example, we might say that what it is to be C is to be treated in certain ways, and those “ways” change (albeit very slightly) when we change how we talk about C. On a view like that, people who use the predicate ‘womxn’ and people who use the predicate ‘woman’ are talking about very similar properties, but not the same property. Similarly, on this view, people who use the predicate ‘African American’ and people who use the predicate ‘Black’ are talking about very similar properties, but not the same property.
To see how this could be the case, consider a world which is exactly like ours up until a time $t$. At $t$, the meanings of ‘man’ and ‘woman’ are switched by the people in power, such that everyone that we would say is a woman has to be called a man at that world, and everyone that we would say is man has to be called a woman at that world. After time $t$ at this possible world, sexist people say things like “Men are so irrational and hysterical sometimes!” and “You silly man!” when they are speaking derogatorily toward people we would call women. Simply put, I don’t think individuals at that possible world who say “Men are so irrational and hysterical sometimes!” are treating the people we would call women in the same way as we do when we say “Women are so irrational and hysterical sometimes!” at the actual world.

In this case, despite the abrupt change in language use, societies in the possible world just described still have a patriarchal structure—individuals who appear to have male bodies still enjoy a certain amount of privilege, and the word ‘man’ is still associated with those privileges. So, even though the people at that world who we would call women are still oppressed, they are now being referred to with a word to which privilege is attached. And it’s possible that this is enough to make it the case that those individuals are treated slightly differently than the women at our world, even though both groups of individuals are still subordinated in very similar ways.\(^8\)

\section{Conclusion}

In first half of this essay, I highlighted an interesting relationship between predicates like ‘cool’ and ‘woman’, and the properties they express. Then I suggested that this relationship might be unique, in that it is a relationship which never obtains between intuitively non-social properties and the terms we use to talk about non-social properties. Why would this relationship hold between

\footnote{Note that this view also has the consequence that to call something “cool” in a language other than English is to treat that thing in a slightly different way. So for example, if I describe my bike as “cool”, and my Spanish-speaking friend describes it as “guay”, we are each treating my bike in a different way.}
predicates like ‘woman’ and the properties they express, but not between predicates like ‘electron’ and the properties they express? Perhaps because this relationship is a mark of the social.

In the second half of this essay I attempted to spell out what that might mean in a bit more detail with SOCIAL. SOCIAL is a sufficient condition on sociality for properties. This proposal has issues (§5), and so it may turn out that other ways of capturing the relationship highlighted here are more promising. But even in that case, my earlier arguments hold: there is something interesting about the relationship between some social properties and the words we use to talk about them, something that isn’t present in the non-social case. And that, I think, is worth exploring.

Acknowledgements

My deepest thanks go out to Karen Bennett, Elisabeth Camp, Carolina Flores, Anthony Gillies, Cal Howland, Zachary Kofi, Katherine Ritchie, Jonathan Schaffer, Theodore Sider, Christopher Willard-Kyle, and Isaac Wilhelm for helpful discussion about this essay.
CHAPTER THREE
Gender Metaphysics

Abstract. The metaphysics of gender is a rich and rapidly expanding area of analytic philosophy. This essay takes up one of its core questions, *What is it to be gender g?*, and argues that there are actually two very closely related questions in the area of this one sentence. The first is a question about language: what does the term ‘g’ mean? The second is a question about metaphysics: which property is g? These two questions are not asking after the same thing, or so I will argue here. Furthermore, once we distinguish these two questions, another question arises regarding the relationship (if any) between our answers to each. My aim in this essay is not to provide definitive answers to these questions, but instead to make the case for asking them each independently, and considering them in relation to one another.

§1 Introduction

These days, many people think that gender is real. In feminist philosophy, this general view often goes by the name of *gender realism* (Barnes 2017; Mikkola 2001). There are a number of ways to be a gender realist (Haslanger 2012; Mikkola 2011; Witt 2011), but on many of these views gender is real *at least* in the sense that there are gender properties.¹ Why be a gender realist? Some people endorse this view because they find claims like the following plausible, claims which they take to imply the existence of gender properties:

(i) Our gender-talk accurately describes the world in a metaphysically robust sense.

(ii) Some people are materially privileged/oppressed on account of their gender.

¹ Note that, on the very general version of gender realism just characterized, one need not be committed to the view that all people of a particular gender g share one property in common. This is a form of gender realism (one famously critiqued by Spelman (1988) and defended by Mikkola (2001)) but not the specific form of gender realism I am discussing here.
(iii) By dressing in certain ways and acting in certain ways it is possible to express a gender identity.

Gender properties are often identified as those properties which are expressed by certain predicates, like ‘woman’, ‘man’, and ‘genderqueer’. Much of the literature on the metaphysics of gender is concerned with what it takes to have certain gender properties. For example, philosophers have spent a lot of time with questions like *What is it to be a woman?*. In this essay, I’m not so much interested in specific answers to questions like *What is it to be a woman?* as I am with the question itself. What are we really asking when we pose this question?

The central contention of this essay is that there are actually two closely related, but nevertheless *distinct* questions in the vicinity of the sentence, ‘What is it to be a woman?’. The first is a question about language: what does ‘woman’ mean? The second is a question about metaphysics: which property (if any) is the woman property? (Or for those who think we’re often dealing with a number of gender properties in these conversations: which properties (if any) are the woman properties?)

In order to distinguish the woman property from ‘woman’, I’ll give it a name: ‘WOMAN’. So really, our second question is: which property (if any) is WOMAN?

Essentially, the distinction between the two questions I’ve raised here is a distinction between properties and terms. Terms (like ‘woman’) are distinct from properties (like WOMAN). ‘Woman’ is a linguistic item, a string of letters which is part of the English language. WOMAN is a property—e.g., a universal, a trope, a set of possible objects—which is part of our non-linguistic reality. And it is something more besides: as it will be considered here, WOMAN is the target of theoretical analysis in the literature on the metaphysics of gender. That is, in this literature, when we ask questions like *Who

---

2 Throughout this essay I will use the singular version of this question in order to streamline my discussion. If you are a pluralist about woman properties, please read this question in its plural form.
is a woman? and What does it take to be a woman? we’re asking questions about who has this property, WOMAN, and why.

I want to remain neutral on whether WOMAN needs to be expressed by ‘woman’. That means I will also remain neutral on whether someone who has WOMAN is in the extension of ‘woman’. On some characterizations of what WOMAN is, it may turn out that everyone with WOMAN has to be in the extension of ‘woman’. On other ways of characterizing WOMAN, it may turn out that someone can have the property WOMAN without being in the extension of ‘woman’. In this essay, I will not be giving an analysis of WOMAN—I will not argue for any particular ways of answering the questions Who is a woman? and What does it take to be a woman?. I’ll just be making a case for asking about what the property WOMAN is like, independently of asking about the meanings of gender terms like ‘woman’.

Of course, here it feels natural to say that ‘WOMAN’ is just a name for the property expressed by ‘woman’; or in other words, a property is WOMAN iff it is a property expressed by ‘woman’. This is a view that many gender scholars hold (Díaz-León 2018; Jenkins 2016; Barnes 2019). In this paper, I will not argue against this view; I will simply argue that it is not the only view on the table. Notice also that this claim about the relationship between ‘woman’ and WOMAN is a way of responding to both of the questions raised here. What does ‘woman’ mean? ‘Woman’ expresses WOMAN. Which property (if any) is WOMAN? The one expressed by ‘woman’. The fact that the answers to these questions are related does not prove that the questions are the same.

§2 One Sentence, Two Questions

Our one sentence is:

(A) ‘What is it to be a woman?’

And our two questions are:
(B) What does ‘woman’ mean?

(C) Which property (if any) is \textit{WOMAN}?

My contention is not that (A) literally means both (B) and (C). This is not a paper about the semantic content of a question. Instead, my claim is that, when someone says (A) in theoretical contexts, it is not always obvious whether they are asking just (B), asking just (C), or somehow asking both (B) and (C) at once.

This is largely because many scholars in the philosophy of gender either explicitly or implicitly posit a relationship between (B) and (C). For example, on Sally Haslanger’s (2012) view, the question of what ‘woman’ means is related to the question about \textit{WOMAN}. Very generally, \textit{WOMAN} is the property which plays a role in (metaphysically) explaining certain important social phenomena (e.g., gender oppression), and in part because it plays this explanatory role, \textit{WOMAN} is a kind of “reference magnet” for our term ‘woman’. In addition, \textit{WOMAN} plays this explanatory role partly because of the ways that we think and communicate using gender terms and concepts. It turns out, then, that \textit{WOMAN} is a reference magnet for ‘woman’ in part because of how gender terms (like ‘woman’) are used.

I mention Haslanger’s view here as an example of how (B) and (C) can be treated as related questions. And there are views on which (B) and (C) are treated as even more closely related than they are on Haslanger’s picture: for example, some people working in gender metaphysics are \textit{deflationists} in the sense that they effectively take both (B) and (C) to be questions about the meaning of ‘woman’. We can understand the basic deflationist move in this literature as follows:

\begin{itemize}
  \item [Step 1] Observe a metaphysical dispute in the philosophy of gender in which parties to the dispute disagree on what it takes for a given person S to be a woman.
\end{itemize}
Step 2 Focus on the term ‘woman’ and investigate its meaning. This usually involves characterizing a number of properties, and arguing that each is, in at least some respects, acceptable as a candidate meaning of ‘woman’.

Step 3 Argue that these properties, as candidate meanings of ‘woman’, are metaphysically on a par—that is, there are no metaphysical reasons to prefer one property over the others.

Step 4 Appeal to linguistic and conceptual considerations to break this parity by asking questions like the following: who will be harmed if we say that ‘woman’ expresses (or should express) property $P$ rather than property $Q$? Who will benefit? How will this affect the distribution of certain necessary materials and resources? If the proposed reading of ‘S is a woman’ would result in the most favorable overall distribution of these materials and resources, and would result in the best overall balance of benefit-to-harm, then that proposed reading is the one we should go with.3

The general point here is that answers to the question What is it to be a woman? are often decided by appealing to how defining ‘woman’ as having this-or-that meaning will impact vulnerable populations. This way of breaking the parity between available answers to What is it to be a woman? suggests that, pragmatic considerations aside, the available readings of ‘S is a woman’ are on a par. That is what makes this strategy deflationist.

3 For examples of this strategy see, Díaz-León (2018), Jenkins (2016), Barnes (2019), and Dembroff (2018), among others. The framework provided here is loosely based on the general deflationist framework presented in Sider (2011).
But we do not need to be deflationists when we address this question. There may be at least one other way to break the apparent parity between candidate answers to the question *What is it to be a woman?*. To be clear, I am not here arguing that the deflationary strategy is a bad strategy. I’m pointing out that the prevalence of this strategy has eclipsed the fact that there is another question here, too—a metaphysical question about properties.

§3 Why This Matters

*What does ‘woman’ mean? Which property (if any) is WOMAN?* Perhaps these two questions do come apart. So what? Before I make a case for distinguishing the questions at the heart of this paper, I’m going to say something here about why this project matters. In a nutshell: if the ‘woman’ question does come apart from the WOMAN question, this reveals that certain assumptions about the relationship between gender and gender terms are not at all obvious, and so require argumentative support. In short, being aware of this distinction will allow us to do more careful gender metaphysics.

To see how, suppose I think that the meaning of ‘woman’ is completely determined by how everyday people use ‘woman’, and I also think that whichever property ‘woman’ expresses is WOMAN. If I thought that, then I could make an argument like this:

1. Everyday people use ‘woman’ to talk about a property, $F$.
2. ‘Woman’ expresses $F$.
3. Whichever property ‘woman’ expresses is WOMAN.
4. Therefore, $F = WOMAN$.

Arguments like this sometimes appear in the philosophy of gender literature. For example, Alex Byrne (2020) has recently made an argument with roughly this structure. Byrne argues for a view he calls AHF, according to which “S is a woman if and only if S is an adult human female” (Byrne 2020,
p. 2). Byrne’s paper has received a lot of attention for its highly controversial thesis, which is explicitly trans exclusionary. I discuss his view briefly here to point out that, whatever one thinks about the moral and political tenor of his view, his arguments are constructed on a substantive assumption, one which, as far as I can tell, finds no argument in his paper. This assumption is premise (3): whichever property ‘woman’ expresses is \textit{WOMAN}.

Byrne is right, of course, that, granting certain disquotational principles, if ‘woman’ means “adult human female”, S is a woman iff S is an adult human female. But Byrne effectively talks past other scholars in the philosophy of gender in at least two important ways: first, he assumes that other scholars are concerned with capturing how the majority of English speakers use ‘woman’. But many gender scholars do not think that the meaning of ‘woman’ is determined fully by use.

Second, and more importantly for our purposes here, Byrne is interested in a different property than other scholars in the gender literature. Byrne is interested in the property expressed by his usage-defined sense of ‘woman’, but that is not the property that other scholars in the literature are concerned with. For example, as we’ve seen, Sally Haslanger (2012) is interested in whichever property plays a specific role in metaphysically explaining certain oppressive social structures. As such, in order for Byrne to correctly claim that his view has succeeded where Haslanger’s has failed, Byrne would need to argue either that (1) the property he has identified plays the metaphysical role that Haslanger is interested in (and

---

4 Where the token ‘woman’ immediately to the left of the biconditional is used in the ordinary English sense.
5 Dembroff (2020) discusses this point in more depth.
hers does not), or else that (2) Haslanger shouldn’t be interested in a property that plays that metaphysical role, she should be interested in a different one.⁶

In order to successfully argue for (1), Byrne would need to show that the property of being an adult human female plays the metaphysical role Haslanger has identified, and also that her property fails to play that role. In order to successfully argue for (2), Byrne would need to argue against Haslanger’s semantic externalism, and show that in fact, ‘woman’ does not express Haslanger’s property. He does neither of these two things, it seems, because he doesn’t think he needs to. And this is because he assumes that ‘woman’ expresses \textit{WOMAN}.

But this is a hefty assumption. To see this, consider a central difference in the approaches of Byrne and Haslanger: where Byrne’s approach to gender metaphysics foregrounds the meaning of ‘woman’ and works from there to draw certain metaphysical conclusions about \textit{WOMAN}, Haslanger’s view foregrounds the metaphysical role played by \textit{WOMAN}, and effectively works in the reverse direction—using the resources of semantic externalism, she argues that ‘woman’ means \textit{WOMAN}. Note that Haslanger’s commitment to semantic externalism allows her to effectively bridge the “gap” between metaphysics (\textit{WOMAN}) and

⁶ As far as I understand, Byrne does think that his view has succeeded where Haslanger’s has failed, at least where he reads Haslanger as engaged in a descriptive project (Byrne 2020, p. 4 fn8). Quotations like the following indicate that Byrne takes himself not just to be arguing that ‘woman’ expresses the property of being a woman (which follows from disquotational principles (Byrne 2020, p. 2)), but rather that by figuring out what property ‘woman’ expresses, we can also settle which property is \textit{WOMAN}. Byrne writes:

A defender of the view that \textit{woman} is a social category needs to explain how the lexicographers could have erred so badly, but this issue is never even raised. On closer examination, appealing to dictionaries is not so bad after all. They are hard to reconcile with the idea that \textit{woman} is a social category. (Byrne 2020, p.5)

The argument Byrne makes in this quote only has a chance of succeeding if he assumes that ‘woman’ (as it is defined in the dictionary) expresses a property identical to \textit{WOMAN}, the target of theoretical analysis in the philosophy of gender.
language (‘woman’). Byrne needs something similar in order to secure his conclusion that ‘woman’ means WOMAN. But what, for Byrne, plays the role that semantic externalism plays for Haslanger? It’s hard to tell, and this, I think, is because he just assumes that ‘woman’ expresses WOMAN.

To be clear, I have not here argued that Byrne’s assumption is incorrect or wrong in some sense. Indeed, as I mentioned previously, many people in the philosophy of gender literature make this assumption. Some of them, like Haslanger (2012), defend it; others do not. For example, in the passage quoted below, Elizabeth Spelman (1988) appears to be making a similar mistake:

[If] we examine the use of “woman” in particular contexts, then we might be encouraged to ask when descriptions of what-it-is-to-be-a-woman really are descriptions of what-it-is-to-be-a-woman-in-culture-X or subculture-Y. Being a woman, as we surely know by now from cross-cultural studies, is something that is constructed by societies and differs from one society to another. (Spelman, 1988, p. 36)

Byrne (2020) and others have used this assumption (that ‘woman’ expresses WOMAN) to purchase very substantive metaphysical conclusions about the nature of womanhood. This assumption needs an argument, however, and acknowledging the difference between the two questions posed in the present paper helps us see that.
§4 ‘Woman’

In order to show that our question about ‘woman’ is distinct from our question about \textit{WOMAN}, in these next two sections I will show that it is possible to address one question without addressing the other. In this section, I’ll do this by considering an argument by Elizabeth Barnes (2019) and framing it within the present discussion. Barnes’ argument is one to the conclusion that the so-called \textit{exclusion problem} in the philosophy of gender is a problem about language, and (perhaps surprisingly) not a problem about metaphysics. If Barnes’ argument succeeds, then it provides us with one example of how the question \textit{What does ‘woman’ mean?} might be treated without engaging the question \textit{Which property (if any) is WOMAN?}, which is precisely what we’re after.

\textit{What does ‘woman’ mean?} ‘Woman’, as it appears here, is a predicate and predicates express properties. So, put a bit more precisely, our question is: which property does ‘woman’ express? We’ll start here with a general overview of the literature: contemporary work in the philosophy of gender offers a number of candidate gender properties which we might look to in answering questions about the meaning of ‘woman’. For example, according to some accounts, an individual is a woman just in case she occupies a certain social position in society (Ásta 2013, 2018; Haslanger 2012; Witt 2011). These are often called \textit{social position} accounts of gender. According to other accounts, an individual is a woman just in case she identifies in a particular way (Bettcher 2009, 2013; McKitrick 2015). These are often called \textit{self-identification} accounts of gender. There are additional accounts which incorporate a social position component as well as a self-identification component (Jenkins 2016). And there are those accounts which identify gender with biology (Byrne 2020).

However, as answers to the question \textit{What does ‘woman’ mean?} each of the available accounts runs into a similar difficulty. This difficulty sometimes goes by the name of the \textit{exclusion problem} (Barnes 2019; Jenkins 2016; McKitrick 2015). Generally speaking, an account of gender faces the exclusion problem just in case it does at least one of the following two things: (1) excludes some people from a
particular gender which it should not exclude, or (2) includes some people within a particular gender which it should not include.

As just one example of this, consider Sally Haslanger’s (2012) theory of gender. Haslanger is often understood as offering an answer to the question *What should ‘woman’ mean?* with the following account:

\[ S \text{ is a woman if and only if} \]

1. \( S \) is regularly and for the most part observed or imagined to have certain bodily features presumed to be evidence of a female’s biological role in reproduction;

2. that \( S \) has these features marks \( S \) within the dominant ideology of \( S \)’s society as someone who ought to occupy certain kinds of social positions that are in fact subordinate (and so motivates and justifies \( S \) occupying such a position); and

3. the fact that \( S \) satisfies (i) and (ii) plays a role in \( S \)’s systematic subordination, that is, along some dimension, \( S \)’s social position is oppressive, and \( S \)’s satisfying (i) and (ii) plays a role in that dimension of subordination. (Haslanger 2012, p. 234)

This is a social position account of gender. On this view, someone is a woman just in case other people see her in a certain way, and in virtue of seeing her that way, treat her in certain ways, and in so doing, subordinate her.

Haslanger presents this account within the framework of an ameliorative project. That is, here, Haslanger is not attempting to capture the ways in which we typically conceive of womanhood, or
respect the ways in which we typically use the term ‘woman’ in everyday discourse. Rather, she here proposes a new meaning of ‘woman’ which she argues we should adopt, given that doing so facilitates the achievement of certain pressing social and political ends, among which is the goal of ending gender oppression. Roughly, she argues that adopting the meaning she characterizes here will serve to draw our attention to certain harmful social practices in our society, and thereby facilitate their amelioration (Haslanger 2000, p. 47).

It is now common to think that Haslanger’s view faces the exclusion problem (Bettcher 2007; Jenkins 2016). For one thing, Haslanger’s account is problematically exclusionary. That is, Haslanger’s account fails to classify some prima facie women as women. For example, there are a number of contexts in which some transwomen are not observed or imagined to have certain bodily features presumed to be evidence of a female’s biological role in reproduction. Haslanger’s account implies that ‘woman’ should not apply to those transwomen in those contexts.

Haslanger’s account is also is problematically inclusionary. That is, Haslanger’s account classifies some prima facie non-women as women. For example, there are a number of contexts in which some transmen are observed or imagined to have certain bodily features presumed to be evidence of a female’s biological role in reproduction. Haslanger’s account implies that ‘woman’ should apply to those transmen in those contexts.

---

7 Of course, her view is a bit more nuanced than this: Haslanger is a semantic externalist, and (as was discussed earlier) she thinks that ‘woman’ actually does express the property she’s characterized with conditions (i)-(iii). It’s just that we don’t realize it expresses that property. So, in a sense, ‘woman’ already does mean what Haslanger says it “should” mean—I think it’s best to interpret her as saying that we should realize that, for externalist reasons, ‘woman’ actually has this surprising meaning. If this reading of Haslanger is correct, then she is also correctly understood as offering us an answer to the question What does ‘woman’ mean?
Recently, Elizabeth Barnes (2019) has argued that we should see the exclusion problem as a problem for how we use our gender terms, not as a problem for gender metaphysics. According to Barnes,

[the] problems are, instead, primarily ones of labeling. The issue isn’t inadequacies of social position accounts [like Haslanger’s] per se, but rather with their inability to serve as proxies for or definitions of our gender terms. (Barnes, 2019, p. 19)

To see this, consider the property of having blue hair, and call that property ‘B’. Now consider a scenario in which no transwoman in the world has B, and every ciswoman in the world has B. Note that at this point it wouldn’t make much sense to talk about transwomen being excluded from having B. The issue comes in when we give B as an answer to the question What does ‘woman’ mean?. For, this view has the consequence that ‘woman’ only applies to ciswomen, and not to transwomen. In other words, the issue here is not one with B itself, but one with whether ‘woman’ should apply to all and only the people who have B.

Similarly, with her account of womanhood, Haslanger is describing a property— a property one has just in case one is subordinated in certain ways for certain reasons. Call that property ‘W’. If we were going to object to the metaphysics of Haslanger’s account, we might do that by arguing that W is very disjunctive, or very unnatural, and no property can be that disjunctive or unnatural. Or we might object to something else about W itself. But if W is indeed a property which it is possible for people to have, then it is true of each of us either that we have W, or that we do not. However, the issue here is not one with W itself, but one with whether ‘woman’ should apply to all and only the people who have W.

This brings out that the exclusion problem for Haslanger is not one that targets the property she characterizes as being wrongfully exclusionary, but rather that property as an answer to the question
What should ‘woman’ mean? For as an answer to that question, her account has the result that ‘woman’ should not apply to some transwomen, and ‘woman’ should apply to some transmen, a result which many find unacceptable.

An important upshot of this discussion is that the exclusion problem should be understood as a problem about language, not a problem about metaphysics. Barnes seems to take this to imply that questions about what ‘woman’ in fact means, and what ‘woman’ should mean, are questions primarily about language, and accordingly we shouldn’t look to metaphysics for the answers. And this is in part because she thinks metaphysics doesn’t have the answers to these questions.

In Barnes’ view, the task of gender metaphysics is to say what, if anything, metaphysically explains gender. But that task can and does come apart from the task of providing application conditions for our gender terms. Questions about who is a man and who is a woman are questions about language which cannot be fully answered by metaphysics, because on Barnes’ picture “there aren’t any deep, language-independent facts about which people are women, which people are genderqueer, etc.” (Barnes, 2019, p. 22).

But what about our metaphysics question? \((Which\ property\ (if\ any)\ is\ WOMAN?)\) Does that question make sense to ask, on Barnes’ view? It depends on how we read the passage quoted above: does Barnes mean that there are no deep, language-independent facts about which people are classified under ‘woman’? Or does Barnes mean that there is no property WOMAN? If Barnes meant the first thing, then it would seem there is still space to ask our question, \(Which\ property\ (if\ any)\ is\ WOMAN?\). But if Barnes meant the second thing, then the answer to this question is just this: there is no property WOMAN.
I agree with Barnes that the exclusion problem is a problem of labeling, and questions about the extension of ‘woman’ are questions in the domain of language in at least one important sense. I depart from Barnes, however, with my interest in this further question about metaphysics: *Which property (if any) is woman?*. I’ll spend some time with this question in the next section.

§5 WOMAN

In this section, I will work through two examples, each of which emphasizes the distinction between the question *Which property (if any) is woman?* and the question *What does ‘woman’ mean?.* Again, the strategy in this section will be to show that it is possible to address the *WOMAN* question without addressing the ‘woman’ question. The first example comes from debates about naturalism in metaethics: there, the basic idea will be that, just as metaethicists have distinguished between questions about the term ‘natural fact’ and questions about the nature of natural facts, so too can gender metaphysicians distinguish questions about terms like ‘woman’ from questions about properties like *WOMAN* (§5.1). The second example comes from the literature on the metaphysics of gender: there I will consider one prominent view— Ásta’s (2018) view of gender properties as conferred properties— and show how, on this view, it is possible to investigate the nature of gender properties without first addressing the question of what ‘woman’ means (§5.2).

5.1 Naturalism and Nonnaturalism in Metaethics

_Are normative facts natural facts?_ Historically, the answer to this question has been hotly debated in metaethics. Very generally speaking, naturalists in this debate think that normative facts are natural facts, or at the very least normative facts are ultimately grounded in natural facts. Nonnaturalists do not think that normative facts are natural facts, and many think that normative facts are not ultimately grounded in natural facts. Of course, your view here will depend in part on what you think it takes for a fact to be normative, and in part on what you think it takes for a fact to be natural (Parfit 2011; Schafer-Landau 2003).
There are at least two ways to understand the question *Is F a natural fact?*. On the first reading, this is a question about whether *F* is in the extension of the term ‘natural fact’. On the second reading, this is a question about what *F* itself is like: what features does *F* share in common the fact that murder is wrong? What features does *F* share in common the fact that water is H₂O? Which facts is *F* grounded in? What are those grounding facts like? The metaethics literature recognizes both readings of this question. This means that sometimes, the question *Is fact F a natural fact?* is treated as a question about the metaphysical status of certain normative properties and facts, and not as a question about the meanings of the terms ‘normative fact’ and ‘natural fact’ (Parfit 2011; Lenman & Lutz 2018; McPherson 2012).

My contention here is that, just as the question *Is fact F a natural fact?* is a question that it makes sense to ask about facts in metaethics, so too does it make sense to ask of some property *P*, *Is P identical to WOMAN?*. Just as in the metaethics case, *Is P identical to WOMAN?* is not a question about the meaning of ‘woman’. It is not a question about which property or properties ‘woman’ expresses. It is a question about *P*. In asking this question, we might pose related follow-up questions, e.g. which conditions does someone need to satisfy in order to have *P*? Why think those conditions are conditions on having the property *WOMAN*? Notice that a version of this question can be asked of gender properties generally. Consider a higher-order property, the property of being a gender property, and call that higher-order property ‘*GENDER*’. Then we might ask of some property *P*, *is P identical to GENDER?*. This is not a question about the meaning of the term ‘gender’; it is a question about the nature of a property.

This analogy with discussions in metaethics shows us that questions like *Is P identical to WOMAN?* and *Is P identical to GENDER?* differ from the questions *Does ‘woman’ express P?* and *Does ‘gender’ express P?* This much is hopefully uncontroversial, and this much is all I need in order to establish the distinction between our questions *What does ‘woman’ mean?* and *Which property (if any) is WOMAN?*. 
The more controversial move here is this: posing the question *Which property (if any) is WOMAN?* raises the idea that this property, *WOMAN,* exists. Of course, that is not what I am claiming here. I am not making an argument for the existence of *WOMAN.* Instead, I am proposing that the question *Which property (if any) is WOMAN?* (1) is different from the question *What does ‘woman’ mean?*, and (2) is a question worth asking.

And indeed, some feminist metaphysicians have asked it (although it is not clear that they have been understood as having asked it). In the next section, I'll consider one serious answer to the question *Which property (if any) is WOMAN?* My hope is that this discussion will shed some more light on the distinction between questions about gender terms, and questions about gender.

5.2 Conferred Gender Properties

Consider Ásta’s (2013, 2018) account of gender. On Ásta’s account, gender properties are socially constructed properties, and socially constructed properties are *conferred properties.* Ásta distinguishes between institutionally conferred properties and communally conferred properties, but here I’ll be working with communally conferred properties because that process is a bit more straight-forward.⁸

(Communally) *Conferred properties are so-called because they are conferred onto individuals by other individuals in a particular community. Other people can confer a property onto you by treating you in certain ways, where this treatment socially constrains and/or enables your activity. Importantly, the individual with the conferred property is treated in these ways because they are taken to have some further property (called the *base property*). For this reason, Ásta describes*  

---

⁸ A property is *communally conferred* if it is conferred by a person, entity, or group with social standing or authority (note: this need not be institutionally recognized authority, although it can be). These sorts of properties can be either implicitly or explicitly conferred by the attitudes and/or behavior of the individual(s) executing the conferral. (Ásta, 2018, p. 21-22)
conferred properties as the *social significance* of their base properties. Gender properties are conferred properties, which means that they are the social significance of further base properties.

Let’s work with a simple example: in a particular context $C$, your perceived biological sex might matter a lot. Because other people in $C$ think you have a particular biological sex, they may behave toward you in certain ways, and in so doing, they may constrain and/or enable your activity in $C$. By constraining and enabling you in these ways, other people in your context confer onto you another property: a gender property, WOMAN. That is, this constraining and enabling on account of being perceived to have a particular biological sex *grounds* your having a further property, WOMAN. (Ásta 2018, p. 44, 74).

For Ásta, ‘WOMAN’ names a great many properties, where which properties ‘WOMAN’ names is context-dependent. But more on that in just a bit. What is important here is that, on Ásta’s view, a property is not identical to WOMAN merely because it is a property expressed by ‘woman’. She writes:

> What . . . is going on when there are disputes over what it is to be a woman and whether a person is a woman? On my view, disputes over what it is to be a woman are disputes over what feature should be the base property for the conferral of the institutional or communal status woman in a particular context. (Ásta 2018, p. 90)

Here, “the communal status woman” is not a communal status which is such that one has it just because one is the referent of ‘woman’. We might say that, for Ásta, our technical predicate ‘WOMAN’ names a property which one has in a context iff they have a certain communal status in that context. We then have to do further work to figure out what our ordinary predicate ‘woman’ means in that
context. In other words, Ásta’s analysis of gender properties recognizes a distinction between ‘woman’ and *WOMAN*.

How are ‘woman’ and *WOMAN* related on Ásta’s view? There is a deflationist answer to this question: roughly, someone has *WOMAN* in a context $C$ just in case they are in the extension of ‘woman’ in $C$. But there is another way to read Ásta here, one according to which it takes more than being in the extension of ‘woman’ for a property to be identical to *WOMAN*. To see this, consider the sense in which Ásta is a contextualist—

In the philosophy of gender, semantic contextualism is quite popular (Barnes 2019; Bettcher 2013; Saul 2012). Semantic contextualists about the meanings of gender terms like ‘woman’ typically accept that (1) the meaning of ‘woman’ may vary across contexts, and (2) in many contexts, ‘woman’ may have more than one meaning. Ásta is a contextualist, but she does not explicitly argue for a form of semantic contextualism. Instead, she is what we might call a *metaphysical* contextualist about gender properties. ⁹ Ásta is a contextualist in this sense because she thinks that (1) ‘*WOMAN*’ names many different properties, and (2) whether someone has one of those properties will vary across contexts (Ásta 2018, p. 75, 117, 123).

For example, in one context, people might attach a lot of social significance to having a certain body type. Call the property which is communally conferred in this context ‘*WOMAN*₁’. In this context, you can only have the property *WOMAN*₁ if you have that body type (that is, people in this context only confer *WOMAN*₁ onto individuals with that body type). But in another context, people might not care at all about body type, and instead attach a lot of social significance to the way individuals are dressed. Call the property which is communally conferred in this context ‘*WOMAN*₂’. There, you can

---

⁹ Ásta does say that the metaphysical contextualist view she articulates can be used to support a semantic contextualist view of gender terms (Ásta 2018, p. 110, fn. 19).
only have the property $WOMAN_2$ if you dress in a certain way (that is, people in this context only confer $WOMAN_2$ onto individuals who dress that way). And importantly, the only property you can get in the first context is $WOMAN_1$ (and so you cannot get the property $WOMAN_2$ in that context), and the only property you can get in the second context is $WOMAN_2$ (and so you cannot get the property $WOMAN_1$ in that context). That is the sense in which Ásta is a metaphysical contextualist.

Now consider some questions that Ásta’s contextualist view raises: Why does she single out the particular set of communal statuses she does when discussing womanhood, rather than an alternative set? What do the communal statuses that she does select share in common which makes them appropriate targets of her analysis? These are questions about the properties Ásta chooses to discuss, they are not questions about what she means by ‘woman’. We want to know why she has chosen the properties she has and what those properties are like.

If Ásta and I take a plane from California to Iceland, and once we arrive I ask her, “Which property in this context is identical to $WOMAN$?”, she might give a deflationist response, and say “Well, it’s just whichever property is expressed by ‘woman’ in this context.” In that case, she would tell me something about the relationship between a predicate and a property. But she might also say “Well, it’s the one you have just in case you have communal status X”, where X is a particular communal status in that context. In this second case, Ásta would tell me about a property, without telling me anything at all about ‘woman’.\(^\text{10}\)

### §6 Gender Properties and Gender Classification

Recent work in the philosophy of gender has encouraged us to focus less on metaphysical questions about who has the property $WOMAN$ ($MAN$, $GENDERQUEER$, etc.) and focus more on

\(^{10}\) Supposing, of course, that having that communal status X does not essentially involve being called a woman.
normative and linguistic questions about gender classification, questions like: Who should be in the extension of ‘woman’? Which normative considerations bear on that question? Which way of assigning meanings to gender terms would produce the best social and political outcomes?

In closing, I want to emphasize that the distinction I’ve argued for in this paper is complementary to these projects in the philosophy of gender. That is, the idea that there are two questions to be asked here—a question about ‘woman’ and a question about \textit{WOMAN}—is not at odds with the claim that the linguistic question is deserving of our attention when it comes to revising the rules of gender classification or policy making. Rather, this difference between ‘woman’ and \textit{WOMAN} is actually one that serves everyone well, not just people who are interested in the metaphysics of \textit{WOMAN}.

To see this, consider one recent project in this area: Robin Dembroff (forthcoming) has argued against something they call \textit{The Real Gender Assumption}, or the idea that “our gender classifications ought to track the operative gender kind membership facts” (Dembroff, forthcoming, p. 8). According to the Real Gender Assumption, “someone should be classified as a man only if they ‘really are’ a man—that is, only if \textit{man} is a recognized gender, and they meet its membership conditions” (Dembroff, forthcoming, p. 2). Dembroff argues that the Real Gender Assumption should not guide our classification practices, because the gender kinds that operate in dominant contexts are oppressive, at least insofar as they marginalize many trans and nonbinary people by rendering their identities unintelligible, and by denying them membership in gender kinds like \textit{woman} and \textit{man}.\footnote{I have been using the all-caps \textit{WOMAN} as a name for a gender property, but in keeping with Dembroff’s focus on kinds rather than properties, I use the lowercase italicized \textit{woman} here to refer to a gender kind.}

One might think that, by arguing for a distinction between \textit{WOMAN} and ‘woman’, I am thereby endorsing the Real Gender Assumption. However, by making a case for asking the question \textit{Which...}
property (if any) is \textit{WOMAN}? in the present essay, I am not thereby making a case for this assumption. The Real Gender Assumption is an assumption about how people should be classified. Which gender labels should go with which kinds? Who is in the extension of the term ‘woman’ in this context? These are questions about classification. On the other hand, the question \textit{Which property (if any) is WOMAN?} is a question about properties and what they are like. Someone might have the property WOMAN and yet not be classified under the label ‘woman’. You might think, for example, that this is exactly what happens to some transwomen in dominant contexts: they are women (because they have the property WOMAN) but they are not classified as such.

Note that Dembroff and I agree on this last point. That is, Dembroff thinks that someone can be in a particular gender kind (i.e., have a particular gender property), and yet fail to be classified as such in their present context. In their discussion of modest ontological pluralism (the view that we are each a member of a variety of different gender kinds, but the social relevance of these memberships changes across contexts) Dembroff considers the following example:

Suppose that someone is a woman relative to dominant gender kinds, but a man relative to trans-inclusive gender kinds. According to modest ontological pluralism, this person has both of these gender kind memberships in all contexts. For this reason he can truthfully say, “I am a man.” This is because the term ‘man’ in his claim refers to the trans-inclusive gender kind -- a kind he retains membership in even when in contexts where this kind is not operative. (Dembroff, forthcoming, p. 19)

In this example, someone is a member of a gender kind despite not being classified as a member of that kind in their present context. Dembroff writes in terms of kinds whereas I prefer to write in terms of properties, but for the purposes of our discussion here, the terms are interchangeable. The person in Dembroff’s example has the property \textit{MAN}, despite the fact that they are not in the
extension of ‘man’ in their present context. That is, in this example, the person speaking is a man

because they are a member of a trans-inclusive gender kind, where that trans-inclusive gender kind is the kind

man.

What makes it the kind man? Here are two plausible answers: (1) the kind in question is the kind man

because it is the referent of the term ‘man’ in at least one (trans-inclusive) context. Alternatively, (2)

the kind in question is the kind man because something about the kind itself makes it that way.

Dembroff’s discussion later in their paper suggests to me that they would say something like (1), but

their project is also consistent with (2). Regardless, both ways of answering this question

acknowledge a distinction between gender kinds and classification under a gender term.

I read Dembroff as arguing against the view that we should look to whichever gender properties are

salient in our context when we’re trying to figure out who gets access to personally and politically

significant terms like ‘man’, ‘woman’, and ‘genderqueer’. I do not read them as arguing against the

idea that there are gender properties, properties like WOMAN, MAN, and GENDERQUEER. I hope it is

now clear that our projects complement one another. Dembroff is relying on a distinction between

membership in a gender kind and being in the extension of a gender term. That is precisely the

distinction I’m making in the present essay. Having a gender property (e.g., WOMAN) is one thing;

being classified under the label ‘woman’ is another.

§7 Conclusion

What is it to be a woman? The central contention of this essay has been that there are

actually two closely related, but nevertheless distinct questions in this vicinity. The first is a question

about language: what does ‘woman’ mean? The second is a question about metaphysics: which

property (if any) is WOMAN? I have not argued in favor of treating the question What is it to be a
woman? as a question about language rather than a question about metaphysics (or vice versa). I have just pointed out that we have a choice between those two options.

Acknowledgements

My deepest thanks go out to Karen Bennett, Esa Díaz-León, Katharine Jenkins, Andrew Rubner, Jonathan Schaffer, Ted Sider, Isaac Wilhelm, and the participants of the 2020 Early Career Metaphysics Workshop for their many helpful comments and discussion on this essay.
BIBLIOGRAPHY


Mason, R. (2020/manuscript). *Social Ontology Naturalized*.


